RESEARCH

What Works for Children with Literacy Difficulties?

The Effectiveness of Intervention Scheme

Greg Brooks
University of Sheffield

department for
education and skills
creating opportunity, releasing potential, achieving excellence

Research Report RR380

What Works for Children with Literacy Difficulties?

The Effectiveness of Intervention Schemes

**Greg Brooks
University of Sheffield**

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS OF INFORMATION PROVIDED FOR THE SECOND EDITION

These acknowledgments include many for information which, sadly, could not be used.

Local Education Authorities

Bradford	Derbyshire	Norfolk
Bolton	Devon	Redcar and Cleveland
Bridgend	Co. Durham	Sefton
Bristol	Hampshire	Wakefield
Cornwall	Haringey	Worcestershire
Cumbria	Knowsley	

Individuals

Where individuals' affiliations are not shown in the left-hand column, their affiliations are with the LEA mentioned in the right-hand column.

Pauline Bentote	SIDNEY in Hampshire
Mel Blackburn	Stockport's Wave Three scheme
Mary Charlton, consultant	Tracks Literacy
Alan Davies, Director of THRASS	THRASS
Sue Derrington	Phono-Graphix™ in Bristol
Dr Angela Fawcett, University of Sheffield	Interactive Teaching and Assessment, Reader's Intelligent Teaching Assistant, three reviews on D*f*ES website
Ann Foster	Better Reading Partnerships in Co. Durham
Pam Fudge	Reading Recovery in Bristol
Dr Peter Hatcher, University of York	Reading Intervention
Carol Greenway	Reciprocal Teaching inWestminster
Peter Hayes, Director, CSV Education for Citizenship	Reading Together
Mike Johnson, Manchester Metropolitan University	Multi-sensory Teaching System for Reading
Bruni Jones & Eileen O'Dowd	Underattainers Can Achieve!, Great Crosby Catholic Primary School, Sefton
Brian Kermode, consultant to McGraw-Hill Publishers, UK	Reading Mastery Fast Cycle
Angie Kotler	Talking Partners in Bradford
Julie Lawes, Project Director, The Catch Up Project	The Catch Up Project in Cornwall, Norfolk and Wakefield, and the National experimental study
Prof. Ann Lewis, University of Birmingham	Use of ILS with children with literacy difficulties
Kerry Longhorn	Family Literacy in Hampshire
Anthea Main	Better Reading Partnerships in Worcestershire
Dr Irene Mersh	Tracks Literacy in Wokingham
Dr Martin Miles, University of Exeter	AcceleRead AcceleWrite in Devon

Jim McGowan	City Reading Intervention Scheme (CRIS) in Nottingham
Dr Martin McPhillips, Queen's University of Belfast	His Four-Movement Programme
Dr Penny Munn and Sue Ellis, University of Strathclyde	Longitudinal evaluation of Reading Recovery in Northern Ireland
John Mursell	SIDNEY in Poole
Linda Perry	Blitz! in St Helens
Dave Philpot	Phono-Graphix™ in Wigan
Carol Price, consultant	Blitz! nationally and in East Sussex
John Rack, Dyslexia Institute, York	SPELLIT
Christa Rippon	Reciprocal Teaching in Haringey
Dr Jonathan Solity, University of Warwick	The Early Reading Research
Heather Swain	Reading Intervention in Cumbria
Andrew Taylor	Better Reading Partnerships in Redcar and Cleveland
Carol Taylor	Better Reading Partnerships in Derbyshire
Prof. Keith Topping, University of Dundee	Cued Spelling, Paired Reading, Paired Thinking, Paired Writing
Trudy Wainwright	South Gloucestershire's scheme promising 'a flying start with phonics'
Jacqui Worsley	The Catch Up Project in Norfolk
Prof. David Wray, University of Warwick	Lewisham Literacy 2000

CHAPTER ONE
FOCUS AND INTENTION OF THIS REPORT

1.1   The focus

Most children learn to read and write satisfactorily through high-quality classroom teaching, but what of those who don't? How are they to be helped? This research report reviews early intervention schemes that have been devised to help struggling readers and writers, and is intended to inform schools' choices among such schemes.

More exactly, the questions this research report is addressing are:

- **What intervention schemes are there which have been used in the UK in an attempt to boost the reading, spelling or overall writing attainment of lower-achieving pupils in at least one of Years 1-6, and which have been quantitatively evaluated here?**

- **What are those schemes like, and how effective are they?**

The restriction to schemes used and evaluated in the UK is partly intended to avoid a deluge of information on schemes used elsewhere in the world, but mainly to circumvent the objection 'How do we know that it will work here?'

The intention is to make clear and analytic information on such schemes available in order to inform practice and choices of approach. Those choices should be guided not only by the evidence assembled and analysed here, but also by careful matching of the needs of an individual school or even class to the specifics of particular schemes.

Most of the schemes covered in this report are 'Wave Three' initiatives within the current structure of the National Literacy Strategy in England, as defined on the following page, though some are also in use as Wave Two initiatives, and a few are 'preventive' measures, aimed at preventing young children developing difficulties in the first place.

Within that structure, there is an obvious need for schools to have clear information, in order to make principled decisions about which approach to adopt for children who experience difficulties in literacy.

The National Literacy Strategy's Three Waves

The NLS assumes three 'waves' of support for pupils, as described below. The SEN Code of Practice (2001) describes a 'graduated response' to identifying and meeting special educational needs which may be mapped onto the NLS's three waves as shown.

Wave One:

The effective inclusion of all pupils in a high quality Literacy Hour (Quality First Teaching)

Pupils may be at any point on the 'graduated response' – i.e. the usual differentiated curriculum, School Action or School Action Plus

Wave Two:

Small group intervention (Early Literacy Support, Additional Literacy Support, Further Literacy Support, intervention programmes, Booster classes, equivalent LEA- or school-based programmes) for pupils who can be expected to 'catch up' with their peers as a result of the intervention – that is, who do not have special educational needs related specifically to learning difficulties in literacy or mathematics.

Wave Two interventions are not primarily SEN interventions. Where intervention programmes are delivered without modification within the designated year group, there is no requirement that the pupils involved should be placed on School Action.

Pupils included in Wave Two interventions may on occasion already be at School Action or School Action Plus, or have a statement of special educational needs. This may be where they have special educational needs such as emotional and behavioural difficulties, communication and interaction difficulties, or sensory or physical impairment, for which they are receiving other forms of intervention.

Wave Three: Specific targeted intervention for pupils identified as requiring SEN support.

Pupils at Wave Three may have particular needs related specifically to literacy, or needs associated with other barriers to their learning.

Provision at Wave Three is likely to draw on specialist advice. It may involve the adjustment of learning objectives and teaching styles, and/or individual support. It aims to reduce gaps in attainment and facilitate greater access to Waves One or Two.

Pupils receiving Wave Three support will always be placed on School Action, and on School Action Plus if an external agency is involved in assessment, planning and review.

1.2   The need

What proportion of children experience literacy difficulties? An estimate (for England only, since the D*f*ES's remit is only for England) can be based on the results of National Curriculum assessments. Table 1 gives the percentages of children not yet achieving level 2 in reading at the end of Key Stage 1 (age 7), or not yet achieving level 3 in English at Key Stage 2 (age 11), in 1998-2002.

Table 1:   Percentage of children in England achieving below level 2 in Reading in Key Stage 1 National Curriculum tests, or below level 3 in English at Key Stage 2, 1998-2002 *

Year	Key Stage 1 Percentage	Key Stage 2 Percentage
1998	19%	7%
1999	17%	7%
2000	16%	6%
2001	16%	7%
2002	15% (provisional)	7% (provisional)

* excluding absent pupils

This shows that significant numbers of children experience literacy difficulties and are likely to have difficulty in coping with the steadily increasing demands of the curriculum in Key Stage 3 (and beyond).

So what can be done for these children? We have identified 25 studies which are useful in attempting to answer this question.

1.3   The schemes covered

The titles of the 25 studies we have analysed, and the shorter names by which they are mainly referred to in this report, are as shown in Table 2. Several of the studies contained evaluations of more than one scheme, so in order to show the full coverage of the report, relevant studies are shown with their 'alternative interventions' (that is, the other approaches with which the main ones were compared) listed in italics below their full title.   These alternative interventions bring the total of approaches evaluated up to nearly 40. In addition, many of the studies contained 'no intervention' ('ordinary classroom teaching') control groups, and these are also analysed here: studies with a well-defined no-intervention control group are marked with an asterisk.

Table 2: Full and abbreviated names, and outline structure, of the 25 studies

A. General schemes

1	* Better Reading Partnerships (in 5 LEAs)	BRP	

2	* The Catch Up Project
	    (pilot study, national study, and 3 LEAS)
	    Matched Time (within pilot study, national study)
		The Catch Up Project	

3	* Cued Spelling	Cued Spelling	

4	Family Literacy
	    (Basic Skills Agency's Demonstration Programmes,
	    Hampshire, Basic Skills Agency's New Groups initiative)
		Family Literacy	

5	* Individual Styles in Learning to Spell	Individual Spelling	

6	Inference Training
	    Comprehension exercises
	    Rapid decoding
		Inference Training	

7	* Interactive Assessment and Teaching	IA&T	
8	Knowsley Reading Project	Knowsley	
9	Multi-Sensory Teaching System for Reading *Beat Dyslexia*	MTSR	
10	* Paired Reading in Kirklees	Paired Reading	(1)
11	* Parental Involvement in Haringey *Extra reading*	Parental Involvement	(2)
12	Phono-Graphix™ (2 LEAs)	Phono-Graphix™	
13	* Phonological Awareness Training	PAT	

14	* Reading Intervention
	    (formerly Cumbria Reading with Phonology Project – where
	    this is meant it is called 'original')
	        (also in general use in Cumbria, and for statemented children)
	    Reading-only
	    Phonology-only
		Reading Intervention	

15	* Reading Recovery
	    (in London and Surrey; separately in Bristol)
	    Phonological Intervention
		Reading Recovery	(3)

16	Reciprocal Teaching	Reciprocal Teaching	

17	* Somerset Self-esteem and Reading Project
	    Self-esteem counselling only
	    Remedial phonics only
	    Remedial reading only
	    Drama plus DISTAR
	    DISTAR only
		Somerset	(4)

18	* SPELLIT	SPELLIT	
19	THRASS	THRASS	
20	* Time for Reading	Time for Reading	

B.   Schemes involving ICT

21   AcceleRead AcceleWrite (formerly Jersey Computer Assisted
      Reading Development Programme)   AcceleRead AcceleWrite
           (2 LEAs)

22   * Integrated Learning Systems   ILS

23   * Reader's Intelligent Teaching Assistant   RITA
      IA&T

C.   Schemes for improving writing

24   * Paired Writing   Paired Writing
           (3 studies)

25   Family Literacy   Family Literacy

Key: * = *scheme with well-defined no-intervention control group in at least one study – see Appendix and section 2.3*
        Approaches mentioned in italics under each main scheme are alternative interventions investigated within the same evaluation

Notes to Table 1:

(1)   Topping and Lindsay (1992) reviewed dozens of Paired Reading schemes from all over the English-speaking world. For this report, the Kirklees scheme, which was not only used in England but also by far the biggest of those reviewed by Topping and Lindsay, has been taken as representative of all the findings, and is referred to simply as 'Paired Reading'.

(2)   Similarly, since there have been many Parental Involvement schemes, the original and best-known, Haringey, has been taken as the exemplar for this report.

(3)   Reading Recovery has been the subject of nearly 20 evaluations in the UK, mostly by individual LEAs. However, by far the best reported is that carried out in six London LEAs and Surrey by researchers at the University of London Institute of Education. This is taken as the exemplar for this scheme in the UK (though one local evaluation is also analysed). Where Reading Recovery itself is meant, the title 'Reading Recovery' is used; but where it is necessary to refer to the control or Phonological Intervention conditions in the London Institute of Education study, the abbreviation RR is used instead.

(4)   Somerset was a series of four studies; where necessary these are distinguished by a number in brackets, e.g. Somerset (1).

1.4   Forms of data

In order to judge whether an initiative has really made a difference, it is not enough just to ask the participants – they will almost always say it has. This 'feel-good' factor is valid in its own terms, but doesn't always correlate with measured progress, and certainly doesn't convince policy-makers and funders.   So quantitative data on the learners' progress are essential, measured by appropriate tests of (in this case) reading, spelling or writing.

But not just any test data will do: if the test provides only raw scores, the average gain may look impressive, but what does it mean?   How good is it, compared with gains in other projects and/or with national norms?   We need some way of comparing the impacts of different initiatives. The two forms of impact measure used in this report are ratio gains and effect sizes. These are explained in more detail in the early part of the Appendix; briefly,

- a ratio gain is a group's average gain in reading or spelling age in months divided by the time between pre- and post-test in months. A ratio gain can only be calculated where the test provides reading or spelling ages;

- an effect size is the experimental group's gain minus the control group's gain divided by the control group's standard deviation. An effect size can be calculated whether the scores are reading/spelling ages, standardised scores, or even raw scores – there are examples of all three in the Appendix. An effect size can even be calculated in the absence of a control group, *provided that the test used yields standardised scores*. In these circumstances the standardisation sample is treated as an implicit control group and the standard deviation of the test is used (see, for example, Family Literacy (1) and (3) in the Appendix).

Both forms of impact measure make it possible to put different initiatives on the same scale, despite their having used different tests, and therefore to compare their effectiveness. Unfortunately, ratio gains and effect sizes can't be translated into each other, so in Tables A4-8 in the Appendix you will find some that list ratio gains and others that list effect sizes. A very few evaluations provide the information to calculate both forms of impact measure, but so few that the two scales cannot be correlated.

There are, of course, other forms of data out there, and a word of explanation is needed on why they have not been used. Basically, it's because they don't allow different initiatives' impacts to be put on a common scale and compared. This is not to deny the usefulness of these forms of data for their own purposes. For example, all Reading Recovery schemes calculate and report 'the percentage of children successfully discontinued', and this enables Reading Recovery providers to see how closely they are matching up to the high targets they are set. But this measure cannot be used to provide external comparisons. This is also true of another measure frequently reported now, namely how far children have progressed up the Bookbands scale. Attempts have been made to translate Bookband levels into reading ages, but the translation is in terms of very broad reading age bands (e.g. '5:00-6:06'), and these are too blunt to provide estimates of progress for which either statistical significances or impact measures can be calculated.

Similarly, some initiatives now report in terms of children's greater achievement at (say) Key Stage 1 than their teachers had predicted, and this is meaningful to providers within the

system that operates in England. But this measure provides no comparisons beyond England, or beyond the few schemes that have so far reported in these terms.

1.5   New features

The first edition of this report was published by the National Foundation for Educational Research in 1998 under the title *What Works for Slow Readers? The Effectiveness of Early Intervention Schemes* (Brooks *et al.*, 1998), and the field has moved on considerably even in four years. So in this edition a great deal of the information from the first edition has been retained, though some is no longer included, and a great deal more has been added.   The schemes which are no longer included are:

> Docklands Learning Acceleration Project
> Dyfed Improving Reading Standards in Primary Schools Project
> Leeds Sustained Reading Intervention
> Lewisham Literacy 2000
> Pause, Prompt and Praise
> Saint Lawrence School, Towcester, Northants
> Shropshire: Raising Attainment In Shropshire Education

Pause, Prompt and Praise was not included because the UK evidence on it was very slight (although there is strong evidence from outside the UK that it can be effective). The four LEA schemes (Dyfed, Leeds, Lewisham, Shropshire) were not included because they are no longer available. Docklands went out because it seemed rather ineffective – this also applied to the Shropshire scheme and, on re-analysis, to Lewisham Literacy 2000.

However, the scheme run at Saint Lawrence School, Towcester, Northants, should be singled out for special mention: in the first edition it was the only initiative run by a single school. It has been dropped only because too many demands might be made on the school.

Schemes primarily focused on reading which have been added in this edition are the following:

> Interactive Assessment and Teaching
> Knowsley Reading Project
> Multi-Sensory Teaching System for Reading
> Phono-Graphix™
> Reciprocal Teaching
> SPELLIT
> Time for Reading
> Reader's Intelligent Teaching Assistant

Schemes for spelling

The first edition was concerned solely with reading, and this still takes up the bulk of the report. However, one of the requirements for the revision was that it should consider what information was available on spelling and on writing more generally. Accordingly, there are two schemes which have spelling as their sole focus:

> Cued Spelling (2 studies)
> Individual Styles in Learning to Spell

and other data on spelling will be found in the entries for:

> The Catch Up Project (one of the initiatives in Norfolk)
> Interactive Assessment and Teaching
> Multi-Sensory Teaching System for Reading (2 studies)
> Phono-Graphix™ (in Bristol)
> Reading Intervention (both the original Cumbria Reading with Phonology Project
> > and two other studies within the county)
> Reading Recovery (the evaluation in London and Surrey)
> THRASS
> Time for Reading
> AcceleRead AcceleWrite (2 studies)
> Reader's Intelligent Teaching Assistant.

All the schemes which focus on reading and/or spelling are in parts A and B of chapter 3. Those which are not, or not primarily, computer-based are dealt with in Part A, while part B describes three schemes which are wholly computer-based. Some of those in part A have computerised as well as paper-based versions, for example THRASS.

Schemes for writing

The requirement to cover the general process of writing was much more difficult. Much less quantitative research is done on writing than on other aspects of literacy, and there is very little that stands up to scientific scrutiny. In the end, the new section on schemes for improving writing (part C of chapter 3) mainly consists of just three studies on the Paired Writing technique, plus some insights from less rigorous evaluations of Family Literacy and Reading Recovery. The studies on Paired Writing were all conducted by Keith Topping and colleagues at the University of Dundee. Their example should lead many other researchers to help fill this gap.

The way in which the information summarised in this report was analysed is described in the Appendix. The schemes are all described in chapter 3, but first a guide through them is provided in chapter 2.

SIGNPOSTS

2.1   Finding your way

Reading research is a jungle, and quantitative evaluations of early interventions are among its densest thickets. This chapter is intended to help you find the schemes which may be most relevant to your situation. All the schemes mentioned are described in chapter 3 – but remember that these are only the schemes analysed for this report.

Before going on to the remainder of the chapter it would be advisable to read the caveat in the next section on the scale of the various evaluations.

When reading this chapter **it is particularly important to remember that it mentions all schemes, whether effective or not;** however, as a rough guide, schemes we consider less effective are shown in [square brackets].

2.2   Scale of the evaluations

When considering the interventions and what is said about them here, it would be well to bear in mind that the evaluations differed vastly in scale. To emphasise this, Table 3 shows the numbers of children involved. Where more than project is covered by a heading, the various studies' numbers are summed.

The reason for the huge numbers for Paired Reading is explained in the description in chapter 3. Better Reading Partnerships and The Catch Up Project have been evaluated on a large scale in some areas since the first edition, when their numbers were modest. The numbers against some well-known names in the list are comparatively small – comments on this are in chapter 4. But there is no simple correlation between size and quality here – some small studies (in terms of number of children in the experimental group) were meticulously designed and reported, while the reporting of some with much larger numbers was considerably less full.

2.3   The impact of ordinary classroom teaching

As shown in Table 2, 13 of the 25 studies provided evidence on well-defined control groups who received no extra intervention, in other words ordinary classroom teaching. Several of the studies providing information on the impact of ordinary teaching were among the largest, and the total number of children in the relevant groups was at least 2500 (and probably considerably larger, since the control group numbers in Better Reading Partnerships in Bradford and [Integrated Learning Systems, Phase II] were not known).

Table 3:   Numbers of children involved in the studies, in decreasing order of number in experimental group

Ref	Name of study	Numbers of children			
		Experimental groups	Control groups	Alternative Intervention(s)	Total
1	Better Reading Partnerships	2897?	822?		3719?
10	Paired Reading in Kirklees	2372	446		2818
2	The Catch Up Project	1484	60	50	1604
22	Integrated Learning Systems	(953*)	(964*)		(1917*)
4, 25	Family Literacy	598			598
14	Reading Intervention	459	31	61	551
8	Knowsley	302			302
12	Phono-Graphix™	242			242
15	Reading Recovery	234	153	91	478
19	THRASS	160			160
17	Somerset	143	91	225	459
9	MTSR	109		17	126
7	IA&T	98	87		185
21	AcceleRead AcceleWrite	91			91
23	RITA	74	103		207
20	Time for Reading	68	72		140
24	Paired Writing	55	75		130
11	Parental Involvement	51	152	45	248
18	SPELLIT	51	58	41	150
5	Individual Spelling	36			36
3	Cued Spelling	32	10		42
6	Inference Training	26		26	52
13	PAT	24	24		48
16	Reciprocal Teaching	16			16

Key: ? = number not stated clearly or at all in a contributing report
 * = numbers for Integrated Learning Systems are for numeracy as well as literacy; literacy numbers were not given separately

Note: *Where no number is shown, there was no control or alternative intervention group.*

As the Appendix again shows, most control groups made normal progress. This finding is, however, circular: children receiving ordinary teaching mostly made the progress to be expected of children receiving ordinary teaching. What is more interesting is that some control groups made *better* than expected progress despite, apparently, receiving no extra intervention – see especially Paired Reading. What secret might this scheme have had? It seems that in Kirklees (the LEA where the Paired Reading study was conducted) the experimental intervention affected a significant proportion of schools.

So it may be that Paired Reading affected a high proportion of the schools in the area in which it took place, and the experimental schools were observed by others. This may have influenced non-participating schools to 'raise their game', and provide 'ordinary teaching' of a higher effectiveness than usual. Density of implementation seems not to have been a feature of schemes where the control groups made normal progress, and it might be reasonable to conclude that this is more like the normal situation, and therefore that ordinary teaching provides extra impact only in exceptional circumstances.

On the other hand, if the 'density' effect is real, it would support an argument for implementing initiatives at a fairly high density and/or with great publicity.

The evidence on ordinary teaching therefore proves the need for early intervention schemes: in general, ordinary teaching does not enable children with literacy difficulties to catch up.

2.4 Focusing on phonological skills for reading

Phonological skills, including spelling, were the focus of the largest number of studies. Among those analysed here the following mainly phonological schemes focused on reading:

- four main schemes: AcceleRead AcceleWrite, [PAT], Phono-Graphix™ and THRASS, and

- five alternative interventions [Phonology-only in Cumbria], Rapid decoding in Inference Training, [Phonological Intervention in the Reading Recovery in London and Surrey study], Phonics-only in Somerset (1) and DISTAR-only in Somerset (4).

Those shown in square brackets are relatively ineffective, while the rest are at least reasonably effective. So the overall evidence on the effectiveness of schemes which focused on phonological skills is mixed.

However, a generalisation can be drawn from the five schemes mentioned above which were alternative interventions within larger evaluations. In all of these except Somerset (1), the phonological scheme was substantially less effective than the main experimental approach; and the main approach was broader and incorporated work on phonological skills.

This description also fits Phono-Graphix™ and THRASS, which give explicit attention to grapheme-phoneme relationships within a broad framework, and were effective.

For greatest impact with struggling readers, therefore, work on phonological skills should be embedded within a broad approach.

2.5   Focusing on writing and spelling

There were too few studies on the compositional aspect of writing to justify broad conclusions, but Paired Writing has potential, and the approach adopted in Family Literacy seems to have been effective.

As pointed out in chapter 1, 12 schemes provided data on spelling, including two, Cued Spelling and Individual Spelling, which were concerned exclusively with spelling. Those two schemes and at least two others, Phono-Graphix™ and [THRASS – less effective for spelling than for reading] paid explicit attention to phoneme-grapheme relationships. However, it is less obvious for spelling than for reading that embedding within a broad framework is crucial, since Phono-Graphix™, [THRASS] and Cued Spelling did this, but Individual Spelling did not – rather, it encouraged teachers to find the precise approach that worked best for individual children.

Two other schemes were highly effective for spelling: IA&T and RITA. What these seem to have in common with the other effective schemes for spelling is that they are highly structured.

For greatest impact with children who struggle with spelling, highly structured schemes work best.

2.6   Focusing on comprehension skills

Most aspects of reading improvement are under-researched in the UK, but this is the most under-researched of all. There are innumerable studies of comprehension, in its pure cognitive and psychological-process aspects, and many quantitative studies on how to improve it have been conducted outside the UK. In the United States a powerful meta-analysis of rigorously conducted randomised controlled trials (Rosenshine and Meister, 1994) found a satisfactory effect size in favour of working on comprehension. But here there have been few quantitative studies of how to help children who can read accurately, in the sense of decoding fluently, but who appear not to understand much of what they read.

The only studies in this set which addressed this issue directly were Inference Training and Reciprocal Teaching. Inference Training included only 13 'less skilled comprehenders' and 13 'skilled comprehenders', but had a detailed and tightly organised experimental design. The Inference Training which was designed specifically for the experiment was effective for less skilled comprehenders, and more so than for skilled comprehenders. However, extra comprehension exercises were equally effective.

Reciprocal Teaching was even smaller – 16 children with no controls. However, its result was very clear – a strong impact on reading accuracy and an even stronger one on comprehension.

There is a small further insight from a project with a different main focus. Within Integrated Learning Systems, school A achieved good improvements in children's comprehension when only the comprehension section of the computer program was switched on - see section 2.8.

There appeared to be no tendency for schemes other than Inference Training to bring about greater improvements in children's comprehension than in their reading accuracy, or for the opposite to occur.

However, there were three other schemes which had substantial impacts on comprehension, namely Paired Reading, Phono-Graphix™ and THRASS. Again, these are highly structured schemes which seem to have targeted comprehension.

Bentley and Reid (1995, pp.21-22) have several useful suggestions for (less formally researched) methods of supporting struggling readers who are not 'reading for meaning'.

From the limited evidence available it can tentatively be deduced that children's comprehension skills are benefited most by being directly targeted, and not indirectly through work on reading accuracy.

2.7   Focusing on self-esteem

A series of four studies on this topic was carried out over a decade or more from 1970 to 1984 in Somerset - see pp.43-46. The results can be summarised as follows.

Self-esteem counselling by a professional educational psychologist was effective in raising reading attainment. But since this was too expensive an approach for general use, the researcher then trained non-professionals to deliver self-esteem counselling. When used alone, this was not effective, but when combined with a specific reading intervention it was very effective. And in the final study drama teaching designed to boost self-esteem plus a specific reading intervention was also very effective - and could be seen as even more cost-efficient.

Even though no further comparable studies seem to have been done for nearly 20 years, working on self-esteem and reading in parallel would seem to have definite potential.

2.8   Focusing on ICT

Results from the three studies using ICT as their main resource (AcceleRead AcceleWrite, [ILS], RITA) are again mixed. These studies differed significantly in scale. AcceleRead AcceleWrite's two studies involved only a few children in each of several year groups. At the other extreme, [Integrated Learning Systems] was a national evaluation led by the University of Leicester for the (then) National Council for Educational Technology, now the British Educational Communications and Technology Agency.

RITA did produce significant progress, but no more than the non-computer-based intervention to which it was compared, IA&T – hardly the result that advocates of the technology expect.

The main result of [ILS] was that its impact on reading was non-significant. The children in the experiments did make some progress, but no more than would be expected from normal schooling. And in one primary school where an Integrated Learning System had been targeted on children with SEN, the children in the project made significantly *less* progress than the controls (see NCET, 1996, pp. 19 and viii, school U). Even more generally, Ann Lewis's (1999) review of using ILS with children with low attainments in reading concluded that its effectiveness had not been demonstrated.

However, there were two contrary findings. Both of AcceleRead AcceleWrite's studies produced highly significant gains, the Devon study just for reading, the Jersey study for spelling also. Both studies were small, but what is striking about the approach is how precisely targeted it was. Children read and re-read a sentence from a card until they could type it into the (talking) computer from memory with high accuracy. Thus the approach stressed the accuracy of both reading and spelling.

And within the generally non-significant results from [ILS], there was one school (see NCET, 1996, pp. 19, 29 and ii, school A) where project pupils made three times as much progress in six months as the controls. It is worth quoting at length the researchers' description of how this was achieved *(op. cit.,* p.29):

> *The teachers ... found that pupils had completed Initial Reading without having mastered the comprehension strands. In some cases pupils had avoided comprehension completely but were still able to finish the module. Teachers decided to re-enrol pupils on Initial Reading, switching on only the comprehension strands and increasing their support of pupils. Pupils then moved on... better prepared to cope with the comprehension level demanded of them.*

Thus where the technology was used with precision and backed up by teachers, gains were made. In other circumstances, pupils were left to find their own routes and targets. This left them at risk of what Hurry (1996, p.26) has called 'the butterfly approach or the smorgasbord approach', either flitting unproductively from one item to another, or trying to digest too much all at once.

Given the financial investment that all ICT approaches require, technology used to boost literacy attainment deserves to be targeted as precisely as possible.

2.9   Large-scale programmes

Four of the schemes covered here merit this description: Family Literacy, Phono-Graphix™, Reading Intervention, and Reading Recovery. All were effective, but all are relatively expensive, since they require considerable training for teachers, and either a good deal of individual tuition for children, or recruitment of parents also onto the courses. But since US evidence suggests that every dollar spent on early intervention saves seven dollars on social remediation later (Schweinhart *et al.,* 1993), such schemes may well be good value.

Wright (1992) contrasted the one-off cost of Reading Recovery (then) in Surrey of £600 with the £15,000-£25,000 needed for a Statement of Special Educational Needs and resulting support over many years. Similarly, Hurry and Sylva (1998) suggested that, although Reading Recovery is expensive at the point of delivery, averaged out over a five-year period the cost of support for Reading Recovery children was only 10 per cent more than the cost of learning support which schools normally provide, as calculated for the control schools in the London and Surrey evaluation. Brooks *et al*. (1996a) calculated the cost of each participant learning hour in the Basic Skills Agency's Family Literacy Demonstration Programmes as £3.47 (1996 prices), and judged this to be good value.

Large-scale schemes, though expensive, can give good value for money.

2.10 Partnership approaches for reading

Where resources do not permit such large-scale schemes, partnership approaches may be effective. We have given the general label 'partnership' to schemes in which children who are poorer readers are tutored one-to-one, or in small groups, by better readers of the same age, or by older children, or by adult volunteers, or (in some cases) by their teachers.

The best-known partnership approach is Paired Reading, and the effectiveness of this approach has been fully demonstrated. Its siblings, Cued Spelling and Paired Writing, are much less researched to date but seem promising, and operate on the same principle. Anyone interested in following these schemes up is recommended to contact the Paired Learning Centre at the University of Dundee.

The Catch Up Project is perhaps the most precisely designed of the partnership schemes. Originally it was targeted specifically on children who achieve level 1 in reading at the end of Key Stage 1 - see chapter 1 of this report for the evidence that about a sixth or a fifth of children achieve below level 2 - and very practically designed to be delivered by Year 3 teachers in a few minutes per child per week. Though it is increasingly being adapted for and used with other age-groups there is a lot of evidence for its effectiveness.

Other schemes of this general type whose evaluations have demonstrated their effectiveness are Better Reading Partnerships, Knowsley, and Parental Involvement.

All effective partnership schemes rely on providing poorer readers with substantially increased time for reading, supported by a sympathetic, more skilled reader who has received structured training for the purpose, and receives ongoing support. In order to ensure that the increased time on task is effective, focused training for the tutors is essential, so that they know what to do when a reader falters or makes an error. As Bentley and Reid (1995, p.21) put it: 'Hearing children read is **not** teaching them to read.'

Where resources are limited, and partners are available and can be given appropriate training and ongoing support, reading partnership approaches deserve close consideration.

2.11 How can those with the greatest difficulties be helped?

Most of the schemes analysed here worked well for many children with what might be called 'moderate' literacy difficulties. However, there are several indications in the reading data that a number of schemes worked less well for children with more severe difficulties:

- The local adviser in Worcestershire stated that Better Reading Partnerships were not working there for children who were non-readers when the intervention began;

- In the special school in [ILS, School U], intervention produced a negative result, since the children in the control group outperformed those in the experimental group;

- More generally, Ann Lewis's (1999) review on the literature on using ILS with children with low attainments in reading (see the end of section 3.22) showed that no benefits from the technology could yet be proved;

- In Knowsley, children who had to take the lowest level of the test series at pre- and post-test (because they could not have managed any higher level) made no gain at all, on average;

- [Phonological Awareness Training] only helped children in the experimental group (who were all children with MLD or dyslexia) to hold their own, rather than catch up.

On the other hand, Phono-Graphix™ in Surrey did work well for children with severe difficulties; in Reading Intervention (third study) the children with MLD and dyslexia made reasonable gains; and in Reading Recovery in London and Surrey, unlike the rest of the experimental group, children who were non-readers at the start maintained their gains right through to the three-year follow-up. No obvious reason for this discrepancy is apparent.

Therefore success with some children with the most severe problems is elusive, and this reinforces the need for skilled, intensive, one-to-one intervention for these children.

2.12 Duration and impact

Do shorter interventions produce bigger gains, or do gains continue to be mount up during longer interventions? There is a dual problem of publication bias here: non-significant results are much less likely to be published, except in the case of longer-term interventions, which are likely to be more expensively funded and therefore to be under more pressure to publish, whatever the results. It is certainly the case that there are very few moderate to weak impact measures (ratio gain less than 2.0, effect size less than 0.50) in this report for interventions running for a term or less (4-13 weeks).

That said, there are more moderate to weak impact measures from longer-term interventions: see Better Reading Partnerships in Worcestershire, The Catch Up Project (national experimental study, Cornwall, Wakefield, and two out of six RGs in Norfolk), Individual Spelling, Knowsley, [Phonological Awareness Training], [Reciprocal Teaching's RG for comprehension], Somerset (2), [SPELLIT], [Time for Reading], and [ILS, School U in Phase II, Phase III overall, and for pupils with low attainments in reading].

Interventions longer than one term do not necessarily produce proportionally greater benefits.

2.13 Immediate benefits: satisfactory versus good

Both this and the first edition have used the rule of thumb that ratio gains of 1.4 or more and effect sizes of 0.25 or more represent gains that are definitely more than standard progress, and therefore educationally significant. But it now seems appropriate to draw a further distinction, between interventions with an RG of more than 2.0 or an effect size of more than 0.50, and those with less impact. Those with impact measures above these values seem to represent not just satisfactory but good extra progress.

How reasonable is it to expect local initiatives to meet this definition of good extra progress? The following 18 schemes all provide evidence of impact of this order in at least one study: Better Reading Partnerships, The Catch Up Project, Cued Spelling, Family Literacy, Inference Training, IA&T, MTSR, Paired Reading, Parental Involvement, Phono-Graphix™, Reading Intervention, Reading Recovery, Reciprocal Teaching (for reading accuracy),

Somerset, THRASS (for reading, though less reliably for spelling), AcceleRead AcceleWrite, RITA, and Paired Writing.

Good impact – sufficient to double the standard rate of progress – can therefore be achieved, and it is reasonable to expect it.

2.14 Lasting benefits

Finally, do children sustain the improvements they make in intervention experiments, or do the gains tend to 'wash out' afterwards? If quantitative evaluations of the sort analysed here are rare, studies in which the participating children are followed up after the intervention are even rarer. However, 10 of those covered in this report did follow children up at least once. The details are summarised at the end of the Appendix, and may be glossed by saying that

> **- in three cases the children at least maintained the improvements, and**

> **- in four cases the children made a further relative gain, but**

> **- in three the gains partly washed out.**

No generalisations seemed reasonable about when gains might wash out or not.

2.15   Conclusions

- Ordinary teaching ('no intervention') does not enable children with literacy difficulties to catch up.

- Work on phonological skills for reading should be embedded within a broad approach.

- Schemes for children who struggle with spelling work best when highly structured.

- Children's comprehension skills can be improved if directly targeted.

- Working on children's self-esteem and reading in parallel has definite potential.

- ICT approaches only work if they are precisely targeted.

- Large-scale schemes, though expensive, can give good value for money.

- Where reading partners are available and can be given appropriate training and support, partnership approaches can be very effective.

- Success with some children with the most severe problems is elusive, and this reinforces the need for skilled, intensive, one-to-one intervention for these children.

- Interventions longer than one term do not necessarily produce proportionally greater benefits.

- Good impact – sufficient to double the standard rate of progress – can be achieved, and it is reasonable to expect it.

- Most of the schemes which incorporated follow-up studies showed that the children maintained their gains.

CHAPTER 3
THE SCHEMES AND THEIR EVALUATIONS

This chapter describes the 25 schemes, in three sections:

- general schemes targeting reading and, in some cases, spelling
- schemes which involve the use of ICT
- the few experiments of any rigour which have investigated how to improve writing.

Within each section the schemes are dealt with in alphabetical order. Each description contains an outline of the scheme itself, followed by a few details of its evaluation, and references. Where the report which is referenced may be difficult to obtain (for example, if it is an unpublished mimeograph), a contact address is also given. First, some general characteristics of the 25 schemes are summarised in Table 4 on page 20.

Table 4: General characteristics of the schemes

Ref no.	Programme		Y1	Y2	Y3	Y4	Y5	Y6	Duration (weeks)	Number of Sessions for Each Child in Experimental Group	Taught By
1	Better Reading Partnership		•	•	•	•	•	•	10	2 or 3 x 15 mins a week	Other adults, 1-1
2	The Catch Up Project			•	•	•	•	•	39	10 mins a week indiv. + 15 mins/week group	Teachers, 1-1
3	Cued Spelling					•	•		6	3 x 15 mins a week	Parents, other pupils, 1-1
4, 25	Family Literacy		• N R	•		•			12	8 hours a week	Other adults, group
5	Individual Styles in Learning to Spell			•	•				26, 22	5 words daily	Teacher, 1-1
6	Inference Training				•				4	2 x 20-45 mins a week	Other adults, group
7	Interactive Assessment and Teaching			•	•				10	variable	Teachers, group
8	Knowsley Reading Project							•	52	variable	Other adults, 1-1
9	Multi-Sensory Teaching System for Reading			•		•			8	variable	Teacher, group
10	Paired Reading		•	•	•	•	•	•	9	variable	Other adults / pupils, 1-1
11	Parental Involvement			•	•				52	variable	Parents, 1-1
12	Phono-Graphix™		•	•	•	•	•	•	12-26	12 x 1 hour sessions 3 x 20 minute follow up for each 1 hour session.	Teachers & other adults, 1-1
13	PAT					•	•	•	20	10 mins daily	Teachers, 1-1
14	Reading Intervention			•	•	•	•	•	25	2 x 30 mins a week	Teachers, group
15	Reading Recovery		•	•					20	30 mins daily	Teachers, 1-1
16	Reciprocal teaching							•	16	20 sessions	Teachers, group
17	Somerset self–esteem and reading.	(1)				•			26	20 mins week/indiv + 30 mins a week/group	Other adults, group / 1-1
		(2)				•			18	30 mins a week/indiv. + remedial	
		(3)			•	•			17	30 mins a week/indiv. + remedial	
		(4)			•				20	45 mins a week/pair + 3 x lhr a week group	
18	SPELLIT			•	•	•			30	37 hours total	Teachers, group
19	THRASS			•	•	•	•	•	22	30 mins daily	Teachers, group
20	Time for Reading		R						26	variable	Volunteers, 1-1
21	AcceleRead AcceleWrite				•	•	•	•	4	20 mins daily	Computer & supervising adult, 1-1
22	Integrated Learning Systems			•	•	•	•	•	26	variable	Computer & supervising adult, 1-1
23	Reader's Intelligent Teaching Assistant			•	•				17	variable	Computer & supervising adult, 1-1
24	Paired Writing		R		•			•	8	Variable	Other pupils, 1-1

Key: **N** = Nursery, **R** = Reception

Note: **The year-groups shown are those within Years 1-6 for which the scheme has been evaluated and analysed for this report. There are also some entries for Nursery and Reception.**

A. General schemes

3.1 Better Reading Partnerships

The Better Reading Partnership, developed originally in Bradford but no longer in use there as a Wave Three intervention, aimed to help children to become better readers by providing explicit training for adults. The 15-minute sessions, which occurred three times a week, followed a common structure and focused on the development of independent reading strategies.

(1) Bradford

Scheme
In Bradford LEA the Better Reading Partnership developed out of the realisation that those experiencing difficulties in reading were not finding current strategies of simplifying text very helpful. Such approaches were found to result in poor quality learning and a heavy dependence on the teacher.

The national charity Volunteer Reading Help was a vital component in this partnership. They recruited adults from existing curriculum support staff and parent volunteers already helping in the school. The two-day training course included a direct observation using a one-way viewing facility. The ongoing training was supported by a project co-ordinator who met with the partners to discuss the development and progress pupils were making and consider new aspects of the reading process. In 1997, the partners had already worked with 1649 children.

The partners worked with the pupils for ten weeks. They read together for 15 minutes, three times a week. Each reading session followed a common structure of re-reading a known and familiar text, re-reading a book recently taken home, and introducing a new text. The focus was on the development of independent learning strategies. The reading partners were encouraged to discuss the text with the pupil and were trained to prompt the pupil to problem solve difficulties and to develop reading behaviours that would have maximum pay-off.

Evaluation
Bradford LEA carried out its own evaluation. The Suffolk reading test was used to monitor the effectiveness of the programme for the cohort of pupils who took part in 1995-96. They made significant gains in the 10-week period.

Reference
Collins (1996)

(2) Derbyshire

The ROWA! (Read On – Write Away!) initiative in Derbyshire took up BRP from Bradford under the name Better Reading Partners as one of its schemes in 1998 – it was adapted by two educational psychologists. By July 2002 they had trained just over 2000 partners, of whom half were volunteers, the rest being teaching assistants, etc., and over 8000 children had been partnered.

Evaluation
ROWA! has carried out its own evaluation every year. The clearest data provided were for 656 children the single year 1998-99, when most of the partners were teaching assistants, and this is what is analysed in the Appendix. Average gains were excellent in each year group.

Reference
Taylor (1999)

Contact
Carol Taylor
Director
Read On – Write Away!
County Hall
Matlock
Derbyshire DE4 3AG

carol@rowa.co.uk
www.rowa.co.uk

(3)   Co. Durham

In Durham LEA the Better Reading Partnership aims to improve the skills of average and below average readers, enabling them to become more successful in their independent reading.

Adults working on the programme receive two days training and two follow-up interviews with the literacy consultant in charge of the programme to discuss their sessions and their pupil's progress.
Adult partners may be teachers, learning support staff, teaching assistants, ancillary staff such as caretakers, and cooks, parents, grandparents, community workers, governors and friends of the school.

As in the Bradford model, which is followed with only slight variation, the partners work with the pupils for ten weeks. They read together for 15 minutes, three times a week. Pupils' reading includes a variety of fiction, non-fiction and poetry texts and is drawn from reading schemes as well as 'off-scheme' books and magazines and newspapers for older pupils.

Evaluation
The LEA had collected its own monitoring data, and the impact was substantial, especially in KS1.

Reference
Unpublished data supplied by Ann Foster

(4) Redcar and Cleveland

Very few details were available on this.

Evaluation
However, the LEA had collated data on over 1000 pupils, and the impact was substantial.

Reference
Unpublished data supplied by Andrew Taylor

(5) Worcestershire

Worcestershire began using BRP in 1999, as the approach adopted for an Innovative Development Project (IDP) funded by the Basic Skills Agency. Modifications to the Bradford model were that the partners were parents, and they were asked to read two books at each session (not three), and they read with children twice a week for 15-20 minutes. One book was familiar, and the second was new, and the new text from each session became the familiar book for the next. In 2000-02, the LEA has trained many parents, and about 60 have achieved accreditation for this work. About 25 schools were using the scheme in 2002; in addition 10 schools in an Education Action Zone use it with teaching assistants.

Evaluation
The 1999-2000 IDP was evaluated by two researchers from the National Foundation for Educational Research. The scheme was generally successful, but the LEA adviser commented (Anthea Main, personal communication, 29 September 2002) that 'There were gains across all year groups as long as the children had started reading – it is not successful with non-readers.'

Because the evaluation used a cross-over design, the phase 2 data from the first group to receive the intervention are effectively follow-up data. The phase 2 data from that group show that that group continued to make approximately standard progress. They were not making any further relative gain, but were maintaining the gain made in phase 1.

Reference
Brooks and Hutchison (2000)

3.2   The Catch Up Project

The Catch Up Project targets struggling readers in Years 2-6. It is recognised that for these children to meet National Curriculum requirements, an intensive yet manageable programme is of critical importance. The Catch Up Project is a 10-minute structured teaching programme that is carried out once a week with individual children by the class teacher or teaching assistant.

Scheme
The Catch Up Project was initially developed in 1998 at Oxford Brookes University, in partnership with the Caxton Trust, for struggling Y3 readers, as a result of a study undertaken by the project consultants, Diana Bentley and Dee Reid. The research helped to identify a systematic method for supporting struggling readers which could be readily adopted by classroom teachers. The child must complete a comprehensive assessment procedure before the programme begins in order to determine the correct level, as well as to provide some pre-intervention data.

The 10-minute individual sessions are divided into three discrete sections. The first two minutes are called the *prepared reading approach.* A book at instructional level is selected, and the text and pictures are scanned to introduce vocabulary and familiarise the story. In the next four minutes the child reads the story whilst the teacher records progress and identifies points to follow up. The final four minutes are the follow-up, a linked writing or spelling activity. This is where the teacher acts upon the information gleaned and decides which skill should be worked on.

The aim is to enable the children to read with accuracy and understanding. The Catch Up Project has produced various support materials, including a CD ROM, a Parents Link book, and a videotape. By 2001-02 it was being implemented in about 3,000 schools and supporting children in Y1-4. Most of the recorded gains were substantial. In the national experimental study, however, they were less so – this may have been because the scheme was (apparently) less firmly supported with these schools. From September 2002 it was extended to support struggling readers in Y5-6 (some evaluation data already exist for those year groups) and in 2003 is due to be extended to secondary school pupils.

Contact
Julie Lawes
Project Director
The Catch Up Project
Thetford EAZ
Baxter Healthcare
Caxton Way
Thetford
Norfolk IP24 3SE

catchup.eaz@virgin.net
www.thecatchupproject.org

Evaluations

The full programme of the Catch Up Project is intended to last a whole school year. However, the initial evaluation was a one-term pilot study carried out by the programme developers in the Autumn term of 1997, with some statistical advice from the National Foundation for Educational Research in the school year 1998-99.

Reference

Clipson-Boyles (2000)

Since then, the National Experimental Study replicated the original Pilot Study over an academic year with a larger group of schools. Three groups of children were monitored over this period to compare the progress of those on The Catch Up Project with (a) those given a non-specific matched time equivalent intervention by the teacher, and (b) a control group. However, the total number of children in this study was only 98, and the results were not clear-cut in favour of the programme.

Other evaluation data from Cornwall, Norfolk and Wakefield provide stronger evidence on the basis of samples totalling over 1,300. The Norfolk projects involved the extensive use of teaching assistants, supported by teacher coordinators.

References

Worsley (2001), and unpublished data supplied by Julie Lawes

3.3   Cued Spelling

Cued Spelling is a procedure designed by Keith Topping and colleagues at the University of Dundee for two individuals working together. The pair might be parent and child working at home or two children working together in school. In school, the children can be of the same or different age and spelling competence. They may remain in role as tutor and tutee, or the roles may reverse at intervals. Cued Spelling can also be used for whole-class tutoring.

Scheme
According to the authors, the technique consists of 10 steps, 4 points to remember, and 2 reviews – a chart setting all this out can be downloaded from the website. The most accessible description of the method is in Topping (2001). He admits (p.181) that it looks 'rather complicated' but maintains that 'You can train seven-year-olds to do it half an hour – it is a lot simpler than it looks.' It is usually done three times a week for an initial trial period of six weeks. Each session takes about 15 minutes.

Evaluation
Two evaluations by Topping and colleagues are analysed in the Appendix. Both are summarised in Topping (1995, 2001). France *et al*. (1993) trained 47 parents as tutors, and gathered test data on 22 children who were Cued Spelling tutees and on 10 better spellers in the same class as a comparison group. Watt and Topping (1993) gathered similar data on 10 children tutored by their parents, with a comparison group of only 6. In both cases, the results were promising.

References
France *et al*. (1993), Topping (1995, 2001), Watt and Topping (1993)
www.dundee.ac.uk/psychology/TRWresources

3.4   Family Literacy

(1) Basic Skills Agency's Demonstration Programmes

The aims of the scheme were intergenerational; they balanced intended benefits for the parents' literacy with intended benefits for their children. It was hoped that improving parents' skills would enable them to help develop their children's language and literacy.

Scheme
The scheme was devised at the Basic Skills Agency in 1994, and stemmed from the fact that children whose parents experience problems with literacy are themselves more likely to experience literacy difficulties, thus continuing the cycle. The Basic Skills Agency devised the initiative with the aims of raising standards of literacy among adults with difficulties, boosting their ability to help their children, and increasing the children's literacy skills.

The programme recruited those most in need of help. The participating parents were therefore, in general, poorly qualified and not employed outside the home. The programme was set up in four areas of multiple deprivation in Cardiff, Liverpool, Norfolk and North Tyneside. Any parent who had a child aged between three and six years was welcome, as long as both parent and child attended the course. Between the four programmes 361 parents and 392 children completed the course during the period of the evaluation.

The course ran 8 hours a week for 12 weeks. Each week there were two separate sessions (parents in one room, children in another) and one joint session. In their sessions parents worked on their own literacy skills and towards accreditation for their achievements, and learnt how best to help their children. The children's sessions were a mixture of nursery and infant school practices and approaches, as appropriate to the ages of the particular children attending. In the joint sessions the parents worked with their own children, applying what had just been learnt.

Evaluation
A team at the National Foundation for Educational Research (NFER) was commissioned to evaluate the initiative in the four school terms from Summer 1994 to Summer 1995. For the purposes of this report, only the data on the children's reading attainment are considered. All children aged at least 5 on entry to the course were given the Reading Recognition subtest of the Peabody Individual Achievement Tests (PIAT) both at entry and on completion of the course. Data were available from this test on 147 children at pre- and post-test. Varying numbers of children were re-tested on the PIAT at three points: 12 weeks and 9 months after the end of the intervention, and between January and April 1997, which was between 20 and 34 months after the end of the intervention for individual children.

The pre-test showed the children as disadvantaged and at great risk of educational failure. During the courses, they made an average gain of over 4 standardised score points, and the educational outlook for many of them was improved. At 12-week follow-up, the Summer and Autumn 1994 cohorts had made further relative gains, but not the Spring and Summer 1995 cohorts. At 9-month and 1997 follow-ups, the children had on average sustained their gains.

References
Brooks *et al.* (1996a, 1997); Gorman and Brooks (1996)

(2) The Basic Skills Agency's model in Hampshire

A replication of the Demonstration Programmes was mounted in Hampshire in 1996-97, and evaluated by Dwynwen Stepien of the University of Portsmouth with a small sample (27). However, the impact seemed even more powerful than in the Demonstration Programmes.

Reference
Stepien (1997)

(3) The Basic Skills Agency's model adapted for New Groups

For reasons beyond the Agency's control, the Demonstration Programmes contained hardly any families from linguistic minorities. By design, the programmes were limited to families with a child aged 3-6. In a further initiative in 1997-98, the Agency set up pilot programmes for linguistic minority families and for families with a child in Y4. These were again evaluated by a team from NFER. The adaptations for linguistic minority families and those with a child in Y4 were judged appropriate, with successful adaptation for linguistic minority families requiring close attention to issues of bilingualism. Both groups of children made substantial gains.

Reference
Brooks *et al*. (1999)

3.5   Individual Styles in Learning to Spell

Scheme
In 1996-99, the D*f*EE funded a three-year project by the Helen Arkell Dyslexia Centre investigating individual styles in learning to spell. Several pilot and exploratory phases led to the production of a teaching pack offering 10 different teaching approaches, which teachers were to select from for individual children based on a simple assessment of their preferred styles.   The approaches were: Neurolinguistic Programming, Onset-Rime, Look-Cover-Write-Check, Own-Voice, Tracing, Simultaneous Oral Spelling, Picture Association, Mnemonics, Phonics, and Look-Say. All the approaches were described briefly in a teaching pack which is reproduced at the end of the research report.

Evaluation
The developers of the project evaluated it both in special schools (with groups too small to be reported here), and in three mainstream schools, of which two provided data that could be analysed for this report. A cross-over design was used, involving two groups of children. In phase 1, one group received the intervention while the other acted as a comparison group. In phase 2, the second group received the intervention, while the children in the first group also continued to do so. The results were in line with prediction. Both groups made significant gains in both phases. In phase 1 the first group made significantly more gain than the comparison group; in phase 2 the difference was non-significant. Thus both groups made good gains while receiving the intervention, while the second group made only standard progress in phase 1 before receiving it, and the first group continued to make better than standard progress in phase 2 (follow-up).

Reference
P. Brooks and Weeks (2000)

3.6   Inference Training

This scheme focuses upon the band of children who fall within the normal range of cognitive ability yet fail to comprehend fully what they read. The authors break down the many skills needed to understand a text into manageable chunks: lexical elaboration, question generation and comprehension monitoring. Tasks are designed so as children can make links between the text and its meaning. Sessions last between 20-45 minutes, twice a week for four weeks.

Scheme
Previous studies by Yuill and Oakhill (researchers at the University of Sussex) had shown that less skilled readers have difficulty in making inferences from text. They argue that word recognition and decoding skills are not always adequate in developing good reading skills. The meanings of individual sentences and paragraphs have to be integrated so as to understand the main ideas of the text. It has been suggested that working memory plays a part in this skill.

Yuill and Oakhill tested the effect of children's reading comprehension using three types of intervention: 1. Inference skills training (this consisted of lexical inference, question generation and prediction); 2. Comprehension exercises; 3. Rapid decoding practice. The same narrative texts were used in all three intervention conditions. The experimenter saw children in groups of three to five, twice a week over three and a half weeks. Length of sessions varied from 20 to 45 minutes. Training sessions lasted slightly longer than control sessions, since subjects had to spend time thinking of questions, whereas the control group had precise tasks to perform that did not involve long periods of silence.

Evaluation
This was an experimental study, rather than an evaluation of a separately devised project. The results showed that less skilled comprehenders benefited from Inference Training more than skilled comprehenders. The authors concluded that for less skilled comprehenders Inference Training was both more beneficial and more helpful than decoding practice. However, comprehension exercises appeared to be as beneficial as Inference Training.

This is of interest as few studies have tackled comprehension improvement directly. It is believed that children in the Inference Training groups gained new confidence and enjoyment from the reading tasks, and motivation was high relative to the repetitive tasks required in the decoding groups. However, observations showed that the decoding group found the rapid reading task challenging, and they had faster reading speeds. This suggests that less skilled comprehenders' deficits are not a result of slow decoding.

Reference
Yuill and Oakhill (1988)

3.7 Interactive Assessment and Teaching

This entry needs to be read in conjunction with that for RITA – section 3.23.

Scheme
The Interactive Assessment and Teaching (IA&T) approach is the programme advocated by Reason and Boote (1994). IA&T is a pragmatically based approach that is designed for children with special educational needs, and is compatible with current theoretical approaches to the teaching of reading. It is an individually adaptive, curriculum-based, support programme with the emphasis on word building and phonics skills in the broad reading context. The model of literacy development includes both writing and reading, and their interaction. Furthermore, three separate aspects of literacy are considered – meaning, phonics and fluency. While beginning readers often have difficulties with 'meaning' - understanding the 'rules of the reading game', it is with phonics and fluency that most poor readers struggle. On the basis of this theoretical and pragmatic analysis of development of the skills and knowledge in reading, Reason and Boote developed their 'step by step' approach to tailoring the reading support to the individual capabilities of each reader. This involves five steps: First, make an initial assessment in terms of the four stages for Meaning, Phonics and Fluency separately. Second, decide on priority teaching areas. Third, develop a support plan, in terms of the objectives and the learning steps involved, making sure that each step is achievable. Fourth, select appropriate teaching methods and teach each step, trying to ensure variety and motivation. Finally, record and evaluate progress, keeping records for each step.

Evaluation
The IA&T programme was studied in the first of two phases of a research programme devised by researchers from the Psychology Department of the University of Sheffield. Both phases involved children in Y2 and Y3.

The results of phase 1 were mixed. The IA&T children made significantly greater gains than those in the control groups during the 10 weeks of the intervention in both reading and spelling. For Y2 the gains in spelling were maintained in follow-up tests six months later, but almost completely lost in reading (the control group had slipped back even further). For Y3 the gains in reading were maintained, while the spelling gain was partly lost.

In phase 2, the pre- and post-test data from phases 1 were used for 'alternative intervention' comparisons with the same research team's computer-based version of IA&T, RITA (Reader's Intelligent Teaching Assistant). RITA has its own entry later in this chapter, and the comparisons between IA&T and RITA are considered there.

References
Nicolson et al. (1999), Fawcett et al. (1999), Reason and Boote (1994)

3.8 Knowsley Reading Project

This project was excluded from the first edition because it concerned Y6 children, and the first edition's coverage was Y1-4. However, as Elliott *et al.* (2000) pointed out, it is one of very few detailed studies of the use of adult volunteers conducted in the UK.

Scheme
The key feature of the project was the recruitment and training of large number of adult volunteers who helped primary (Y6) pupils with their reading on a regular basis. Other components were:

- training for parents and other volunteers. This was believed to be one of the most important components, because in a pilot project it had seemed to make the most significant difference to raising levels of reading attainment;

- a two-day residency for teachers at each participating school which included demonstrations of range of techniques, and opportunities for teachers to observe and evaluate the demonstrations using OFSTED criteria;

- an agreed common approach to teaching reading, which defined the range of reading skills pupils needed, including word recognition, phonic skills, prediction, memory skills, recall of previous reading, comprehension, and the ability to choose books for different purposes;

- an audit of school policy and resources;

- a consistent and continuous system of assessment; and

- a support network.

Evaluation
Knowsley LEA commissioned an evaluation from NFER. It focused on the 13 schools which entered the project in the school year 1994-95, and within those schools on pupils in Y6. None of the schools had been involved in the project before, and the evaluation had a one-group, pre-/post-test design. Almost all the groups involved made good progress, but the small group who had the lowest scores to begin with made almost none.

Reference
Brooks *et al.* (1996b)

3.9   Multi-sensory Teaching System for Reading (MTSR)

Scheme

MTSR is described by its UK providers as 'a fully scripted, multi-sensory, structured, sequential package for teaching word level reading. It is based round the teaching of three elements of the reading process – phoneme/grapheme relationships, rules of English, and how to tackle irregular words.' It is derived from a scheme developed in Texas by Margaret Taylor Smith and called Multi-sensory Teaching System (MTS). MTSR was developed and produced at Manchester Metropolitan University in collaboration with the British Dyslexia Association; development was financed by a research grant from the (then) Department for Education and Science. The published teaching pack consists of a teacher's book, two books of teaching materials with cards, and a videotape.

Evaluation

The developers have so far mounted four evaluations: a pilot study in 12 primary schools in three LEAs in the North West of England, and three larger studies, in Rutland, Ireland and Bolton. The pilot study including a comparison with a scheme called Beat Dyslexia (also with Phonological Awareness Training, but the numbers using that were too small to analyse here). The pilot and the Bolton study (which was financed by the D*f*ES through its 'SEN Small Programmes Fund') are analysed in the Appendix. The impact on reading was good. For spelling the results were curiously contradictory: the pilot study had a large negative ratio gain, while the Bolton study had a massively positive one. They are among the most extreme impact measures that appear in this report.

References

Johnson *et al*. (1999) and unpublished data supplied by Mike Johnson

3.10   Paired Reading in Kirklees

This is one of the simplest schemes yet devised, and the subject of one of the largest evaluations.

Scheme
Paired Reading was devised by Morgan (1976) to meet the needs of children who were finding reading difficult and to involve non-professionals in helping them. He designed it to be simple to administer after the minimum of training, and flexible, in that it could be applied to any form of reading material. The fullest description is Morgan's (1986) book, and it is summarised in diagrammatic form in Topping and Lindsay (1992, p.200) and on the website. Essentially, it is a 'scaffolding' approach in which tutor and child begin by reading aloud together, and the tutor gradually withdraws and leaves the child to read aloud alone. Techniques are specified for intervening when the child falters or makes an error, and praise is given regularly.

Evaluations
Topping and Lindsay (1992) reviewed dozens of studies from across the English-speaking world, and Topping (1990) himself carried out the largest evaluation, of the Kirklees project. That evaluation covered not just one project in that LEA, but 155 projects spread across 71 schools, both primary and secondary. The results consistently showed that the technique was effective, and other partnership approaches have imitated, incorporated or adapted it.

Topping's work has led on to other forms of Paired Learning: Cued Spelling and Paired Writing (which have new entries in this edition) and Paired Thinking (which does not).

References
Morgan (1976, 1986); Topping (1990, 1995, 2001); Topping and Lindsay (1992); Topping and Wolfendale (1985); Wolfendale and Topping (1996). Also, for an evaluation in Dublin with children with moderate learning difficulties, Nugent (2001).

Contact
Prof. Keith Topping
Centre for Paired Learning and Graduate Educational Psychology
Department of Psychology
University of Dundee
Dundee DD1 4HN

Tel: 01382 344628
http://www.dundee.ac.uk/psychology/TRWresources

3.11 Parental Involvement in Haringey

'Learning at a parent's knee' must be among the world's oldest teaching methods, but only in recent years has it been formalised and generalised with this title, and been researched.

Scheme
There have been many schemes bearing or deserving the title Parental Involvement both in the UK and beyond. For the purposes of this report, the first well-known and -researched scheme of this sort in the UK, the Haringey project, has been taken as the paradigm:

> *The experimental innovation in the Haringey project consisted in asking all parents of children in certain top infants' [Y2] classes to listen to their children read aloud for a short period, several times a week, from reading material selected and sent home by the child's class teacher... It was found that... the great majority of parents provided constructive help and support for their children, and avoided counterproductive behaviour such as pushing their children too hard, or confusing them with inappropriate information.* (Hewison, 1985, pp.47-48)

One major factor in that success seems to have been the motivation provided by the parent's close attention to the child's development. The original project provided little in the way of guidance to parents on how exactly to share books with their children, and subsequent schemes have refined this part of the approach.

Evaluation
The original research in Haringey LEA was conducted by a team from the University of London Institute of Education led by Tizard. The approach was found to be highly effective, and much more effective than extra teacher help with reading in school. In addition to pre- and post-tests, the design included two follow-ups, one year and three years on; both showed the gains had been maintained.

References
Hewison (1985); Tizard *et al.* (1982); Topping and Wolfendale (1985); Wolfendale and Topping (1996). Also, for a failure to replicate the Haringey effect in a different context, Hannon (1987), Hannon and Jackson (1987).

3.12   Phono-Graphix™

'Letters are pictures of sounds.' Phono-Graphix™ is based on this beguilingly simple insight.

Scheme
Diane McGuinness (1997, 1998) surveyed what she thought was wrong with initial reading instruction and concluded that the main fault was not taking seriously the core design feature of English orthography – that it is an alphabet, a system for representing (originally and in principle) each distinctive speech sound with one symbol. She began developing a teaching system, and this was fully developed by her daughter-in-law and son, Carmen and Geoffrey McGuinness (1998), into the system known as Reading Reflex or Phono-Graphix™. The scheme arrived in the UK in a blaze of publicity in 1998. Its essential features are that it

- develops the concept that written English is a phonemic code – each sound in a spoken word is represented by some part of the written counterpart

- teaches the phonological skills of blending, segmenting and phoneme manipulation that are needed to use a phonemic code

- teaches knowledge of sound-to-symbol relationships ('correspondences') explicitly.

The scheme is supported by detailed training and materials, and by a network and website.

Evaluation
Phono-Graphix™ has been taken up quite widely in the UK, and substantial data were received from Bristol LEA. When accessed on 14 August 2002, the website gave details of recent work in Scotland and Sunderland, and of a small-scale evaluation at an independent specialist school for children with dyslexia in Surrey. The Surrey and Bristol data are analysed in the Appendix. The impact measures were substantial, including the largest ratio gain for reading of all the studies reviewed in this report.

References
Derrington (2001a, b), C. McGuinness and G. McGuinness (1998), D. McGuinness (1997, 1998), and unpublished data supplied by Sue Derrington
http://www.readamerica.net

3.13   Phonological Awareness Training

In developing the PAT programme (Wilson, 1993), originally for Buckinghamshire LEA, the aim was to provide a new approach to phonics teaching. The idea is that by using the child's existing knowledge about letter-sounds and words, new words which contain familiar 'rimes' (identically-written word endings) should not present a problem in reading or spelling. Training worksheets containing specific rimes, supported by reading lists and sentences for dictation, were intended to enable children to have a better understanding of relationships between sounds and spoken words.

Scheme

Poorly developed phonological skills have been suspected as one of the causes of reading difficulties. The PAT programme is designed to help children in reading, spelling and writing phonically regular words. It is acknowledged that children need to experience a wide variety of stimuli such as story books, poems and rhymes to develop literacy skills. The programme aims to enhance children's literacy skills by making analogies. Children who have existing knowledge of word sounds can develop these by applying them to other words, thus using commonly occurring rimes. This way of teaching encourages a problem-solving based approach rather than the traditional 'listen and learn' way.

The daily 10-minute programme provides intensive work on three skills within the same activities: identifying sounds, blending phonemes together, and segmenting or isolating sounds in words. PAT is made up of training worksheets containing specific rimes supported by reading lists and sentences for dictation. Training requires children to generate their own words from rimes. This is based on the idea that once children can understand the concept of reading and speech made by analogies, all they have to remember is how to form the rimes. There are no pictures or visual cues of any sort due to the fact that the child is learning to focus on sound and to develop a problem solving approach to generate words. Pictures would deflect from developing their own strategies for remembering.

Evaluation

The developer of the programme and a colleague designed and implemented the evaluation. Three schools participated in the study. Equal numbers of programme and comparison children from each of Years 4 to 7 were selected in order to test the hypothesis that children with literacy difficulties who completed the programme would make better progress than their counterparts. All the children in both experimental and control groups were on at least Stage 2 (now School Action) of the SEN Code of Practice; the two groups of 24 children were carefully matched. The pre-tests were carried out between two and five weeks prior to the PAT programme. The post-tests were carried out by educational psychologists who did not know which intervention group the children were in. The programme ran for 20 weeks. Groups of six met four days a week for 20 minutes.

The results were not clear-cut. The children in the experimental group did make significantly more progress than those in the control group; but the children in the experimental group made scarcely any more progress than would have been expected from ordinary classroom teaching and development.

References
Wilson (1993); Wilson and Frederickson (1995)

3.14   Reading Intervention (formerly Cumbria Reading with Phonology Project)

The Cumbria Reading with Phonology study supports the view that it is the combination of phonological training and reading that is important for helping poorer readers. Children are helped to isolate phonemes within words to appreciate that sounds can be common between words and that specific sounds can be represented by particular letters. Lessons also include story work.

Scheme
It is well documented that children who exhibit good phonological skills appear to make the most progress in learning to read. This study illustrates that a combined phonological and literacy skills training programme effectively boosts the reading skills of reading-delayed seven-year-olds.

Poor readers in Y2 were assigned randomly to one of four groups. Group one received training in phonological skills and help in learning to read. Pupils in the second group received teaching in reading alone – the teaching of reading here and in the first group was similar to Reading Recovery. The third group received teaching in phonological skills alone. A control group received normal teaching. During the intervention period, which lasted 20 weeks, each of the experimental groups received forty 30-minute teaching sessions.

The reading with phonology package combined a highly structured reading scheme with systematic activities to promote phonological awareness. The first section of a session was devoted to re-reading a familiar book whilst the teacher kept a record of the child reading. This allowed for rehearsal of familiar words in different contexts. Phonological activities and letter identification were also involved in the first part of the session, accomplished using a multi sensory approach (feeling, writing and naming.) The second part of the session involved writing a story and cutting it up. The last part of the session introduced a new book.

Evaluation
This was a very tightly designed and administered study, carried out by an adviser in Cumbria LEA and two colleagues from the University of York. The four groups were matched on reading age at the pre-test, and teaching time for the three experimental groups was equated as closely as possible. The 93 children in the three experimental groups were taught by 23 teachers. Each teacher worked with groups of two to nine children in order to reduce the effect of differentiation. The time of day at which children received their intervention was systematically varied. The testers were unaware of the children's experimental status.

The Reading with Phonology group made significantly more progress in reading than the other three groups; the other groups did not differ - in other words, neither reading-only nor phonology-only brought about any greater progress than normal schooling.

Much the same finding emerged from a very similar study in Rhode Island, USA (Iversen and Tunmer, 1993), in which the reading intervention was the authorised form of Reading Recovery.

The reading with phonology approach, now known as Reading Intervention, or as Sound Linkage, the name of published materials derived from and supporting it, continues to be in widespread use in Cumbria, and Peter Hatcher has published further research on it. This

shows that the initiative continues to be effective for the generality of poor readers, and even for children with moderate learning difficulties or dyslexia.

References
Hatcher *et al.* (1994), Hatcher (2000)

3.15   Reading Recovery

Reading Recovery arose out of an extensive research project carried out by Clay in New Zealand. It identifies children who are having difficulty in acquiring literacy skills at an early stage of their school career and aims to provide help before problems become consolidated. The programme is delivered for 30 minutes on a daily basis, by a specially trained teacher. The lessons consists of a series of activities, including reading two or more books, one familiar and one new. It encourages children to monitor their own reading.

Scheme

Reading Recovery is aimed at children who after one year of schooling show they are having difficulty with reading. Children identified as being in the bottom 20 per cent of the class in reading receive daily 30-minute individual lessons for up to 20 weeks from a specially trained teacher, who provides highly responsive instruction tailored to the needs of each child. Throughout the lesson the teacher's interventions, based on daily diagnoses, are carefully geared to identify and praise successes, promoting confident and independent behaviour. This ensures that a range of strategies are brought to bear whenever problems arise. Children leave the programme (are ' successfully discontinued') when reading improves to the level of the average class reading group, enabling them to work in class without additional support. Children who are not successfully discontinued are referred for more detailed assessment and specialist help.

The first LEA in the UK to introduce Reading Recovery was Surrey, in 1990. In 1992 a group of 20 other LEAs in England and Wales received funding to introduce it, and it has since been taken up by further LEAs in England and Wales, and by all the Education and Library Boards in Northern Ireland.

Evaluations

The original request for information for the first edition of this report produced more replies about Reading Recovery than about any other initiative, and these constituted about a quarter of all of the information received then. The following LEAs sent reports on their local evaluations: Bradford, Cambridgeshire, Gwent and South Glamorgan (jointly), Hammersmith and Fulham, Lambeth, Northamptonshire, and Rotherham. Also available were reports on local evaluations in Surrey (Prance, 1992; Wright, 1992) and Northern Ireland (Gardner *et al.,* 1997). More recently, a further and larger-scale evaluation in Northern Ireland has been carried out and reported (Munn and Ellis, 2001), and Bradford LEA again sent local data in response to the request for information for the second edition. However, none of these reports provided either an impact measure or data from which such a measure could be calculated. The only reports which did provide such information were the London Institute of Education study of Reading Recovery in six London LEAs and Surrey (Sylva and Hurry, 1995a, b) and one stage of local data-gathering in Bristol (Fudge, 2001). Since the Sylva and Hurry study also included both carefully-chosen control groups and an alternative intervention condition (Phonological Intervention), it has been taken as the main representative of Reading Recovery for the purposes of this report, and the Bristol report is also analysed.

(1)   London and Surrey

The Phonological Intervention condition in the London and Surrey study gave children additional tuition in the specific area of phonological awareness.

The Sylva and Hurry study showed that, during the intervention, Reading Recovery children made significantly greater progress than either Phonological Intervention children or those in the relevant control group; and that Phonological Intervention children did not make significantly greater progress than those in their control group. The Bristol study also showed a strong impact.

Sylva and Hurry (1995) also reported a follow-up conducted in May-July 1994, one year after the end of the intervention, and Hurry and Sylva (1998) reported a further follow-up after three years (Summer 1996). At the one-year follow-up, the Reading Recovery children were slightly less further ahead of, but still significantly better than, the main control group. And by this point the Phonological Intervention group were significantly better than their control group, so the gap was narrowing.

At the three-year follow-up, neither the Reading Recovery nor the Phonological Intervention group was significantly better overall than their respective control groups, so the main effect had washed out, but children receiving free school meals, and Reading Recovery children who had been complete non-readers at the pre-test in 1992 were still maintaining their gains.

Other evidence on the long-term effectiveness of Reading Recovery, from Australia and New Zealand (Lowe, 1995; Moore and Wade, 1998), shows more lasting benefit and less wash-out.

References
Clay (1979, 1985, 1993); Gardner *et al.* (1997); Hurry and Sylva (1998); Lowe (1995); Moore and Wade (1998); Munn and Ellis (2001); Prance (1992); Sylva and Hurry (1995a, b); Wright (1992)

(2)   Bristol

This local study contained 145 children – more than the main experimental group in the London and Surrey study. It showed strong impact of Reading Recovery.

Reference
Fudge (2001)

3.16   Reciprocal Teaching

Scheme
The Reciprocal Teaching Method is a teaching approach first described by Palincsar (1982) and then further developed by her and Brown (Palincsar and Brown, 1984; Palincsar, 1986). They describe it as:

> A procedure … where teacher and student took turns leading a dialogue concerning sections of a text. Initially the teacher modelled the key activities of summarising (self-review), questioning (making up a question on the main idea), clarifying and predicting. The teacher thereby modelled activities: the students were encouraged to participate at whatever level they could. The teacher could then provide guidance and feedback at the appropriate level for each student.
> (Palincsar and Brown, 1984, p.124)

The four activities are seen as having two functions, 'comprehension-fostering and comprehension-monitoring' (p.121). Pupils are gradually encouraged to take over the teacher role as they gain confidence, and the whole approach is predicated on the idea that poorer comprehenders can improve by being shown and explicitly understanding and adopting good comprehenders' strategies.

Evaluation
There has been a large amount of research on the technique in North America, but scarcely any in the UK. Rosenshine and Meister (1994) did a meta-analysis on the 16 most rigorous studies and produced an effect size of 0.32 on standardised tests. In the UK the technique is in use at least in Westminster LEA (where Greenway, 2002 reported an experiment with seven children, sadly too small a number to analyse here) and Haringey – where an evaluation with 16 children (data supplied by Christa Rippon) showed a very strong gain in reading accuracy and an even stronger impact on comprehension.

References
Greenway (2002); Palincsar (1982, 1986); Palincsar and Brown, 1984; Rosenshine and Meister (1994); and unpublished data supplied by Christa Rippon

3.17 Somerset Self-esteem and Reading Project

This was a series of four studies all conducted by Denis Lawrence, who was at first a specialist remedial teacher, and then from 1974 to 1981 (see Lawrence, 1988, biographical note opposite title page) Chief Educationalist Psychologist, in the county. He was convinced that low self-esteem had a detrimental effect on children's attainment, and that it was absurd to tackle only the low attainment.

Schemes
Lawrence's studies were specially-designed experiments rather than initiatives to which evaluations were added. And in this analysis they constitute the only *series* of studies, where each built on the preceding. The fourth study was on a large scale, involving 335 children in all; the others were much smaller. In each study, the children in the experimental group received some form of Rogerian self-esteem counselling plus a specific reading intervention. Otherwise the studies have in the main to be described separately.

Somerset (1)
Here the counselling was provided by a professional psychologist who was a remedial specialist (presumably Lawrence himself), and the specific reading intervention was the remedial teaching already provided within the school, which was mainly phonics. A control group received no extra intervention, and there were two alternative intervention groups: one received only counselling, while the other received only the remedial phonics teaching. Lawrence (1971, p.120) gave a half-page description of the counselling, and summarised it as follows: 'This involved a responsible, sympathetic adult, with status in the eyes of the child, communicating to the child that he enjoyed his company.' The fullest account of Lawrence's approach and recommendations is in Lawrence (1988).

Somerset (2 and 3)
In each of these, there were only two groups. The children in the experimental groups received counselling plus remedial teaching, while those in the control groups received only remedial teaching (and were therefore alternative intervention, rather than no-intervention groups). The counselling appears to have been intended to be identical to that in the first study, except that it was provided by non-professionals: 'The head teacher of each school contacted a woman in the area whom he considered to be a suitable person for the experiment' (Lawrence, 1972, p.49). These non-professionals were trained by 'the psychologist' (presumably again Lawrence himself).

Somerset (4)
There were four groups of pupils: a no-intervention control group, and three groups who all received DISTAR. One group received only that intervention, while the other two received in addition one of two 'therapeutic' interventions designed to boost pupils' self-esteem about reading.

The DISTAR-only group received instruction in the skills of reading through the Direct Instructional Teaching technique devised by Engelmann *et al.* (1969). The teachers involved in using DISTAR with this and the other two relevant groups were all trained in the technique by a manager of the scheme's UK promoters. The children were taught in groups of 6-10, according to the number identified in each school as low attainers, for one hour, 3 times per week. In this technique, children sit in a semi-circle within touching distance of the teacher. The lowest-attaining children are placed in the centre. They interact continuously with the

teacher, learning word patterns out loud. The sequences are highly structured, and are taught until all children have mastered them.

The children in the experimental group receiving the first of the 'therapeutic' interventions in this study received DISTAR as above, plus counselling once a week for 20 weeks from one of 35 non-professionals. The children were seen in pairs, for 45 minutes each time. The counsellors were selected by the head teachers of the schools involved. They had four meetings beforehand at which they were given handouts on how to structure the sessions with games and activities. These had been designed by the experimenter, or were those described by Canfield and Wells (1976). The counsellors were also briefed on self-concept theory and on the establishment of empathy as described by Rogers (1975) and 'modelling' as described by Bandura (1977). The essence of the intervention was an accepting and non-judgmental relationship between counsellor and children.

The group receiving the second of the 'therapeutic' interventions received DISTAR as above, plus a weekly drama session designed to enhance self-esteem. Groups varied in size from 7 to 15. The sessions lasted about 45 minutes, and were taken by the County Adviser for Drama. They were structured to allow the children to experience success, and included role-playing of 'experts', for example they would be on an imaginary journey and would each be given a different expert role. The rule was that no criticism of experts was allowed. The drama condition was intended to illustrate an even simpler method of delivering self-esteem improvement than the delivery of counselling by non-professionals.

Evaluations
These were all carried out by Lawrence. The following summary gives the main results for each study, and then some overall conclusions.

Somerset (1)
Professional counselling only was effective. The evidence for counselling plus remedial phonics teaching, and for remedial phonics only, was less clearcut. Professional counselling plus remedial phonics was no better than counselling only or remedial phonics only; but professional counselling only was better than remedial phonics only or no intervention and equal to professional counselling plus remedial phonics.

Somerset (2)
Counselling by non-professionals plus remedial teaching was no better than remedial teaching alone.

Somerset (3)
Counselling by non-professionals plus remedial teaching was better than remedial teaching alone.

Somerset (4)
The two 'therapeutic' conditions (counselling by non-professionals plus DISTAR, drama plus DISTAR) did not differ, and were better than DISTAR only and no intervention, which also did not differ.

Overall
Counselling by *a professional* alone was effective in study 1.

Counselling by non-professionals *plus* a specific reading intervention (remedial teaching, DISTAR respectively) was effective in studies 3 and 4, but seemed less so in study 2 (where the reading intervention was also remedial teaching).

Drama teaching designed to boost self-esteem plus DISTAR was effective in study 4.

The specific reading intervention alone was ineffective in study 3, reasonably effective in study 4 (though no more so than no intervention, and significantly less so than the therapeutic interventions), and not particularly effective in study 1 (where it was phonics) and study 2.

Normal classroom provision was ineffective in study 1.

Finally, the conclusions just stated are distinctly less strong than the claims made by Lawrence. Nevertheless, he did show that self-esteem counselling by non-professionals plus a specific reading intervention can be effective, and that the boost to self-esteem can also (and perhaps more cost-efficiently) be delivered through appropriate drama teaching. The need for motivational factors in poor reading to be re-explored is heightened by the recurrent anxiety over boys' low achievement compared to girls, and the possibility that part of the reason may be boys' negative attitudes to reading and writing.

References
Bandura (1977); Canfield and Wells (1976); Engelmann *et al.* (1969); Lawrence (1971, 1972, 1973, 1985, 1988); Rogers (1975)

3.18 SPELLIT

Scheme

SPELLIT stands for Study Programme to Evaluate Literacy Learning through Individualized Teaching. It was a research and development project funded by the D*f*ES, the Community Fund, WHSmith, and the Dyslexia Institute. Within the programme of work were a number of distinct, but interconnected projects concerned with providing support for young children experiencing difficulties in literacy learning. The main aims of these projects were:

- to provide a scientific evaluation of structured multi-sensory teaching
- to explore ways of supporting parents to enable them to help their children learn
- to work in partnership with Local Education Authorities in order to help disseminate good practice.

SPELLIT's programme of activities included:

- Development and publication of new learning programmes and materials
- Development and delivery of structured programmes of support for parents to support their children at home
- Delivery of training courses to participating schools
- Observational study of children in schools and in individual teaching sessions
- Interview and feedback meetings with participants – pupils, parents, tutors, classteachers
- Production of Practical Guidance and Case Studies to inform wider educational practice
- Exploration of the application of the support programmes with other groups who are socially disadvantaged by literacy difficulties.

The pupils involved were aged about 7 at the start of the study, and took part in the programme over Y2-3 or Y3-4. There were three different learning programmes:

- Structured multi-sensory teaching using the Dyslexia Institute's approach, twice weekly over a 24-week period in sessions each lasting one hour – this was in effect the 'experimental' condition

- A Home Support Programme consisting of activities and exercises to be done at home for around 15 minutes per day, for 5 days per week over a 30-week period an 'alternative intervention'

- A Combined programme involving 1 hour per week of structured teaching for 24 weeks and Home Support Activities in 15-minute sessions, 3 times per week over a 30-week period.

There was also a No Teaching 'control' group of children who received no additional support but went on to receive a programme involving structured teaching later.

Evaluation

The programme was evaluated by its developers at the Dyslexia Institute in York. For reasons beyond the researchers' control, the Combined programme did not operate as planned, and provided no data, leaving the experimental, control and alternative intervention (Home Support) groups. The control group made poor progress, falling further behind standard rates of progress. The Home support group progressed at exactly the standard rate. The experimental group made more progress than this, but not significantly.

References

Rack and Hatcher (2002a, in press, b, in press)

3.19   THRASS (Teaching Handwriting, Reading and Spelling Skills)

THRASS is a structured multi-sensory literacy programme, devised by Alan Davies, which teaches children about letters, speech sounds and spelling choices. It is divided into the three main areas of handwriting, reading and spelling. It increases understanding of the way the English language is structured, with 44 phonemes, of which 20 are vowel sounds and 24 are consonant sounds. Children learn immediately that the same sound can be represented by different letters or groups of letters (graphemes), eliminating any confusion.

Scheme
THRASS was developed by Alan Davies, an educational psychologist then at Manchester Metropolitan University. The programme has been continuously developed and revised, and in 1997 became available on computer.

Davies found that the problem many people have whilst learning to read and write is that there are 44 sounds or phonemes in most well-known accents of English, yet only 26 letters to represent them. Therefore the central feature of the scheme is that children are taught explicitly about the variety of grapheme-phoneme and phoneme-grapheme correspondences of English. Teachers are given training in the use of materials (video, workshops, audio cassettes and an instruction booklet). A typical THRASS lesson might include identifying upper and lower case letters by name, and writing each letter while listening to verbal instructions. Children are introduced to common sequences such as days of the week or seasons. During each lesson new learning is introduced but there is always practice of material already covered. Children are encouraged to work together whilst the teacher provides positive encouragement and reinforcement for correct responses.

Evaluation
Though THRASS has been extensively studied both in the UK and in Australia, almost all the work has considered its use as an initial, across-the-board scheme, and there is little evidence on its value as a 'catching-up' intervention. In the first edition, the entry on THRASS was based on an investigation mounted in Sheffield in 1994-95 (Johnson, 1995) – but the version of THRASS used in that evaluation has long been superseded. For this edition, that account has been replaced by data from the only piece of more recent research on THRASS as a 'catching-up' intervention, namely the 'Special Initiative to Enhance Literacy Skills in Bridgend' conducted there with pupils in Y3-6 in 1998. Both reading and spelling were assessed. The results showed considerable impact on reading for all year groups, and on spelling in Y3, but not on spelling in Y4-6, where the children made standard progress.

References
Johnson (1995); Matthews (1998)

Contacts
THRASS (UK) Ltd
Units 1-3 Tarvin Sands
Barrow Lane, Tarvin
Chester CH3 8JF

http://www.thrass.co.uk

3.20   Time for Reading

Do volunteers really make a difference to children's progress in reading? This intervention was designed as part of a research project investigating this question.

Scheme
Time for Reading was designed and evaluated by staff at the University of Sunderland. The evaluators describe it as

> 'vary[ing] from other volunteer studies in several ways:
> (i)   the project was designed to operate with children of 4 and 5 years of age whose schools were located in areas of particular socio-economic disadvantage. It was hypothesised that work with younger children from such communities might avoid the negative impact of an experience of reading failure;
> (ii)   The focus of the volunteer inputs concerned the development of the children's phonological awareness, letter knowledge, the gaining of experience of a wide range of reading related behaviours and the heightening of the enjoyment of stories. All of these, it was considered, were likely to be particularly important at the early stages of reading acquisition;
> (iii) A detailed training programme was supplemented by a specially prepared manual for volunteers and teachers … It was considered that the provision of detailed training and guidance might help to overcome a major reason for the failure of the intervention evaluated by Loenen's (1989) study;
> (iv) ongoing supervision, monitoring and management was provided by the research team in collaboration with senior staff in the volunteer schools.'
> (Elliott *et al.*, 2000, p.232)

The intervention took place over six months when the children were aged 4-5. A total of 31 volunteers worked with 68 children.

Evaluation
Testing was conducted at the beginning and end of the intervention (two phonological awareness tests, of rhymes and initial phonemes), and then again 2½ years later ('3-year follow-up') (reading and spelling), when the children were aged 7-8 (Y3). Both the participating children and a randomly-assigned control group were studied. There were no significant differences between the groups in either test used either at the end of the intervention or at the 3-year follow-up.

The evaluators put forward several possible reasons for the non-significant results. The length of the intervention may have been too short; the children may not have received the most appropriate reading programme;   the programme may have been 'overly comprehensive'; the tutors may not have grasped important aspects of book sharing; the children with most need often received less input because the tutors found it difficult to persevere with them; many volunteers did not gain a sophisticated understanding of the programme; and 'liaison between teachers and volunteers proved rather superficial'.

The findings need to be judged alongside those from the Knowsley Reading Project and various Better Reading Partnerships, where volunteers did make a difference. However, a

systematic review of randomised controlled trials on volunteers helping children to read (Torgerson *et al.*, 2003, in press) found only a modest and non-significant effect size (0.19).

References
Elliott *et al.* (2000); Torgerson *et al.* (2003, in press)

B.   Schemes involving ICT

3.21 AcceleRead AcceleWrite (formerly Jersey Computer Assisted Reading Development Programme)

(1)   Jersey

In 1993, the education authority in Jersey read about the success of 'Talking Computers' trialled by Martin Miles in Somerset and realised that it would be possible to replicate the study at little cost. Jersey schools already had the appropriate computers along with a good relationship with the software publisher. The level of computer literacy amongst Jersey teachers meant that the training to use the computer element of the programme was readily achievable. All pupils progressed beyond expectation, not just in reading but in self-esteem, computer skills and self-correction.

Scheme

Seventy-one pupils with reading difficulties from 15 primary schools and four secondary schools took part. Each school supplied a project Co-ordinator. Courses were run to train the adults involved in how to use the computer software and the process of delivering the reading material.

The supervising adults worked with an individual child for 20 minutes a day over a four-week period. The child was presented with a card containing four sentences. Each card contained a particular phonic pattern or number of patterns. The child was allowed to read the card until confident of memorising it. The card was then placed face down and the pupil had to say the sentence to the adult, then type it into the computer. The computer said each word as it was entered, giving audio feedback of misspelled words. It also read the complete sentence once the full stop had been typed. Mistakes were rectified by the student until the sentence was completed correctly.

Evaluation

The evaluation appears to have been devised by Martin Miles, but it is not clear from the report who actually implemented it. The pupils were tested for reading, spelling and recall of digits using subtests of the British Ability Scales. Questionnaires about the pupils' levels of motivation and ability in reading, comprehension, spelling, creative writing, speaking and listening were completed by teachers both before and after delivery of the reading programme. A questionnaire aiming to evaluate the feeling of the supervising adults about the overall effectiveness of the scheme was also completed.

Results were available from 61 children in the experimental group (and from nine children in a control group - but this group was too small for the results to be reliable). The children using the program made substantial gains. It was noted that pupils' motivation was high and they were relaxed. There were noticeable improvements in computer skills and confidence. An improved awareness of letter patterns, fluency in reading, employment of letter strategies and independence in their writing were reported. During the same period positive changes in pupil behaviour (which were not the prime focus of the project) were also reported.

Reference
Jersey Advisory Service (1993)

Contacts

Dr Martin Miles	or	Talking Systems
mmilesep@aol.com		22 Heavitree Road
		Exeter
		Devon
		EX1 2LQ
		talksystem@aol.com

(2)   Devon

More recently, Martin Miles used the program, now known as AcceleRead AcceleWrite, in a primary school in Devon with 30 mostly older KS2 children who had been identified as experiencing difficulties with reading and/or spelling. Remarkable gains were made.

Reference
Unpublished data supplied by Martin Miles

3.22   Integrated Learning Systems: the National Council for Educational Technology study

Integrated Learning Systems is the general name for a number of computer-based learning systems, and for initiatives using them. The study summarised here is the one mounted by the National Council for Educational Technology (NCET, now the British Educational Communications and Technology Agency, BECTa) and evaluated by teams led by staff of the University of Leicester (phases I and II) and NFER (phase III).

Scheme

NCET provided a number of schools with both hardware and educational software designed to facilitate independent learning by pupils. The two main systems evaluated were SuccessMaker and Global Learning Systems. Both have extensive programs for both literacy and numeracy. Most schools in the study used the numeracy packages; only about half used the literacy packages. The literacy packages provided massive amounts of guided practice on all aspects of literacy; they also gave feedback on errors, and kept track of pupils' progress for their own and their teachers' benefit.

Evaluations

(1)   Mainstream

As stated, Phase II of the project (1994-96) was evaluated by a team led by staff of the University of Leicester, and Phase III (1996-97) by NFER. (Phase I was smaller, though its conclusion for literacy was much the same as the later phases'.) Phase III was not covered in the first edition because it did not involve any children in Years 1-4, but is included now because of its Y5 group. Though the results of Phases II and III for numeracy were reasonable, those for literacy were unimpressive: no overall benefit compared to control groups. Two exceptions were noted in Phase II: a Special School where *children in the control group outperformed those in the experimental group*, and a mainstream primary school where the ILS group did significantly better than the control group, but only after teachers switched off all but the comprehension modules of the computer program - see section 2.6.

References
BECTA (1998); NCET (1994, 1996); Underwood (1997); Underwood and Brown (1997); Underwood *et al.* (1994)

(2)   For pupils with low attainments in reading

> Within the NCET study special attention was paid to pupils with special educational needs (Lewis, 1995) and other under-achievers (Gardner, 1995), on the hypothesis that they might particularly benefit from the technology. Lewis (1999) reviewed this work, including her own study, and several other UK interventions using ILS with pupils with low attainments in reading. Because of small sample sizes and the existence of Lewis's thorough review and analysis, none of the data from these studies are reproduced or analysed here. Lewis concluded (p.156, emphasis original) that teachers' beliefs about the value of Integrated Learning Systems for pupils with special educational needs or learning difficulties are not supported by the findings

from the major UK evaluations. These, as noted in the final report (BECTa, 1998), do not indicate that pupils were learning *nothing* from Integrated Learning Systems but rather that whatever they were learning was not being transferred to the outcome measures used in these studies. Overall, the use of Integrated Learning Systems was not conspicuously more effective than other approaches in terms of these outcome measures …

In other words, the value of this technology for these pupils is still very much Not Proven – just as with mainstream pupils. Lewis (personal communication, 2002) confirmed that no further evidence had since emerged to modify her conclusion.

For example, under the EPPI programme, Andrews *et al*. (2002) undertook a mapping review of the impact of ICT on literacy learning in English for 5- to 16-year-olds, plus a systematic review of the impact of *networked* ICT on literacy learning for ages 5 to 16. The quality of the 188 mapped studies was extremely variable, and the 16 studies of networked ICT and literacy learning were too heterogeneous to allow meta-analysis. We requested a search within the library of 188 studies for studies which were British and conducted with primary-age pupils: there were 17. Of these, just three dealt specifically with pupils with low attainments in literacy. Two had sample sizes of 10 and nine, too small to be reliable; the third could not be obtained in time to be analysed. The value of ICT for poor readers and writers at primary level still requires to be demonstrated, except in carefully targeted exceptions such as AcceleRead AcceleWrite.

References
Andrews *et al*. (2002); Gardner (1995); Lewis (1995, 1999)

3.23   RITA (Reader's Intelligent Teaching Assistant)

This section needs to be read in conjunction with section 3.7 on Interactive Assessment and Teaching (IA&T).

Scheme
RITA is a computer-based literacy support system. Its authors describe it as follows:

> Interaction takes place in the HyperCard 2.3 environment on the Apple Macintosh computer … Output includes pictures, graphics and high quality 'synthesised speech' or digitised speech in addition to text.
>
> The teacher can specify 'activities', from a single session to a 'workbook' comprising a whole course unit… A single 30-minute session can … be pre-programmed as a series of up to three 'activities'… The teacher may elect to be present at any or all of these activities. For a more autonomous learner [it] is possible to let the computer take over the scheduling …, or to allow the learner some responsibility for deciding what to do next.
> (Nicolson *et al*., 1999, p.197)

Much of RITA was IA&T computerised.

Evaluation
The developers also evaluated the program, in a carefully designed experiment with Y2 and Y3 children comparing RITA against both ordinary classroom teaching (no-intervention control group) and IA&T. RITA was implemented in the same schools as IA&T had been (the previous year?), using the same teachers, same timetable and equivalent children. The data for IA&T were those from the same team's earlier evaluation of IA&T, used here as an 'historical alternative intervention' group. Reading and spelling were tested at the beginning and end of the intervention. The control group made hardly any gain in either reading or spelling. Both RITA and IA&T made significant gains in both areas, but the gains did not differ in either. The researchers stress the interpretation that this shows that RITA was just as effective as IA&T; an alternative view would be that the technology added nothing.

Reference
Nicolson *et al*. (1999)

C.   Schemes for improving writing

In this section we consider the only well-researched scheme for improving writing more generally, Paired Writing, plus some less rigorous data from Family Literacy.

3.24 Paired Writing

Scheme
Paired Writing is another in the suite of innovations devised and researched by Keith Topping and colleagues. He describes it as follows:

> Paired Writing … is a framework and set of guidelines to be followed by pairs working together to generate a piece of writing for a purpose. It gives a supportive structure to scaffold interactive collaborative behaviours through all stages of the writing process…
>
> [It] consists of
>
>> 6 STEPS
>> +
>> 10 Questions   (Ideas)
>> 5 Stages   (Drafting)
>> 4 Levels   (Editing)
> (Topping, 2001, pp.141, 144)

As with Cued Spelling, Topping stresses that Paired Writing 'is a lot simpler than it looks'. And again as with Cued Spelling and Paired Reading, children are provided with a flowchart as an aide-memoire – this is downloadable from the website. On each occasion in each pair, one child has the task of writing ('the writer'), while the other supports ('the helper').

Evaluations
Topping and colleagues have carried out three reasonably well designed and executed, though small, research studies on Paired Writing.

Sutherland and Topping (1999) studied two groups of 16 children in P4 (equivalent to Y3) in one Scottish primary school, with two equivalent groups of 16 in the same classes in the same school who did not receive Paired Writing training. One experimental group had helpers ('tutors') of the same ability (and swopped roles at intervals), the other had helpers of different ability (and did not swop roles). The cross-ability group made a significant gain during the intervention, while the same-ability group did not (at least in absolute terms – this group's control group's post-test score declined so much that the same-ability group's post-test score was significantly better).

Nixon and Topping (2001) studied 10 writers in one P1 class (equivalent to Reception) in a Scottish primary school, plus 30 of their classmates as an 'unpaired' comparison group, and 10 helpers from P7 (equivalent to Y6) in the same school. The 10 writers had a significantly higher average score for writing at pre-test than the comparison group, but had also made a much more significant gain at post-test.

Yarrow and Topping (2001) studied 13 children in one P5 class (equivalent to Y4) in a Scottish primary school, plus 13 of their classmates as a comparison group. The experimental group contained both writers and helpers; their data are analysed together (as the 'Interaction' group) in the Appendix because the groups would otherwise be too small. The Interaction group made significantly more gain than the No Interaction group.

Doing research on reading is difficult enough; research on writing is much more difficult, and in its infancy. Topping and his colleagues have made a start, and others should follow their lead. The three studies combined have a total sample size of 55 children in experimental groups and 75 in control/comparison groups. Other approaches need to be evaluated, but from this evidence Paired Writing looks at least promising.

References
Nixon and Topping (2001); Sutherland and Topping (1999); Topping (1995, 2001); Yarrow and Topping (2001)
www.dundee.ac.uk/psychology/TRWresources

3.25 Family Literacy

Other evaluations besides Paired Writing have gathered data on children's improvement in writing, but less rigorously. For instance, Moore and Wade (1998) reported on a long-term follow-up of 121 former Reading Recovery children from seven schools in Australia and six in New Zealand, plus an equal number of matched comparison children from the same schools. The children were between 10 and 12 years old, and the former Reading Recovery children had experienced the initiative when they were six. The former Reading Recovery children wrote slightly longer texts and made significantly fewer errors in spelling, punctuation and grammar per 100 words, and the overall quality of their writing was rated significantly higher.

Closer to home, the evaluation of the Basic Skills Agency's Family Literacy Demonstration Programmes (see section 3.4) included assessments of the emergent or early writing of all the children in the study – 362 at the outset (when the children's ages ranged between 3:00 and 6:11) and smaller numbers at the end of the courses and at three follow-up points. The assessments were made on a seven-point scale which was empirically derived from analysis of the over 900 scripts involved. The children made significant gains, which were judged by the evaluators to be better than would have been expected.

References
Brooks *et al.* (1996a, 1997); Gorman and Brooks (1996)

CHAPTER 4
COMMENTS ON THE EVALUATIONS

4.1   Schemes not included

A number of schemes on which information was gathered or available were not included in the analysis, for various reasons:

-    A further study in the Somerset self-esteem series (Lawrence and Blagg, 1974) was too small to be included. Six children in an experimental group received counselling from non-professionals plus board games designed to enhance reading; six in a first alternative intervention (AI) group received counselling from non-professionals only; six in a second AI group received the board games only; and six in a third AI group played non-reading games (there was no no-intervention control group). The results were consistent with Somerset (3 and 4): counselling from non-professionals plus playing the board games was significantly better than the other three conditions, which did not differ. Also, only the experimental group made a significant gain in reading age (p=0.01);

-    ARROW (Aural-Reading-Respond-Oral-Written), designed by Lane (1978, 1980), has been widely researched, mainly in Australia, but no research on it was found which was British, focused on reading improvement, and conducted with poorer readers at primary level;

-    'Pause, Prompt and Praise': There have been at least 20 studies on this approach, but the only one located in which the subjects were lower-attaining British readers of a relevant age consisted of a very brief report in Wheldall and Colmar (1990) of a study with 10 experimental pupils and 10 in an alternative intervention group, too small for inclusion this time (though it was included in the first edition);

-    The Early Reading Research is in widespread use in Essex, principally as an initial across-the-board scheme. It is also in limited use as a catch-up intervention, but no systematic data on this were available;

-    Literacy Acceleration (Lingard, 1993, 1997) has gathered copious data but exclusively at secondary level. Other schemes which are exclusively secondary are Corrective Reading, Corrective Spelling, and Reading Together;

-    A number of non-linguistic (medical or physiological) approaches have attracted attention, some for many years (e.g. eye-patching, also known as ocular occlusion), others more recent (e.g. coloured lenses or overlays, movement programmes). An interesting review of such approaches by Dr Angela Fawcett of the Department of Psychology at the University of Sheffield can be found on the D*f*ES SEN website:

     http://www.dfes.gov.uk/sen

     Enter 'dyslexia' as the search term; Review 2 is the relevant one.

Several schemes on which quantitative research has been carried out with children of relevant ages in the UK could not be included because of problems with the data, such as

- reporting only raw scores, either with no control group results at all, or without the standard deviations necessary to calculate an effect size

- internal inconsistencies which could not be resolved from the information available.

Also, Success for All could not be included because no British evaluation data were yet available. In the USA, this programme has been well researched and shown to be highly effective, and it is being implemented in several areas in the UK, especially Nottingham City and Nottinghamshire.

A number of other widely used schemes were not included because, again, no evaluation data could be found, either in the research literature or via direct approach to scheme originators or users. A list of these schemes is given in the Acknowledgments (pp.v-vi).

The project at Saint Lawrence School was singled out for special comment in the first edition. It was designed and implemented by the school itself, which also carried out and reported its own quantitative evaluation – a rare and worthwhile example of teacher-led research. Two other such initiatives, which also could not be included in this edition, deserve the same accolade: a scheme implemented in South Gloucestershire promising 'a flying start in phonics', and Underattainers Can Achieve! at Great Crosby Catholic Primary School in Sefton.

4.2   The quality of the research

Having reviewed the Paired Reading literature, Topping and Lindsay (1992, p.201) commented, with academic restraint: 'The quality of studies was extremely varied.' The literature surveyed for the first edition of this report varied from the meticulous to the very weak. The most meticulous was the only randomised controlled trial, the original Cumbria Reading with Phonology Project, now called Reading Intervention, though there were several other well-conducted experiments. The number of studies excluded from the analysis then (for whatever reason) was considerably larger than the number retained. Most of the excluded studies provided no quantitative data at all, and many of those which did provide such data were unusable, either because of basic design faults (too few subjects, same test used both pre and post over too short an interval, etc.) or because the data were unclear (averages did not correspond with the individual scores, too many children scored zero or maximum, etc.).

The trawl for this edition was more focused and did not produce so much unusable information. There was still some, however. The major deficiencies this time were again inadequate sample sizes, the use of unstandardised instruments, and failure to provide data from which an impact measure could be calculated.

This is not to say that those studies which we have included necessarily told us everything we needed to know. The frequency of the phrase 'not stated' in the Appendix shows how much information was missing, sometimes even from the reports of quite large-scale, independently-funded evaluations.

Three particular problems arose from the tests used in the 25 studies (the tests are listed in the Appendix). First, some of the tests were old even when used in the relevant studies.

Secondly, most of the tests provided only reading/spelling age data and not standardised scores. Though apparently easier to interpret, reading and spelling ages are statistically unsatisfactory - for example, establishing whether a gain in test scores is statistically significant is more difficult for reading and spelling ages than for standardised scores. Reading and spelling age data do allow the calculation of the ratio gain - but this is in itself not a very useful statistic, especially for low-attaining groups. Pupils in such groups might not be expected to make a month's gain in reading or spelling age in one calendar month, for perfectly valid reasons. Standardised scores allow much more direct comparisons of amount of gain. Ratio gains have nevertheless been used in much of this analysis because for most of the studies they were the only impact measure which could be calculated.

Thirdly, for many of the tests used it was impossible to calculate effect sizes, which are statistically much more satisfactory than ratio gains. If a standardised test is used, an effect size can be calculated even in the absence of an explicit control group; but if a non-standardised test is used then an effect size can be calculated only if control group data, including the standard deviation, are reported.

More generally, it is noteworthy that the evaluations included here number just 25. As pointed out in chapter 1, they represent more studies and approaches than that, because for Paired Reading and Parental Involvement in particular one study stands for many, and because several studies contained control groups and/or alternative interventions. But the total is still not impressive, and it is to be hoped that anyone currently devising an intervention will automatically consider the necessity for, and commission, an evaluation. The Government is setting an example here, with its evaluations of Summer Literacy Schools and of the National Literacy Strategy.

4.3   Recommendations

Whenever an educational innovation is devised and tried out, an evaluation should be commissioned.

All evaluations should be based on the gathering of quantitative attainment data, and the data should come from the use of standardised tests, and not non-standardised instruments such as reading- and spelling-age tests.

Properly defined control groups should be set up, through random assignment or at least by matching.

All evaluations should report *as a minimum* the following information:

- the date when the evaluation was carried out (in addition to the date of reporting)

- the exact age-range of the children involved

- salient characteristics of the children involved, for example whether they had special educational needs

- the numbers of children in the experimental and control groups and in any alternative intervention groups

- how children were assigned to the different groups, for example randomly or by matching

- the nature of any alternative intervention

- the exact length of the intervention

- the reading, spelling and any other tests used

- the pre- and post-test average *standardised* scores and standard deviations for every group involved; this would make it strictly speaking unnecessary to report the amount of gain, but this might be interesting in itself

- the statistical significance of the differences between groups at pre-test, so that the initial equivalence of the groups can be shown, or the statistical handling of any significant pre-test differences can be explained

- the statistical significance of the differences between pre-test and post-test scores for each group, so that it can be seen whether or not the *absolute* value of any gains was statistically significant; this is easier for standardised scores than for reading/spelling ages

- the statistical significance of the differences between groups at post-test, or of the differences between their gains, so that the *relative* impact of different interventions (including no intervention) can be seen

- the effect size, so that the impact of the approach can be compared with others.

In order to carry out worthwhile evaluations which provide this information, it may well be necessary to build the cost of the evaluation into the funding of the initiative itself, and this in turn may cause problems for those seeking funds for initiatives. But if evaluations are not properly funded and reported, the current situation will be perpetuated, where initiatives are mainly evaluated on the 'feel-good' factor of participants.

In future, full information ought to be provided, so that non-participants can judge the effectiveness of intervention schemes more objectively, and so that funders (whether government or private) can see that resources have been deployed responsibly and usefully.

REFERENCES

Andrews, R., Burn, R., Leach. J., Locke, T., Low, G. and Torgerson. C. (2002). A systematic review of the impact of networked ICT on 5- to 16-year-olds' literacy in English (EPPI-Centre Review). In: *Research Evidence in Education Library*. Issue 1. London: EPPI-Centre, Social Science Research Unit, University of London Institute of Education.

Bandura, A. (1977). *Social Learning Theory.* Englewood Cliffs, NJ: Prentice Hall.

BECTa (1998). *The UK ILS Evaluations. Final report.* Coventry: British Educational Communications and Technology Agency.

Bentley, D. and Reid, D. (1995). *Supporting Struggling Readers.* Widnes: UK Reading Association.

Brooks, G., Gorman, T.P., Harman, J., Hutchison, D. and Wilkin, A. (1996a). *Family Literacy Works: the NFER Evaluation of the Basic Skills Agency's Family Literacy Demonstration Programmes.* London: Basic Skills Agency.

Brooks, G., Cato, V., Fernandes, C. and Tregenza, A. (1996b). *The Knowsley Reading Project: Using Trained Reading Helpers Effectively.* Slough: NFER.

Brooks, G., Gorman, T.P., Harman, J., Hutchison, D., Kinder, K., Moor, H. and Wilkin, A. (1997). *Family Literacy Lasts: the NFER Follow-up Study of the Basic Skills Agency's Demonstration Programmes.* London: Basic Skills Agency.

Brooks, G., Flanagan, N., Henkhuzens, Z. and Hutchison, D. (1998). *What Works for Slow Readers? The Effectiveness of Early Intervention Schemes*. Slough: NFER.

Brooks, G., Harman, J., Hutchison, D., Kendall, S. and Wilkin, A. (1999). *Family Literacy for New Groups: The National Foundation for Educational Research Evaluation of the Basic Skills Agency's Programmes for Linguistic Minorities, Year 4 and Year 7.* London: Basic Skills Agency.

Brooks, G. and Hutchison, D. (2000). Final report to Basic skills Agency and Worcestershire LEA on Innovative Development Project. Slough: NFER. (mimeograph)

Brooks, P. and Weeks, S. (1999). *Individual Styles in Learning to Spell: Improving Spelling in Children with Literacy Difficulties and All Children in Mainstream Schools.* D*f*EE Research Report no.108. London: D*f*EE.

Canfield, J. and Wells, H. (1976). *100 Ways to Enhance Self-Concept in the Classroom.* Englewood Cliffs, NJ: Prentice Hall.

Clay, M.M. (1979). *Reading: the Patterning of Complex Behaviour.* Portsmouth, NH: Heinemann.

Clay, M.M. (1985). *The Early Detection of Reading Difficulties: a diagnostic Survey with Recovery Procedures.* Third edn. Auckland, New Zealand: Heinemann.

Clay, M.M. (1993). *Reading Recovery: a Guidebook for Teachers in Training.* Auckland, New Zealand: Heinemann.

Clipson-Boyles, S. (2000). 'The *Catch Up Project*: a reading intervention in Year 3 for Level 1 readers.' *Journal of Research in Reading*, **23**, 1, 78-84.

Collins, K. (1996). The Better Reading Partnership Progress Report. Bradford: Bradford Inspection and Consultancy Services. (mimeograph)

Department for Education and Skills (2002). *Research and Development in England: Background Report prepared for the OECD Review*. London: D*f*ES website.

Derrington, S. (2001a). Early Literacy Support Pilot: evaluation report. Bristol: Bristol City LEA Achievement Division. (mimeograph)

Derrington, S. (2001b). [Name of school suppressed]: report on the use of Phono-Graphix™ as a method of helping struggling readers at Key Stage 2. Bristol: Bristol City LEA Achievement Division. (mimeograph)

Elliott, J., Arthurs, J. and Williams, R. (2000). 'Volunteer support in the primary classroom: the long-term impact of one initiative upon children's reading performance.' *British Educational Research Journal*, **26**, 2, 227-44.

Engelmann, J., Osborn, E. and Engelmann, J. (1969). *Distar: an Instructional System.* Chicago: Science Research Associates.

Fawcett, A.J. (2002). Evaluating therapies excluding traditional reading and phonological based therapies. A review for the Department for Education and Skills, the British Dyslexia Association and the Dyslexia Institute. Review 2, February 2002. http://www.dfes.gov.uk/sen Enter 'dyslexia' as the search term and find Review 2.

Fawcett, A.J., Nicolson, R.I., Moss, H., Nicolson, M.K. and Reason, R. (1999). 'Early reading intervention can be effective and cost-effective.' *British Journal of Educational Psychology*, **69**, 1, 47-62.

France, L., Topping, K.J. and Revell, K. (1993). 'Parent-tutored Cued Spelling.' *Support for Learning*, **8**, 1, 11-15.

Fudge, P. (2001). Reading Recovery in Bristol: an evaluation of the first two years of the initiative, 1999-2001. Bristol: Bristol LEA. (mimeograph)

Gardner, J. (1995). *ILS Phase 2 evaluation: ILS and under-achievers; research report to NCET*. Coventry: NCET.

Gardner, J., Sutherland, A. and Meenan-Strain, C. (1997). *Reading Recovery in Northern Ireland: the First Two Years.* Belfast: The Queen's University of Belfast, School of Education.

Glass, G.V., McGaw, B. and Smith, M.L. (1981*). Meta-Analysis in Social Research.* Beverley Hills, CA: Sage.

Gorman, T.P. and Brooks, G. (1996). *Assessing Young Children's Writing: a Step by Step Guide.* London: Basic Skills Agency.

Greenway, C. (2002). 'The process, pitfalls and benefits of implementing a Reciprocal Teaching intervention to improve the reading comprehension of a group of Year 6 pupils.' *Educational Psychology in Practice*, **18**, 2, 113-37.

Hannon, P. (1987). 'A study of the effects of parental involvement in the teaching of reading on children's reading test performances.' *British Journal of Educational Psycholoogy*, **57**, 1, 56-72.

Hannon, P. and Jackson, A. (1987). *The Belfield Reading Project: Final Report.* London: National Children's Bureau, and Rochdale: Belfield Community Council.

Hatcher, P. (2000). 'Sound links in reading and spelling with discrepancy-defined dyslexics and children with moderate learning difficulties.' *Reading and Writing: An Interdisciplinary Journal*, **13**, 257-72.

Hatcher, P.J., Hulme, C. and Ellis, A.W. (1994). 'Ameliorating early reading failure by integrating the teaching of reading and phonological skills: the phonological linkage hypothesis.' *Child Development,* **65,** 1, 41-57.

Hewison, J. (1985). 'Parental involvement and reading attainment: implications of research in Dagenham and Haringey.' In: Topping, K.J. and Wolfendale, S. (eds) *Parental Involvement in Children's Reading.* Beckenham: Croom Helm.

Hewison, J. (1988). 'The long term effectiveness of parental involvement in reading: a follow-up to the Haringey Reading Project.' *British Journal of Educational Psychology*, **58**, 184-90.

Hurry, J. (1996). 'Literacy programmes - how can we choose?' *The Primary English Magazine,* September/October, 24-6.

Hurry, J. and Sylva, K. (1998). The Long-term Effects of Two Interventions for Children with Reading Difficulties. London: QCA.

Iversen, A. and Tunmer, W. (1993). 'Phonological processing skills and the Reading Recovery program.' *Journal of Educational Psychology,* **85,** 1, 112-26.

Jersey Advisory Service (1993). The Jersey Computer Assisted Reading Development Programme. St. Saviour: Jersey Advisory Service. (mimeograph)

Johnson, M. (1995). The Handwriting, Reading, and Spelling Sequence (THRASS): an Evaluation of a Two Term Pilot Study September 1994-April 1995. Sheffield: City of Sheffield Education Department. (mimeograph)

Johnson, M., Phillips, S. and Peer, L. (1999). *A Multi Sensory Teaching System for Reading.* Manchester: Manchester Metropolitan University.

Lane, C.H. (1978). 'The "Arrow" approach for aural rehabilitation.' *The Volta Review,* **80,** 149-54.

Lane, C.H. (1980). The improvement of listening and speech skills in language disordered children. Unpublished M.Ed. dissertation, University of Exeter.

Lawrence, D. (1971). 'The effects of counselling on retarded readers.' *Educational Research,* **13,** 2, 119-14.

Lawrence, D. (1972). 'Counselling of retarded readers by non-professionals.' *Educational Research,* **15,** 1, 48-5 1.

Lawrence, D. (1973). *Improving Reading through Counselling.* London: Ward Lock.

Lawrence, D. (1985). 'Improving self-esteem and reading.' *Educational Research,* **27,** 3, 195-200.

Lawrence, D. (1988). *Enhancing Self-Esteem in the Classroom.* London: Paul Chapman.

Lawrence, D. and Blagg, N. (1974). 'Improved reading through self initiated learning and counselling.' *Remedial Education,* **9,** 2, 61-4.

Lewis, A. (1995) *ILS Phase 2 evaluation: Children with special needs; research report to NCET.* Coventry: NCET.

Lewis, A. (1999). 'Integrated Learning Systems and pupils with low attainments in reading.' *British Journal of Special Education*, **26**, 3, 153-7.

Lingard, T. (1993). The Acquisition of Literacy at Secondary School. Unpublished Ph.D. thesis, University of Exeter.

Lingard, T. (1997). 'Making phonics work: its role within Literacy Acceleration.' *Reading,* **31,** 3, 43 -9.

Lowe, K.J. (1995). 'Factors affecting students' progress in reading: key findings from a longitudinal study.' *Literacy, Teaching and Learning*, **1**, 2, 57ff. Reprinted in S.L. Swartz and A.F. Klein (eds) *Research in Reading Recovery*. Portsmouth, NJ: Heinemann.

Matthews, D. (1998). Special Initiative to Enhance Literacy Skills in Bridgend, Spring 1998. Bridgend: Bridgend County Borough Council Special Needs Services. (mimeograph)

McGuinness, C. and McGuinness, G. (1998). *Reading Reflex: the Foolproof Phono-Graphix™ Method for Teaching Your Child to Read.* New York: Free Press.

McGuinness, D. (1997). *Why Our Children Can't Read: And What We Can Do About It.* New York: Free Press.

McGuinness, D. (1998). *Why Children Can't Read and What We Can Do about It (and Why Teacher Training Doesn't Help).* British edn. London: Penguin.

Moore, M. and Wade, B. (1998). 'Reading Recovery: its effectiveness in the long term.' *Support for Learning*, **13**, 3, 123-8.

Morgan, R.T.T. (1976). '"Paired Reading" tuition: a preliminary report on a technique for cases of reading deficit.' *Child Care, Health and Development,* **2**, 13-28.

Morgan, R.T.T. (1986). *Helping Children Read.* London: Methuen.

Munn, P. and Ellis, S. (2001). Report on the Longitudinal Evaluation of Reading Recovery in Northern Ireland. Glasgow: University of Strathclyde Department of Primary Education. (mimeograph)

National Council for Educational Technology (1994). *Integrated Learning Systems: a Report of the Pilot Evaluation of ILS in the UK.* Coventry: NCET.

National Council for Educational Technology (1996). *Integrated Learning Systems: a Report of Phase II of the Pilot Evaluation of ILS in the UK.* Coventry: NCET.

Nicolson, R.I., Fawcett, A..J. and Nicolson, M.K. (1999). 'Evaluation of a computer-based reading intervention in infant and junior schools.' *Journal of Research in Reading*, **23**, 2, 194-209.

Nixon, J.G. and Topping, K.J. (2001). 'Emergent writing: the impact of structured peer interaction.' *Educational Psychology*, **21**, 1, 41-58.

Nugent, M. (2001). 'Raising reading standards – the Reading Partners approach: cross-age peer tutoring in a special school.' *British Journal of Special Education*, **28**, 2, 71-9.

Palincsar, A.S. (1982). Improving the reading comprehension of junior high school students through the reciprocal teaching of comprehension-monitoring strategies. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Palincsar, A.S. (1986). 'The role of dialogue in providing scaffolded instruction.' *Educational Psychology*, **21**, 73-98.

Palincsar, A.S. and Brown, A.L. (1984). 'Reciprocal Teaching of comprehension-fostering and comprehension-monitoring activities.' *Cognition and Instruction*, **1**, 2, 117-72.

Prance, J. (1992). *Reading Recovery: Early Intervention for At-Risk Six Year Old Children.* Addlestone: Runnymede Centre.

Rack, J. and Hatcher, J. (2002a, in press). *SPELLIT: Summary, Case Studies and Written Guidance.* (Pack for teachers). London: D*f*ES.

Rack, J. and Hatcher, J. (2002b, in press). *SPELLIT: Report of the Evaluation.* D*f*ES Research Report. London: D*f*ES.

Reason, R. and Boote, R. (1994). *Helping Children with Reading and Spelling: A Special Needs Manual*. London: Routledge.

Rogers, C.R. (1975). 'Empathy – an unappreciated way of being.' *Counseling Psychology,* **5,** 2, 2-9.

Rosenshine, B. and Meister, C. (1994). 'Fostering literacy learning in supportive contexts.' *Review of Educational Research*, **64**, 4, 479-530.

Schweinhart, L.J., Barnes, H.V. and Weikart, D.P. (1993). *Significant Benefits: the High Scope Perry Preschool Study through Age 27* (Monographs of the High/Scope Educational Research Foundation, 10). Ypsilanti, MI: High/Scope Press.

Stepien, D. (1997). Family Literacy in Hampshire: a regional application of the Basic Skills Agency family literacy programme. Hampshire County Council. (mimeograph)

Sutherland, J.A. and Topping, K.J. (1999). 'Collaborative creative writing in eight-year-olds: comparing cross-ability fixed role and same-ability reciprocal role pairing.' *Journal of Research in Reading*, **22**, 2, 154-79.

Sylva, K. and Hurry, J. (1995a). *The Effectiveness of Reading Recovery and Phonological Training for Children with Reading Problems: Full Report.* London: SCAA.

Sylva, K. and Hurry, J. (1995b). *Early Intervention in Children with Reading Difficulties: an Evaluation of Reading Recovery and a Phonological Training.* SCAA Discussion Papers: No.2. London: SCAA.

Taylor, C. (1999). Derbyshire Better Reading Partnership: Summary Report, 1998-1999. Matlock: Read On – Write Away! (mimeograph)

Tizard, J., Schofield, W.N. and Hewison, J. (1982). 'Collaboration between teachers and parents in assisting children's reading.' *British Journal of Educational Psychology,* **52,** 1-15.

Topping, K.J. (1990). Outcome evaluation of the Kirklees Paired Reading Project. Unpublished Ph.D. thesis, University of Sheffield.

Topping, K.J. (1995). *Paired Reading, Spelling and Writing. The Handbook for Teachers and Parents.* London and New York: Cassell.

Topping, K.J. (2001). *Thinking Reading Writing. A Practical Guide to Paired Learning with Peers, Parents and Volunteers.* London and New York: Continuum, with accompanying website:
www.dundee.ac.uk/psychology/TRW

Topping, K.J. and Lindsay, G.A. (1992). 'Paired reading: a review of the literature.' *Research Papers in Education,* **7,** 3, 199-246.

Topping, K.J., Nixon, J., Sutherland, S. and Yarrow, F. (2000). 'Paired Writing: a framework for effective collaboration.' *Reading*, **34**, 2, 79-89.

Topping, K.J. and Wolfendale, S. (Eds) (1985). *Parental Involvement in Children's Reading.* Beckenham: Croom Helm.

Torgerson, C., King, S. and Sowden, A.J. (2003, in press). 'Do volunteers in schools help children learn to read? A systematic review of randomised controlled trials.' *Educational Studies*, **28**.

Underwood, J. (1997). 'The role of Integrated Learning Systems.' TOPIC, Issue 18, Item 10.

Underwood, J. and Brown, J. (eds) (1997). *Integrated Learning Systems: Potential into Practice.* Oxford: Heinemann.

Underwood, J., Cavendish, S., Dowling, S., Fogelman, K. and Lawson, T. (1994). *Integrated Learning Systems in UK Schools: Final Report.* Leicester: Leicester University, School of Education.

Watt, J.M. and Topping, K.J. (1993). 'Cued Spelling: a comparative study of parent and peer tutoring.' *Educational Psychology in Practice*, **9**, 2, 95-103.

Wheldall, K. and Colmar, S. (1990). 'Peer tutoring for low-progress readers using "Pause, Prompt and Praise".' In: Foot, H.C., Morgan, M.J. and Shute, R.H. (eds) *Children Helping Children.* Chichester: John Wiley.

Wilson, J. (1993). *Phonological Awareness Training Programme.* London: Educational Psychology Publishing.

Wilson, J. and Frederickson, N. (1995). 'Phonological awareness training: an evaluation.' *Educational and Child Psychology,* **12,** 1, 68-79.

Wolfendale, S. and Topping, K.J. (eds) (1996). *Family Involvement in Literacy: Effective Partnerships in Education.* London: Cassell.

Worsley, J. (2001). The Catch Up programme: a report of the pilot Catch Up project in eighteen Norfolk First and Primary schools – November 2000-July 2001. Norwich: The Norfolk Psychological Service in partnership with The Norfolk Education Advisory Service. (mimeograph)

Wright, A. (1992). 'Evaluation of the first British Reading Recovery programme.' *British Educational Research Journal,* **18,** 4, 351-68.

Yarrow, F. and Topping, K.J. (2001). 'Collaborative writing: the effects of metacognitive prompting and structured peer interaction.' *British Journal of Educational Psychology*, **71**, 261-82.

Yuill, N. and Oakhill, J. (1988). 'Effects of Inference Awareness Training on poor reading comprehension.' *Applied Cognitive Psychology,* 2, 33-45.

APPENDIX
DETAILS OF THE EVALUATIONS

The central part of this Appendix is a log of the 25 schemes, in alphabetical order. Under each scheme are listed the salient statistical and related data used in the analysis in this report. Before the log, abbreviations and the nomenclature of school years are explained, and the organisation of the entries is described; and that description is followed by a number of notes of clarification.

After the log of the 25 interventions, the basis on which schemes have been compared is explained. The final section gives brief conclusions from follow-up data from 10 studies.

Abbreviations:	AI	alternative intervention
	BASWRT	British Ability Scales Word ReadingTest
	c.a.	chronological age
	conts	controls
	exps	experimentals
	m	months
	N	sample size
	n/a	not applicable
	ns	non-significant
	r.a.	reading age
	s.a.	spelling age
	s.d.	standard deviation
	stand.	standardised
	RG	ratio gain

Key to school years:

Label of school year

in England and Wales	in Scotland and Northern Ireland	in North America	Age of pupils (in years)
Reception	P(rimary) 1	Pre-kindergarten	4-5
Year 1	P(rimary) 2	Kindergarten	5-6
Year 2	P(rimary) 3	1st grade	6-7
Year 3	P(rimary) 4	2nd grade	7-8
Year 4	P(rimary) 5	3rd grade	8-9
Year 5	P(rimary) 6	4th grade	9-10
Year 6	P(rimary) 7	5th grade	10-11

A.1   Introduction to the evaluation data
The entries below are organised, as far as possible, in the order shown in Table A.1.

Table A.1: Organisation of entries in log of studies

	See Note
Name of intervention	
Main reference(s)	
Date when it was implemented	
Age-range of children involved, in school years (Y2, etc)	
Type of children involved	1)
Number of pupils in experimental group	
Number of pupils in control group, where there was one	
Number of pupils in alternative intervention group, where there was one	
Nature of alternative intervention	
For each group, number of schools, where known	
Whether groups were equivalent	2)
Length of intervention in weeks	
Reading and/or spelling test(s) or writing assessment used	3, 4)
For each group (where known), pre- and post-test average scores, and units in which these are stated	5)
For each group (where known), difference between pre- and post-test average scores ('gain') in relevant units	6)
For each group, where scores are reading ages (r.a's), ratio gain (RG), stated to one decimal place	7)
Effect size (where this was known or could be calculated), stated to two decimal places	8)
Statistical significance of differences between pre- and post-test scores, and between experimental, control and alternative intervention groups, where known	9)

71

Notes to Table A.1:
One general note needs to be given first: the term 'reliable' is used throughout this Appendix in its everyday and not in its technical sense.

1) **Type of children:** categorised as one of
> SEN - having special educational needs
> Low (reading and/or spelling or writing) attainment, which will in many cases include children with SEN

> Mixed ability - though this still means that the group studied was underachieving, on average, by national standards.

2) **Equivalence of groups:** Eight of the studies (The Catch Up Project (pilot and national), Inference Training, MTSR, Parental Involvement, Reading Intervention (original), Reading Recovery, Somerset (all four phases), RITA) had alternative intervention groups. These were all properly constituted, in the sense that they were created by assigning pupils to groups randomly or by matching, and their results are treated as fully reliable in this analysis.

Eighteen studies had no-intervention control groups: BRP in Bradford and Worcestershire, The Catch Up Project (pilot and national studies), Family Literacy in Hampshire, Individual Spelling, Cued Spelling (both studies), IA&T, Paired Reading, Parental Involvement, PAT, Reading Intervention (original, and for statemented children), Reading Recovery in London and Surrey, Somerset (1, 4), SPELLIT, Time for Reading, AcceleRead AcceleWrite in Jersey, ILS, RITA, Paired Writing (all 3 studies). However, in a few cases (AcceleRead AcceleWrite in Jersey, BRP in Bradford, Cued Spelling (1)) it is clear that the control group had been chosen in a way which did not make it properly equivalent to the experimental group (it seemed usually to consist of children who happened to be available). Also, the control group in Family Literacy in Hampshire was so small that its data are ignored. Data from the rest of these control groups are logged below, and where a ratio gain can be calculated it is taken as reliable, since RGs are always calculated independently for each group, and hence can in a sense be seen as *absolute* statistics. However, results from these control groups are not relied upon for *comparisons* with the experimental groups in the same studies.

In all but one of the remaining control-group studies, the control groups were set up in ways which did make their results reliable as comparisons for the experimental groups.

The exception is ILS. From the report of that study it is clear that there were control groups, but the details of their selection, and even the number of children involved, are unclear. Nevertheless, given the provenance of the study, it is assumed for the purposes of this analysis that the control groups were properly constituted, and comparative statistics from them are used where possible.

3) **Choice of tests to report:** Almost all these studies used more than one instrument to measure impact, and most used several. Only reading and spelling test and writing assessment results have been analysed here, on the grounds that the main focus of this enquiry is interventions designed to boost literacy attainment. Some reading tests yield more than one score (for example, the Neale Analysis gives scores for reading accuracy and reading comprehension); where this is so, both sets of data have been given. All other reading tests cited have been classified as giving measures of reading accuracy.

4) **Range of tests used:** In all, at least 25 different reading tests were used in the studies under consideration. The tests are listed in Table A.2, with the number of studies in which they were used. The number of spelling studies was smaller (12), and they used only seven tests between them; these are listed in Table A.3. Each of the writing studies analysed used a different form of assessment – for details see the separate entries in this Appendix.

Some of the tests listed are rather elderly, especially the Burt and the Schonell.
This may limit the reliability of some of the findings.

Table A.2: Reading tests used in the 21 reading studies

Test	No. of studies
British Ability Scales Word Reading Test (BASWRT), various forms	5
Burt, various editions	4
Carver Word Recognition and Phonic Skills (WRAPS) Test	2
Customised tests within computer programs, yielding US grade levels	1
Hodder and Stoughton Literacy Baseline	2
Hodder and Stoughton Reading Progress Tests	3
Holbom	1
Linguistic Awareness and Reading Readiness (LARR)	1
Macmillan Graded Word Reading	1
Macmillan Individual Reading Analysis	1
Neale Analysis, various editions and forms	9
New Macmillan Reading Analysis	1
NFER Reading Test A	1
NFER-Nelson Individual Reading Analysis	1
NFER-Nelson Progress in English 9	1
NFER-Nelson Reading Ability Series	1
Peabody Individual Achievement Test	1
Primary Reading Test	1
Salford, various forms	4
Schonell	4
Southgate	1
Standard (Daniels & Diack) Reading Tests	1
Suffolk	3
Wechsler Objective Reading Dimension (WORD) Scales	3
Widespan	1
(not stated)	2

Table A.3: Spelling tests used in the 12 spelling studies

Test	No. of studies
British Ability Scales Spelling Test	3
Graded Spelling Test (Daniels and Diack)	1
Schonell	4
Single Word Spelling Test	1
Vernon	3
Wechsler Objective Reading Dimension (WORD) Scales	2
Young's Parallel Spelling Tests	1

5) The units in which average scores and s.d's are stated are almost always either reading/spelling ages or standardised scores, very rarely both. Raw scores were used in a few cases, e.g. Somerset (4).

6) Where the units of measurement are r.a's/s.a's, gain is given in months of r.a./s.a.

7) **Ratio gain (RG):** This is defined by Topping and Lindsay (1992, p.201) as 'the gain in Reading age made by a subject on a reading test during a chronological time span, expressed as a ratio of that time span; that is, ratio gain equals reading age gain in months divided by chronological time in months'. For a group, this can be stated as the formula

$$\frac{\text{(average reading age in months at post-test) - (average reading age in months at pre-test)}}{\text{time elapsed in months}}$$

(The definition and formula are obviously applicable to spelling too.)

Thus calculating an RG does not require data from a control group - but where any non-experimental group is present its RG can and should be calculated too; and the RGs for non-experimental groups are valid for those groups independently of whether they were properly equivalent to the experimental group, because each RG is calculated for the relevant group of subjects without reference to any other group.

The dispersal of scores (as shown in the standard deviation) is ignored in RGs – only the average reading ages at pre- and post-test and time elapsed are used. RG is therefore a statistically unsophisticated device; but, as Topping and Lindsay further point out, using raw gains instead 'renders the highly heterogeneous literature very difficult to summarise'. Also, since over half of the evaluations surveyed here used reading ages as their reporting units (see below) it seemed appropriate to use RGs in attempting to make comparisons between those interventions.

However, RGs do take account of the length of time over which an intervention achieves its impact - as shown in the formula, this is done by dividing the gain in months of reading age (r.a.) by the number of months between pre- and post-test.

RGs were explicitly stated in only a few reports. However, the r.a. data required for calculating RGs were given in many more. Some reports did not use tests which yielded r.a's/s.a's, and therefore RGs could not be calculated for them.

8) **Effect size:** This is a more statistically based metric. It involves dividing the difference between the gain made by the experimental group and that made by the control group by a relevant standard deviation, and the result is expressed as a decimal of an s.d. Therefore the top line of the relevant formula can be stated as:

(average gain of experimental group) - (average gain of control group)

This part of the formula can be applied equally to r.a's, s.a's, standardised scores and raw scores. The problems arise with the choice of control groups and of the appropriate s.d. to use as the divisor.

Where there is a control group but it is not properly equivalent to the experimental group, its data cannot be reliably used in these calculations. In this analysis, where a control group was not properly constituted, generally no attempt has been made to calculate an effect size from its data. An exception has been made for the Paired Writing studies.

Even where the control group was properly set up the choice of s.d. to use as the divisor in the formula is still problematic, since authorities in the field seem to advocate or adopt different approaches:

(a) Glass, McGaw and Smith (1981) seem to recommend the s.d. of the control group's post-test scores (e.g. p.114), and counsel against pooling it with any other s.d. (p.106).

(b) Topping and Lindsay (1992, p.211) give a formula using the s.d. of the control group's *gain*. Very few reports state this statistic, but it can be calculated from the s.d's of the control group's pre- and post-test scores, where both are given. However, effect sizes based on this statistic measure something rather different from those based on the other forms of s.d. mentioned here, and (depending on the size of the correlation between pre- and post-test scores) such effect sizes may be overestimated by a factor between about 1.3 and 2.2.

(c) NCET (1996, pp.6, 10) in the ILS evaluation used a slightly different formula, with the divisor being (apparently) the arithmetic mean of the pre-test s.d's of experimental and control groups.

(d) Where the control group's post-test s.d. is significantly larger than that group's pre-test s.d., the post-test s.d. may be inappropriate as a measure of the variance of the population of interest. In these circumstances, the control group's protest s.d. may be the appropriate divisor.

In practice, for this report, much of the debate over which s.d. to use became irrelevant, since only a few reports stated more than one form of s.d. (and some stated none at all). The following policy on effect sizes was therefore adopted:

- where an evaluation report did not give the data needed to calculate an effect size, we state that 'Effect sizes were not given and could not be calculated';

- where an evaluation report stated an effect size it was accepted, however calculated;

- where no effect size was given but the information to calculate it, according to any of the above methods using a reliable control group s.d., was given, it was calculated.

There were a few studies in which, though effect sizes were not stated and could not be calculated, they would clearly have been close to zero and non-significant, because the gain made by an experimental or alternative intervention group was not significantly greater than that made by the control group. Where this was the case, it has been stated and the information has been taken as reliable.

In a few cases, an effect size has been reported where there was no control group and the authors of the study did not report one. These are all studies which used standardised

tests, and where such a test is used, there is always an implicit control group, the one provided by the standardisation sample. In these circumstances the absence of an explicit control group, or of its data, can be circumvented, since an effect size can be calculated by using the s.d. (usually 15.0) and mean scores of the standardisation sample; and since the mean scores of the standardisation sample are by definition the same at pre- and post-test, the control group term in the formula reduces to zero, and the formula simplifies to:

$$\frac{\text{(average gain of experimental group in standardised score points)}}{15}$$

where the term above the line can be expanded to

(average standardised score of experimental group at post-test) -
(average standardised score of experimental group at pre-test)

While much more sophisticated statistically than RGs, effect sizes (however calculated) take no account of the length of time over which an intervention achieved its impact (though they could be modified to do so). But they do take account of the dispersal of scores (through the s.d.) and of a control group.

9) **Statistical significances:** Two forms of statistical significance data would be relevant, where available, namely on the gains of separate groups (difference between pre- and post-test average scores), and on the differences between gains where there was more than one group.

When the gains of separate groups are tested for significance, the fact that children are older by the time of the post-test has of course to be allowed for. Where standardised tests are used, the standardised scores provide for this automatically. Where r.a. tests are used, allowing for age is more complicated, but can still be accomplished. It is remarkable that within the studies analysed here, hardly any stated the significance of gains within groups. No attempt was made to calculate such statistics for other studies. The absence of such statistical information seems very remiss. It is particularly bothersome where there was neither a reliable control group nor average standardised scores, because then the importance of the result has to be judged 'by eye' from the RG – which was the case in the majority of studies.

Statistical significances of the differences between gains were given (or implied) in several cases, but by no means universally.

**A.2	Data on evaluated schemes**

**1	Better Reading Partnerships**

**(1)	Bradford**

Main reference: Collins (1996)

Date: May 1995-April 1996

Age-range: Y1-7? ('First and Middle schools')

Type of children: Low attainment

N of experimental group: 510(?) in 30(?) schools – information in report unclear

N of control group: not stated

Equivalence of groups: not stated

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 10

Reading test: Suffolk, 1st edition

Pre- and post-test average scores and s.d's: not stated

Gains (months of r.a.) and ratio gains:

	gain	RG
experimentals, First schools	5.9	2.4
experimentals, Middle schools	6.2	2.5
controls	2.1	0.8

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

Follow-up: Children in (apparently) Middle schools made a further gain of 3.7 months of r.a. in 2-3 months post programme; further RG = 1.5; total gain = 9.8 months (in 5 months); total RG = 2.0.

1   Better Reading Partners

(2)   Derbyshire

Main reference: Taylor (1999)

Date: 1998-1999 (though data collected in other years too)

Age-range: Y1–6 (also Y7–8, not reported here)

Type of children: Low attainment

N of experimental group: 683 (for year groups, see below)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 11 (2.6 months used as divisor for RGs)

Reading test: Salford (mainly)

Pre- and post-test average raw scores and gains: not stated

Gains (in months of r.a.) and ratio gains:

	N	Gain	RG
Y1	20	12.30	4.6
Y2	184	9.7	3.7
Y3	120	8.5	3.3
Y4	133	9.0	3.5
Y5	107	8.0	3.1
Y6	90	8.4	3.2

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

1   Better Reading Partnerships

(3)   Co. Durham

Main reference: Unpublished data supplied by Ann Foster

Date: 1999-2001

Age-range: Y1-5

Type of children: Low attainment

N of experimental group: 237 (for year groups, see below)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 10

Reading tests: (Y1–2) Neale Individual Reading Analysis; (Y3–5) Suffolk, 1st edn

Ns, pre- and post-test and 3–month and 12–month follow-up r.a's (s.d's not stated), gains since previous test (in months of r.a.) and ratio gains post vs pre:

Age	N	Pre	Post	Gain	RG	3-month follow-up	Gain	12-month follow-up	Gain
Y1 acc	39	6:0	7:6	18	7.2	7:7	1	8:11	16
comp	34	6:6	7:5	11	4.4	8:0	7	8:4	4
Y2 acc	57	6:6	8:0	18	7.2	7:11	-1	8:9	10
comp	46	6:2	8:1	23	9.2	8:4	3	8:11	7
Y3	48	7:2	7:10	8	3.2	8:0	2	8:8	8
Y4	39	7:6	8:1	7	2.8	8:4	3	8:9	5
Y5	51	9:0	9:5	5	2.0	9:6	1	10:3	9

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

1   Better Reading Partnerships

(4)   Redcar and Cleveland

Main reference: Unpublished data supplied by Andrew Taylor

Date: 1997-2002

Age-range: Y1–6

Type of children: Low attainment

N of experimental group: 1071

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 10

Reading test: Salford

Pre- and post-test average r.a's and s.d's: not stated

Average gain (in months of r.a.): 8

Ratio gain: 3.2

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

1   Better Reading Partnerships

(5)   Worcestershire

Main Reference: Brooks and Hutchison (2000)

Date: 1999-2000

Age-range: mainly Y1-2; some Y4-6

Type of children: Mixed-ability

Ns of experimental groups: (phase 1) 146; (phase 2) 142

Ns of control groups: (phase 1) 142; (phase 2) 146. This was a cross-over design in which the controls from phase 1 received the intervention in phase 2 and thus became a second experimental group, while the phase 1 experimentals became their controls. The phase 2 data from the 1st experimental group are therefore also follow-up data.

Equivalence of groups: pre-test differences handled statistically

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 17

Reading test: Suffolk

Pre- and 1st and 2nd post-test average r.a's and s.d's, gains (in months of r.a.), ratio gains and significances:

Group	Pre-test (October 1999) Mean (s.d.)		1st Post-test (February 2000) Mean (s.d.)		2nd Post-test (July 2000) Mean (s.d.)		Gain	RG	Significance
A	7:3	(1:1)	7:9	(1:5)			6	1.5	
B	7:0	(1:0)	7:3	(1:2)			3	0.9	p<0.007
B			7:3	(1:2)	7:9	(1:5)	6	1.2	
A			7:9	(1:5)	8:0	(1:6)	3	0.8	p<0.011

Pre- and 1st and 2nd post-test average standardised scores, s.d's and gains, effect sizes and significances:

Group	Pre-test Significance (October 1999) Mean (s.d.)	1st Post-test (February 2000) Mean (s.d.)	2nd Post-test (July 2000) Mean (s.d.)	Gain	Effect size	
A	89.5 (9.5)	93.9 (9.6)		4.4		
					0.13	ns
B	88.6 (11.0)	91.6 (11.0)		3.0		
B		91.6 (11.0)	96.4 (11.4)	4.8		
					0.34	p<0.001
A		93.9 (9.6)	95.4 (10.1)	1.5		

Follow-up: The phase 2 (follow-up) data from group A (taking the r.a. and standardised score data together) show that that group continued to make approximately standard progress. They were not making any further relative gain, but were maintaining the gain made in phase 1.

2   The Catch Up Project

(1)   The Pilot Study

Main reference: Clipson-Boyles (2000)

Date: September-December 1997

Age-range: Y3

Type of children: Low attainment (level 1 in reading in Key Stage 1 test)

N of experimental group: 74; 17 in subsample matched to controls

N of control group: 17

N of alternative intervention group: 14

Nature of alternative intervention: 'Teachers were asked to spend time equivalent to The Catch Up Project with selected pupils'

Equivalence of groups: Three of the experimental schools were selected, then matched as closely as possible with two other sets of three schools; then pupils in control, alternative intervention and experimental groups were chosen by the same method in all cases (six pupils in each who had achieved level 1 in reading in Key Stage 1 test)

Length of intervention in weeks: 10

Reading test: Hodder & Stoughton Literacy Baseline

Pre- and post-test average scores and s.d's, average gains and s.d's, all in months of r.a., ratio gains, and effect sizes calculated using differences in average gain over control group's pre-test s.d.:

	pre-test ave.	(s.d.)	post-test ave.	(s.d.)	gain ave.	(s.d.)	RG	effect size
experimentals - all	78.3	(6.0)	84.8	(7.5)	6.5	(5.3)	2.6	*
- in matched schools	79.6	(4.3)	88.2	(6.2)	8.6	(5.9)	3.4	0.78
controls	81.0	(9.6)	82.1	(7.7)	1.1	(6.5)	0.4	
matched time group	77.1	(4.5)	80.6	(8.2)	3.5	(5.4)	1.4	0.25

* The effect size for all experimentals is not given because it would be based on a non-equivalent control group

Ratio gain: n/a

Statistical significances: not stated

2   The Catch Up Project

(2)The National Experimental Study

Main reference: Unpublished data supplied by Julie Lawes

Date: 1999–2000

Age Range: Y3

Type of children: Low attainment (level 1 in reading in KS1 test)

N of schools: 98. The schools were randomly selected from a national database. Three parallel groups were created taking into account variables such as number on roll, free school meals entitlement, ESL, etc. Once the schools chosen as experimental agreed to participate they were sent a pack and offered no further support

N of experimental group: 34

N of control group: 43

Number in alternative intervention group: 46

Nature of alternative intervention: 'Teachers were asked to spend time equivalent to The Catch Up Project with selected pupils'

Equivalence of groups of pupils: not stated

Length of intervention in weeks: 35

Reading test: Hodder Reading Progress Test Series

Pre- and post-test average r.a's and s.d's: not stated

Gains (in months of r.a.) and ratio gains:

	Gain	RG
experimentals	11.5	1.4
controls	8.0	1.0
matched time group	10.2	1.3

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

2   **The Catch Up Project**

(3)   **Cornwall**

Main reference: Unpublished data supplied by Julie Lawes

Date: 1999–2000

Age Range: Y3

Type of children: Low attainment (level 1 in reading in KS1 test)

N of experimental group: 53 in 28 schools. Schools were targeted that had not been successfully awarded funding (staffing) for the ALS (Additional Literacy Support) programme but whose pupils were nevertheless considered to need additional support with literacy.

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 35

Reading test: Hodder Reading Progress Test Series

Pre- and post-test average r.a's and s.d's: not stated

	Gain	RG
Gain (in months of r.a.) and Ratio gain:		
	14.8	1.9

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

2   The Catch Up Project

(4)   Wakefield

Main reference: Unpublished data supplied by Julie Lawes

Date: 1999-2000

Age Range: Y3/Y4

Type of children: Low attainment (level 1/2C in reading in KS1 test)

N of experimental group: 39 in 8 schools

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 35

Reading test: Hodder Reading Progress Test Series

Pre- and post-test average r.a's and s.d's: not stated

	Gain	RG
Gain (in months of r.a.) and Ratio gain:		
	14.5	1.8

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

2   The Catch Up Project

(5)   Norfolk

Main References: Worsley (2001), and unpublished data supplied by Julie Lawes

Date: 2000-2002

Age-range: Y2-6

Type of children: Low attainment. 'Children were selected … according to the following criteria:
- Children from Y2-6 who have literacy difficulties, particularly in reading
- Priority to be given to those pupils who are unlikely to be successful with group interventions
- Priority to be given to those pupils who also have low self-esteem.'

N of experimental group: 1284 in 43 schools (for breakdown, see below)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 17–44 (see below)

Tests: (reading) Salford Sentence Test, revised version 2000; (spelling) Young's Parallel Spelling Test A

Pre- and post-test average r.a's / s.a's and s.d's: not given

Area, date, length in weeks age-ranges, ns, gains in r.a. /s.a. and ratio gains: see below

Area	Date	Length in weeks	Age	N	Reading Gain (in months of r.a.)	RG	Spelling Gain (in months of s.a.)	RG
Thetford	2000-01	44	Y2-4	508	19.8	2.0		
Great Yarmouth	2000-01	30	Y6	544	9.6	1.4		
King's Lynn	2001-02	26	Y3-6	32	14.9	2.5		
Across county (1) 12 schools	2000-01	35	Y2-3	121	19.5	3.3	3.7	0.5
(2) 6 schools	2001	13	Y2-3	79	14.6	3.7		

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

3   **Cued Spelling (1)**

Main reference: France et al. (1993). Also summarised in Topping (2001) and Topping (1995).

Date: not stated (c.1991?)

Age-range: Y4

Type of children: Low attainment (less able spellers, tutored by parents)

N of experimental group: 22 in one class in one school

N of comparison group: 10 better spellers in same class

Equivalence of groups: not matched – comparison group were rest of class

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 6

Spelling test: Graded Spelling Test (Daniels and Diack, 1979)

Pre- and post-test average s.a's, gains (in months of s.a.), s.d's and ratio gains (source: France *et al*., 1993, Figure 2, p.14, lightly edited):

	Cued Spelling (n = 22)	Comparison Group (n = 10)
Mean pre-test spelling age in years	8.39 (1.08)	9.98 (1.56)
Mean post-test spelling age in years	8.90 (1.15)	10.16 (1.48)
Mean gain in spelling age in months	6.1 (6.1)	2.2 (7.9)
Ratio gain	4.1	1.6

Effect size: not stated, and could not be calculated because groups were not matched

Statistical significance: not stated

3   Cued Spelling (2)

Main reference: Watt and Topping (1993). Also summarised in Topping (2001) and Topping (1995).

Date: not stated (c.1991?)

Age range: Scottish Primary 5 (=Y4)

Type of children: Low attainment (less able spellers)

N of experimental group: 10, tutored by parents [also 4 tutored by peers, and 9 in Y4 – both groups too small for their results to be analysed here]

N of comparison group: 6 – very small, therefore results not analysed here

N of alternative intervention group: (no alternative intervention group)

Length of interventions in weeks: 16.5 (3.8 months)

Spelling test: Graded Word Spelling Test (Vernon, 1977)

Pre- and post-test average spelling ages and gains, s.d's and ratio gain (source: Watt and Topping, 1993, Table 1, p.100, edited to show only the relevant group):

Mean spelling age at pre-test (months)	105.6 (20.9)
Mean spelling age at post-test (months)	113.5 (23.9)
Mean gain in spelling age (months)	7.9 (5.2)
Ratio gain	2.1

Effect size: not stated, and could not be calculated because groups were not matched

Statistical significance: not stated

4 Family Literacy

(1) Basic Skills Agency's Demonstration Programmes

Main references: Brooks *et al.* (1996a, 1997)

Date: Summer 1994-Summer 1995 (1 cohort of children in each term)

Age-range: nursery to Y2 (ages 3-6), but reading data only on those in Y1 & 2 (ages 5 & 6)

Type of children: Low attainment

N of experimental group: total 392, but reading data for 147, on about 20 sites. Smaller numbers at each of the three follow-ups because calculations based only on children with complete data ('returners')

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 12

Reading test: Reading Recognition subtest of Peabody Individual Achievement Tests

Sample sizes, average standardised scores and s.d's at pre- and post-test and 12-week, 9-month and long-term follow-ups, gains from pre-test, and effect sizes post-test vs. pre-test calculated using s.d. of standardisation sample:

	N	average score	(s.d.)	Gain	Effect size
pre-test vs post-test	147	84.1 88.5	(17.0) (17.9)	4.4	0.29
pre-test vs 12 week follow-up	101	85.6 92.4	(17.6) (17.5)	6.8	
pre-test vs 9 month follow-up	67	84.2 90.3	(16.2) (18.1)	6.1	
pre-test vs long-term follow-up	107	89.6 93.6	(11.5) (15.2)	4.0	

Ratio gain: n/a

Statistical significances: p<0.05 for all differences from pre-test

4   Family Literacy

(2)   Hampshire

Main reference: Stepien (1997)

Date: 1996-97

Age-range: Reception

Type of children: Low attainment

N of experimental group: 27

N of control group: Control group mentioned (p.30) but insufficient data given for analysis

Equivalence of groups: not stated

N of alternative intervention group: no alternative intervention group

Length of intervention in weeks: 12

Reading test: LARR (Linguistic Awareness and Reading Readiness)

Pre- and post-test average standardised scores (s.d's not stated):

Pre	89.5
Post	118.1

Gain (standardised score points):   28.6

Ratio gain: n/a

Effect size calculated using s.d. of standardisation sample: 1.91

Statistical significance: not stated

4   Family Literacy

(3)   For New Groups

Main reference: Brooks *et al.* (1999)

Date: 1997 - 98

Age-ranges: (linguistic minorities) 3–6, but reading data reported here only on children of Y1 age; Y4

Type of children: Low attainment

N of experimental group: (linguistic minorities) 65; (Y4) 144

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 12

Reading test: (linguistic minorities) Hodder and Stoughton Literacy Baseline; (Y4) NFER–Nelson Progress in English 9

Pre- and post-test average standardised scores, r.a.'s and s.d's, gains, effect sizes calculated using s.d. of standardisation sample, and ratio gains:

	Pre	Post	Gain	Effect size	RG
Linguistic minorities					
Stand. Score	93.5	104.3	10.8	0.72	
(s.d.)	(16.9)	(14.8)			
r.a.	5:2	5:9	9m		3.0
Y4					
Stand. score	87.1	95.8	8.7	0.58	
(s.d.)	(14.5)	(16.4)			
r.a.			14m		4.7

Statistical significances: p<0.05

5   Individual Styles in Learning to Spell

Main reference: P. Brooks and Weeks (1999)

Date: 1996-99

Type of children: Mixed-ability but many with spelling problems

Age-range: Y2-3

Ns of experimental groups: (phase 1) 21; (phase 2) 26, all in 2 schools

Ns of control groups: This was a cross-over design, in which the group of 26 acted as a comparison group for the group of 21 in phase 1, then received the intervention in phase 2. As with BRP in Worcestershire, therefore, the phase 2 data for the first group are follow-up data.

Equivalence of groups: Not equivalent – chosen by teachers – but differences handled statistically (analysis of variance)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: (phase 1) 26; (phase 2) 22

Spelling test: Not stated

Average s.a's at pre-test and 2 post-tests, s.d's for pre-test and gains (other s.d's not stated), and gains (in months of s.a.) and ratio gains for each phase separately:

Group	N	Pre	1st post-test	Gain	RG	2nd post-test	Gain	RG
1	21	7:6 (0:11)	8:4	10 (6)	1.7	9:0	8 (4)	1.6
2	26	6:9 (0:10)	7:4	5 (4)	0.8	8:1	9 (6)	1.8

Effect sizes: were not given and could not be calculated.

Statistical significances: Both groups made significant gains in both phases. In phase 1 the experimental group made significantly more gain than the comparison group; in phase 2 the difference was ns. This was the predicted outcome: both groups made good gains while receiving the intervention, while the second group made only (slightly less than) standard progress in phase 1 before receiving it, and the first group continued to make better than standard progress in phase 2.

6   Inference Training

Main reference: Yuill and Oakhill (1988)

Date: Autumn 1985-Spring 1986

Age-range: Y3

Type of children: Mixed-ability

Ns of experimental groups: (1) 13 less skilled comprehenders
(2) 13 skilled comprehenders

N of control group: (no no-intervention control group)

Ns of alternative intervention (AI) groups: (AI1) 14; (AI2) 12, all in same 5 schools

Nature of alternative interventions: (AI1) comprehension exercises
(AI2) rapid decoding practice

Equivalence of groups: Groups matched on age, reading accuracy and vocabulary; experimental groups deliberately differentiated on reading comprehension

Length of intervention in weeks: 4

Reading test: Neale, form C at pre-test, form B at post-test, thus avoiding specific practice effects over the short interval

Pre- and post-test average scores (r.a's) and pre-test s.d's (post-test s.d's not given), gains (months of r.a.), and ratio gains:

Aspect of Neale	group	pre-test ave. r.a. (yrs & months)	(s.d.)	post-test ave. r.a. (months)		gain	RG
accuracy	exps 1	8:2	(0:6)	8:6	4.3	4.3	
	exps 2	8:3	(0:6)	8:7	3.9	3.9	
	AI 1	8:5	(1:2)	8:10	5.0	5.0	
	AI 2	8:4	(0:5)	8:7	3.0	3.0	
compre.	exps 1	7:3	(0:3)	8:8	17.4	17.4	
	exps 2	8:8	(0:8)	9:4	5.9	5.9	
	AI 1	8:1	(1:0)	8:11	9.6	9.6	
	AI 2	8:1	(0:7)	8:9	8.2	8.2	

Effect sizes: were not stated and could not be reliably calculated because there was no no-intervention control group

Statistical significances:

-    on accuracy, all differences in gains ns; equal gains expected because groups matched on this at outset and intervention not targeted on this skill

-	on comprehension, Inference Training more effective for less skilled comprehenders than for skilled comprehenders; Inference Training more effective than rapid decoding (AI2) for less skilled comprehenders; BUT Inference Training NOT more effective than comprehension exercises (AI 1)

7   Interactive Assessment and Teaching

For data on the computerised version of this approach, see RITA

Main References: Nicolson et al. (1999), Fawcett et al. (1999)

Date: not stated (1996–97?)

Age-range: Y2-3

Type of children: Low attainment

Ns of experimental groups: (Y2) 60; (Y3) 36

Ns of control groups: (Y2) 38; (Y3) 51

Equivalence of groups: control groups matched with experimental groups on reading level (Y2: bottom half of class; Y3: <90 on pre-test) and age

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 10

Tests: Wechsler Objective Reading Dimension (WORD) reading and spelling

Ns, pre- and post-test and 6-month follow-up average standardised scores and s.d's:

	Reading Standard Score				Spelling Standard Score			
	Pre	Post	Effect size post vs pre	Follow-up	Pre	Post	Effect size post vs pre	Follow-up
Y2								
Intervention N=60	89.03 (3.51)	92.76 (7.46)	0.94	89.93 (8.89)	84.26 (8.97)	91.67 (10.57)	0.95	91.53 (12.53)
Control N=38	89.08 (3.96)	88.87 (5.49)		87.03 (10.55)	83.70 (7.82)	85.58 (9.93)		86.18 (10.21)
Y3								
Intervention N=36	79.94 (3.41)	83.31 (3.61)	0.61	83.36 (5.23)	82.64 (5.87)	88.28 (5.11)	0.72	86.64 (6.55)
Control N=51	79.49 (5.55)	80.53 (7.01)		81.05 (10.46)	81.19 (7.87)	83.22 (8.22)		83.00 (8.02)

Ratio gain: n/a

Statistical significances: Experimentals made significantly greater gains than controls from pre to post in both reading and spelling. For Y2 the gains were maintained at follow-up in spelling, but almost completely lost in reading (the control group had slipped back even further). For Y3 the gains in reading were maintained, while the spelling gain had been partly lost.

8   Knowsley Reading Project

Main reference: Brooks et al. (1996b)

Date: 1994-95

Age-range: Y6

Type of children: Low attainment

N of experimental group: 302 in 13 schools (152 others in smaller groups in same schools not reported here)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Nature of alternative intervention: n/a

Length of intervention in weeks: 52

Reading test: Reading Ability Series C (pre-test), D (post-test)

Pre- and post-test average r.a's, standardised scores, gains in months of r.a. and standardised scores (s.d's not reported), ratio gain, and effect size calculated using s.d. of standardisation sample:

	Pre	Post	Gain	RG	Effect size
r.a.	8:11	10:7	20m	1.7	
stand. scores	92.7	97.9	5.2		0.35

Statistical significance: p<0.05

9   Multi-Sensory Teaching System for Reading (MTSR)

(1)   Initial study in 3 LEAs in North-West England

Main References: Johnson *et al*. (1999) and unpublished data supplied by Mike Johnson

Date: not stated

Age-range: Y2, Y5

Type of children: Low attainment

N of experimental group: (Y2) 25 in 4 schools; (Y5) 18 in 3 schools

N of control group: (no control group)

Ns and nature of alternative intervention groups:
(AI1) (Y2 & 5) Phonological Awareness Training – but numbers too small to analyse (6 in
            each year)
(AI2) (Y5) 'Beat Dyslexia' (no details available) N = 17 in 3 schools

Equivalence of groups: not stated, but appear to be non-equivalent opportunity samples in
different schools

Length of intervention in weeks: 8

Tests: (reading) Macmillan Individual Reading Analysis; (spelling) Vernon

Pre- and post-test average r.a's/s.a's and s.d's: not stated

Gains (in months of r.a./s.a.) and ratio gains:

			Gains	RG
Y2	reading	exps.	9.0	4.5
	spelling	exps.	4.2	2.1
Y5	reading	exps.	4.3	2.2
		AI2	3.4	1.7
	spelling	exps.	-7.3	-3.6
		AI2	6.7	3.4

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

9   Multi-Sensory Teaching System for Reading (MTSR)

(2)   Bolton

Main reference: Johnson *et al*. (1999) and unpublished data supplied by Mike Johnson

Date: not stated (2001-02?)

Age-range: Y2

Type of children: Low attainment

N of experimental group: 66 in 12 schools

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 26

Tests: (reading) Neale Analysis; (spelling) Single Word Spelling Test

Pre- and post-test average r.a's / s.a's and s.d's: not stated

Gains (in months of r.a. / s.a.) and ratio gains:

	Gain	**RG**
reading acc.	22.1	3.7
reading comp.	23.4	3.9
spelling	80.5	13.4

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

10   Paired Reading in Kirklees

Main reference: Topping and Lindsay (1992)

Date: 1984-87

Age-range: not stated but known to be across full school age-range (Y1-11)

Type of children: Mixed-ability

N of experimental group: 2372 in 155 projects in 71 schools for main accuracy measure - for other Ns, see below

N of control group: 446 in 37 projects for main accuracy measure - for other Ns, see below

Equivalence of groups: Method not stated, but thought to be matched groups

N of alternative intervention group: (some projects had alternative intervention groups, but too numerous and disparate to report here)

Nature of alternative interventions: (impractical to summarise)

Length of intervention in weeks: (average) 9

Reading tests: many, including Burt, Holborn, Neale, New Macmillan Reading Analysis, Primary, Salford, Schonell, Standard (Daniels & Diack) 1, Standard (Daniels & Diack) 12, Widespan

Pre- and post-test average scores and s.d's, and gains (not given in principal original report because too numerous):

Ratio gains:	accuracy		comprehension	
	N	**RG**	**N**	**RG**
all experimentals	2372	3.3	690	4.3
experimentals in control-group projects	580	3.4	170	4.6
controls in control-group projects	446	2.0	159	2.5

Effect sizes calculated using s.d. of control group gain:

	accuracy	comprehension
N of projects (N of subjects not given)	34	12
effect size	0.87	0.77

Also, mean effect size for published studies in the literature (12 controlled studies) is 2.12 (Topping, personal communication, 4 July 2002)

Statistical significances: all ratio gains were highly statistically significant (p<0.000) for both accuracy and comprehension (Topping, personal communication, 10 August 1998).

Follow-up: The Kirklees project provided follow-up data on 272 children in 17 projects. In follow-ups at less than 17 weeks after the end of the interventions, 102 children in 7 projects averaged RGs during the follow-up period of 2.0 for accuracy and 2.3 for comprehension. In follow-ups at more than 17 weeks, 170 children in 10 projects averaged RGs during the follow-up period of 1.2 for accuracy and 1.4 for comprehension.

11   Parental Involvement in Haringey

Main references: Tizard, Schofield and Hewison (1982), Hewison (1988)

Date: 1976-78

Age-range: Y2-3

Type of children: Mixed-ability

N of experimental group: 51 in 2 schools

N of main control group: 86 in same schools

N of alternative intervention (AI) group: 45 in 2 schools

Nature of alternative intervention: extra teacher help with reading

N of control group for alternative intervention: 66 in same schools as AI group

Equivalence of groups: The 4 schools were assigned at random to experimental and alternative intervention groups; then one Y2 class in each was chosen randomly to receive the experimental or alternative intervention; other Y2 classes became the controls; pre-test data showed that experimental and alternative intervention groups did not differ from their respective controls

Length of intervention in weeks: 104

Reading tests: (pre-test) Southgate; (post-test) NFER Reading Test A

Post-test average standardised scores and s.d's:

	ave. stand. score	(s.d.)
experimentals	104.2	(10.8)
main controls	95.0	(11.0)
AI	99.3	(16.6)
controls for AI	98.1	(13.7)

Gain: could not be stated because unrelated tests used pre and post

Ratio gain: n/a

Effect sizes calculated using differences between post-test average scores over *post*-test s.d's of control groups:

experimentals	0.84
AI group	0.09

Statistical significances: at post-test, experimentals significantly higher than main controls, alternative intervention group vs their controls ns

1-year follow-up, July 1979 (Tizard *et al*., 1982)**:**

Sample sizes. average scores on NFER Reading Test BD and s.d's:

	N	Average score	(s.d.)
Experimentals	66	99.0	(10.5)
Controls	78	91.6	(11.0)
AI (extra teacher help)	37	96.3	(12.3)
Controls for AI	58	92.9	(12.6)

Statistical significances: Experimentals were significantly better than their controls; the AI group and their controls did not differ.

3-year follow-up, July 1981 (Hewison, 1988)**:**

Sample sizes. average scores on London Reading Test (national norms) and s.d's:

	N	Average score	(s.d.)
Experimentals	41	101.0	(11.9)
Controls	69	94.5	(13.5)
AI (extra teacher help)	34	98.9	(14.3)
Controls for AI	56	97.3	(11.3)

Statistical significances: Experimentals were significantly better than their controls; the AI group and their controls did not differ.

12   Phono-Graphix™

(1)   Bristol

Main References: Derrington (2001 a, b) and unpublished data supplied by Sue Derrington

Date: 2000-02

Age-range: Y1-6

Type of children: Low attainment

N of experimental group: 230 in 13 schools – for year groups see below

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: Y1   26
                                            Y4–6   17
                                            Y2–6   12

Tests: (Y1)   Carver WRAPS tests
          (Y4–6) (reading) NFER-Nelson Individual Reading Analysis; (spelling) Vernon
          (Y2–6) (reading) Neale Analysis and Individual Analysis; (spelling) not stated

Pre- and post-test average r.a's/s.a's and gains (s.d's and some other data not stated) and ratio gains:

Age	Date	N		Reading accuracy				Reading comprehension			
			Pre	Post	Gain	RG		Pre	Post	Gain	RG
Y1	2000-01	141			13m	2.2					
Y4-6	2000-01	15	7:0	8:11	23m	5.8		7:11	9:4	17m	4.3
Y2-6	2001-02	74	6:4	8:5	25m	8.3		7:1	9:0	25m	8.3

Age	Date	N		Spelling			
			Pre	Post	Gain	RG
Y2-6	2001-02	60	6:9	7:7	10m	3.3
(subset of reading group above)						

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

12   Phono-Graphix™

(2)   Surrey

Main reference: http://www.readamerica.net - accessed 14/8/02

Date: 1999 -2000

Age-range: Y4

Type of children: SEN (children with dyslexia)

N of experimental group: 12 in one independent specialist school for children with dyslexia

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 26

Reading test: Macmillan Graded Word Reading

Pre- and post-test average r.a's (s.d's not given), gain (in months of r.a) and Ratio gain:

	Pre	Post	Gain	RG
	6:4	8:7	27	4.5

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

13   Phonological Awareness Training

Main reference: Wilson and Frederickson (1995)

Date: 1995?

Age-range: Y4-7 (Y1-3 also using programme but not included in evaluation)

Type of children: Special educational needs (all on at least Stage 2 of Code, including some severe learning difficulties and some mild learning difficulties)

N of experimental group: 24 in 3 schools

N of control group: 24 in 3 schools

Equivalence of groups: 'Allocation to the PAT programme or the comparison group... was made by the special needs coordinators... [They] were asked to try to ensure an even distribution between the PAT and comparison groups..., with comparable levels of reading difficulty and ... of special educational provision ...'

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 20

Reading test: BASWRT

Pre- and post-test average scores (r.a.'s) and pre-test s.d's (post-test s.d's not given), gains (months of r.a.), s.d's of gains, ratio gains, and effect size calculated using difference in average gains over control group's pre-test s.d.:

effect	pre-test average r.a. (yrs & months)	(s.d.)	post-test average r.a. (yrs & months)	gain ave (months)	(s.d.)	RG	effect size
exps	6:7	(0:6)	7:0	5.5	(3.7)	1.1	0.16
controls	6:9	(0:7)	7:1	4.4	(4.3)	0.9	

Statistical significance: difference between gains of experimental and control groups was significant, t=1.73, p<0.05. Control group made slightly less than standard progress; therefore, though experimental group made only slightly more than standard progress, this could be seen as a satisfactory result for these children with identified special educational needs.

14   Reading Intervention (formerly Cumbria Reading with Phonology Project)

(1)   (The original) Cumbria Reading with Phonology Project

Main reference: Hatcher, Hulme and Ellis (1994)

Date: September 1989-May 1990

Age-range: Y2 ('third year of infant schooling')

Type of children: Low attainment (reading quotient (r.a./c.a. x 1 00) on Carver test less than 86, but those with reading quotient less than 71 and percentile rank below 25 on Raven's Coloured Progressive Matrices (1965) excluded)

N of experimental group: 32 (received both reading programme and Phonological Training)

N of control group: 31

Ns of alternative intervention (AI) groups: (AI 1) 31; (AI2) 30

Nature of alternative interventions: (AI 1) reading programme only (similar to Reading Recovery);
(AI2) phonology only (Phonological Training)

Equivalence of groups: Groups matched on reading ability; other factors (IQ, age) treated as covariates in analysis of post-test differences

Length of intervention in weeks: 20 (but 25 weeks between start and end and 30 weeks between pre- and post-test; 7 months used as divisor in calculating RGs)

Tests: (reading) BASWRT form A, Neale revised form 1;
(spelling) Schonell Graded Word Spelling Test, List B

Pre- and post-test and 9-month follow-up average scores (r.a's/s.a's) and s.d's:
(N.B. This table replaces a less complete one in the first edition):

		Reading & Phonology (N = 32)	Reading alone (N = 31)	Phonology alone (N = 30)	Control (N = 31)
BAS word reading	pre	5.85 (0.53)	5.90 (0.47)	5.90 (0.57)	5.96 (0.53)
	post	6.73 (0.85)	6.56 (0.43)	6.55 (0.69)	6.60 (0.67)
Neale acc.	pre	5.10 (0.21)	5.04 (0.19)	5.18 (0.43)	5.11 (0.30)
	post	6.13 (1.00)	5.78 (0.54)	5.81 (0.90)	5.66 (0.80)
	follow-up	6.77 (1.58)	6.22 (0.82)	6.31 (1.03)	6.25 (1.15)
Neale comp.	pre	5.29 (0.30)	5.32 (0.34)	5.43 (0.50)	5.41 (0.49)
	post	6.39 (0.92)	6.00 (0.97)	5.94 (0.80)	5.88 (0.73)
	follow-up	6.99 (1.28)	6.47 (0.94)	6.46 (1.11)	6.35 (0.97)
Schonell	pre	5.78 (0.59)	5.83 (0.50)	5.93 (0.56)	5.77 (0.55)
	post	6.77 (0.93)	6.54 (0.55)	6.66 (0.63)	6.49 (0.74)
	follow-up	7.19 (1.02)	6.90 (0.62)	6.99 (0.82)	6.92 (0.78)

Gains (months of r.a./s.a.), ratio gains, and effect sizes calculated using pre-test s.d's of control group:

test	group	gain (months)	RG	effect size
BASWRT	exps	11	1.5	0.45
	conts	7	1.1	
	AI1	8	1.1	0.04
	AI2	8	1.1	0.02
Neale, accuracy	exps	13	1.8	1.60
	conts	5	0.9	
	AI1	9	1.3	0.63
	AI2	8	1.1	0.27
Neale, compre.	exps	13	1.9	1.29
	conts	6	0.8	
	AI1	8	1.2	0.43
	AI2	6	0.9	0.02
Schonell	exps	12	1.7	0.49
	conts	9	1.3	
	AI1	9	1.3	-0.02
	AI2	9	1.3	0.02

Statistical significances: on all 4 measures post-test, experimentals' gains were significantly better than other 3 groups'; those groups did not differ significantly.

Follow-up: Experimentals made no further relative gain between post-test and follow-up, but maintained the advantage gained during the intervention. However, inspection of the follow-up means reveals that the gains over post-test were slight – all groups, including the experimentals, were making less than standard progress.

14   Reading Intervention (formerly Cumbria Reading with Phonology Project)

(2)   General use in Cumbria since the original project

Main reference: Hatcher (2000)

Date: 1994–98

Age-range: Y2–10

<div align="center">Type of children: Low attainment</div>

N of experimental group: 427, including 73 statemented (see part 3)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 12

Tests: (Reading) Burt, 1974 revision; (Spelling) Schonell

Pre- and post-test average r.a's, s.a's and s.d's: not stated

Gain in months of r.a./s.a. (s.d's not stated) and ratio gains:

	Gain	RG
Reading	6.1	2.0
Spelling	7.9	2.6

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated and could not be calculated

14 Reading Intervention (formerly Cumbria Reading with Phonology Project)

(3) For statemented children (subset of those in (2) above)

Main reference: Hatcher (2000)

Date: 1994-98

Age-range: Y2–10

Type of children in experimental groups: SEN – all statemented (MLD, children with dyslexia)

Nature and Ns of experimental and control groups:

	N
(1) Moderate Learning Difficulties (MLD) (IQ in range 55-75)	28
(2) Controls for MLD (Cont1)	27
(3) Children with dyslexia (DYS)	29
(4) Controls for DYS (Cont2)	29

Equivalence of groups: Each experimental child was matched (from a pool of 351) with a teacher-referred child with an equivalent score on four pooled literacy assessments and of same gender: also of similar age where possible

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 12

Tests: (reading) Burt, 1974 revision; (spelling) Schonell

Average pre- and post-test raw scores and s.d's, gains, and ratio gains as stated by author:

Group		Word reading		Spelling	
		r.a. (yrs)	Stand. score	s.a. (yrs)	Stand. score
MLD (N = 28)	Pre	6.1	18.86 (10.68)	6.2	12.07 (7.93)
	post	6.5	25.96 (10.63)	6.8	18.46 (7.77)
	gain	4.8m	7.10	7.2m	6.39
	RG	1.4		2.4	
Cont1 (N = 27)	Pre		18.30 (9.77)		12.11 (6.70)
	Post		26.85 (9.80)		20.26 (8.31)
	Gain		8.55		8.15
	RG	1.7		3.0	
DYS (N = 29)	Pre	6.6	26.90 (12.61)	6.8	18.41 (8.20)
	Post	7.4	38.24 (13.45)	7.4	23.97 (7.99)
	Gain	9.6m	11.34	7.2m	5.56
	RG	2.9		2.1	
Cont2 (N = 29)	Pre		26.59 (10.81)		18.79 (9.25)
	Post		38.20 (11.60)		27.31 (8.94)
	gain		11.61		8.52
	RG	3.0		3.2	

Effect sizes: Because the experimental groups made less progress than the control groups, but still real progress (as shown by the RGs), effect sizes would be misleading and are not shown.

Statistical significances: For reading, DYS made a significantly greater gain than MLD, but neither experimental group differed significantly from its control group. For spelling, DYS and MLD did not differ, and MLD did not differ from its control group, but DYS made significantly less gain than its control group.

15   Reading Recovery

(1)   London and Surrey

Main references: Sylva and Hurry (1995a, b), Hurry and Sylva (1998)

Date: 1992-93

Age-range: Y2

Type of children: Low attainment

N of experimental group: 89 in 22 schools in seven LEAs in south-eastern England, six in Greater London (Bexley, Greenwich, Hammersmith and Fulham, Islington, Wandsworth, Westminster), plus Surrey

Ns of control groups: (1) 40 in same schools; (2) 153, = 109 in 18 other schools in same LEAs + 44 in alternative intervention schools

N of alternative intervention (AI) group: 91 in 23 schools in same LEAs

Nature of alternative intervention: Phonological Intervention

Ns of control groups for alternative intervention: (1) 44 in same schools; (2) 108 of the same 109 children in 18 other schools

Equivalence of groups: Control schools were matched to experimental and alternative intervention schools, and pre-tests showed equivalence of pupil samples in the three groups of schools; but within-school control groups were not matched to experimental pupils in those schools (therefore data given below only for between-school comparisons)

Length of intervention in weeks: (average) 20 (but average 37 weeks (8.5 months) between pre-test (Sept/Oct 1992) and post-test (May-July 1993)). One-year follow-up took place in May-July 1994, and three-year follow-up in Summer 1996.

Tests: (reading) BASWRT, Neale; (spelling) British Ability Scale Spelling test

Reading results

Pre- and post-test and 12-month follow-up r.a's on BASWRT (s.d's not given), gains over previous test (months of r.a.), ratio gains and effect sizes (N.B. parallel data for Neale not given):

	pre-test r.a.	post-test r.a.	gain	RG	Effect size	1-year follow-up r.a.	gain
	(years & months)		(months)			(months)	
experimentals	4:11	6:4	17	2.0	0.70	7:0	8
main controls	5:0	5:8	8	0.9		6:6	10
AI	5:1	5:11	10	1.2	0.11	6:11	12
controls for AI	5:1	5:9	8	0.9		6:8	11

Statistical significances: between pre- and post-test, experimentals made significantly greater progress than main controls on BASWRT and Neale accuracy (but not Neale comprehension); alternative intervention group and their controls did not differ in progress on these tests

Follow-up: At the one-year follow-up, the Reading Recovery children were slightly less further ahead of, but still significantly better than, the main controls on the BASWRT and Neale accuracy, and now also significantly better on Neale comprehension. And by this point the Phonological Intervention (AI) group were significantly better than their controls on the BASWRT and Neale accuracy, but still not significantly better on Neale comprehension.

At the three-year follow-up, neither the Reading Recovery nor the Phonological Intervention group was significantly better overall than their respective controls. But within both groups, children receiving free school meals had sustained their gains, and were still ahead by about six months of r.a. Also, Reading Recovery children who had been complete non-readers at the pre-test in 1992 had sustained their gains, and were still ahead by about six months of r.a. - but this was not true of such children within the Phonological Intervention group. The researchers commented: 'For children who were non-readers at six [Phonological Intervention] was not enough. It would seem that these children need books as well as phonics.'

Other evidence on the long-term effectiveness of Reading Recovery, from Australia and New Zealand (Lowe, 1995; Moore and Wade, 1998), shows more lasting benefit and less wash-out.

Spelling results

Spelling tested only at one-year follow-up; only average raw scores and effect sizes over relevant control groups given:

	ave. raw score	effect size
experimentals	18	0.32
controls	14	
AI	18	0.27
controls for AI	15	

3.15   Reading Recovery

(2)   Bristol

Main reference: Fudge (2001)

Date: 1999-2001

Age-range: Y1-2

Type of children: Low attainment

N of experimental group: 145 in 21 schools

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 20

Reading test: WRAPS (Word Recognition and Phonic Skills)

Pre- and post-test scores and s.d's: not stated

Gain (in months of r.a): 14.5

Ratio gain: 2.9

Effect size: were not given and could not be calculated.

Statistical significance: not stated

16   Reciprocal Teaching

Main reference: Unpublished data supplied by Christa Rippon

Date: 2001-02

Age-range: Y6

Type of children: Low attainment

N of experimental group: 16

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 16

Reading test: Neale

Pre- and post-test average r.a's and s.d's: not stated

Gains (in months of r.a.) and ratio gains:

	Gain	RG
accuracy	14.0	28.8
comprehension	3.1	6.4

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

17   Somerset Self-esteem and Reading Project

SOMERSET (1)

N.B. Lawrence treated the 'counselling only' group as the main experimental group in this study. However, for consistency with his other studies, that group has here been described as the first alternative intervention group (AI I), while the 'counselling plus remedial teaching' group has been treated as the experimental group.

Main reference: Lawrence (1973, pp.44-56)-statistical data are more accurate in this version than in the original report (Lawrence, 1971), but details of interventions are derived from both, and from Lawrence (1988)

Date: 1970 (see Lawrence, 1988, p.10)

Age-range: Y4 (average ages of groups at beginning, presumably early in school year, were between 8:7 and 8:1 1)

Type of children: Low attainment ('considered by their head teachers to be retarded in reading')

N of experimental group: 12 in 1 school - but 11 at post-test
(received self-esteem counselling (20 minutes' individual interview per week, with a professional psychologist) plus remedial teaching - 30 minutes, once a week, in groups of 5 or 6, primarily phonics)

N of control group: 12 in 1 school

Ns of alternative intervention (AI) groups: (AI1) 12 in 1 school
(AI2) 12 in 1 school - but 11 at post-test

Nature of alternative interventions: (AI1) self-esteem counselling alone (20 minutes' individual interview per week, with same psychologist as experimentals)
(AI2) remedial teaching alone (30 minutes, twice a week, in groups of 5 or 6, primarily phonics, with same teacher as experimentals)

Equivalence of groups: Groups matched on age, sex, mental age (non-verbal IQ) and reading age, but each group was in a separate school

Length of intervention in weeks: 26

Reading test: Schonell Word Recognition Test

Pre- and post-test average scores (r.a's), pre-test s.d's (post-test s.d's not given), gains (in months of r.a.), ratio gains, and effect sizes calculated using differences between gains over control group's pre-test s.d.:

	pre-test average	(s.d.)	post-test average (months)	gain	RG	effect size
exps	6:11	(1:2)	7:9	10	1.7	0.33
conts	6:6	(1:3)	6:11	5	0.8	
AI1	6:10	(1: 1)	7:11	13	2.2	0.53
AI2	6:8	(1:4)	7:5	9	1.5	0.27

Statistical significances (ns = non-significant; p probability; U = Mann-Whitney 'U' test value; ? = not stated):

	conts		AI1		AI2	
	U	p	U	p	U	p
Exps	?	?	50	ns	34	ns
conts			2	<0.001	?	?
AII					30	<0.01

No reason given for not stating control vs experimental and control vs AI2 (remedial phonics only) values; professional counselling plus remedial phonics was no better than counselling only (AI1) or remedial phonics only (AI2); but professional counselling only was better than remedial phonics only or no intervention, and equal to professional counselling plus remedial phonics.

SOMERSET (2)

Main reference: Lawrence (1973,pp.56-65)-some details of interventions derived from Lawrence (1972)

Date: not stated (1970?)

Age-range: Y4 (average ages of groups at beginning ranged from 8:10-9:6)

Type of children: Low attainment ('considered to be retarded in reading')

N of experimental group: 14, 7 in each of 2 schools (received counselling provided by non-professionals, plus remedial teaching)

N of control group: (no no-intervention control group)

N of alternative intervention group: 14, 7 in each of same 2 schools (received remedial teaching only)

Equivalence of groups: Matched on age, sex, mental age and reading age, within schools

Length of intervention in weeks: 18 (but 26 between pre- and post-test)

Reading test: Schonell Word Recognition Test

Pre- and post-test average scores (r.a's), pre-test s.d's (post-test s.d's not given), gains (in months of r.a.), and ratio gains:

	pre-test average	(s.d.) (yrs & months)	post-test average	gain (months)	RG
experimentals	7:0	(1:9)	8:0	12	2.0
AI	6:10	(1: 11)	7:8	10	1.7

Effect size: was not given and could not be calculated because there was no no-intervention control group

Statistical significance: ns

Counselling by non-professionals plus remedial teaching was no better than remedial teaching alone

SOMERSET (3)

Main reference: Lawrence (1973, pp.65-74) - some details of interventions derived from Lawrence (1972)

Date: not stated (1971?)

Age-range: Y3-4 (average ages of the 8 pupil groups at beginning ranged from 7:10 to 9:7)

Type of children: Low attainment ('considered to be retarded in reading')

N of experimental group: 24, 6 in each of 4 schools
(received counselling provided by non-professionals, plus remedial teaching)

N of control group: (no no-intervention control group)

N of alternative intervention group: 24, 6 in each of same 4 schools
(received remedial teaching only)

Equivalence of groups: Matched on age, sex, mental age and reading age, within schools

Length of intervention in weeks: 17

Reading test: Schonell Word Recognition Test

Pre- and post-test average scores(r.a's),pre-tests.d's (post-test s.d's and s.d's of gains not given), gains (in months of r.a.), and ratio gains:

	pre-test average	(s.d.)	post-test average	gain	RG
		(yrs & months)		(months)	
experimentals	6:11	(1:10)	7:11	12	3.0
AI	6:10	(2:0)	7:1	3	0.8

Effect size: was not given and could not be calculated because there was no no-intervention control group

Statistical significance: $p < 0.05$

Counselling by non-professionals plus remedial teaching was better than remedial teaching alone

[N.B. A further study reported in Lawrence and Blagg (1974) was considered too small to be included - see chapter 4.]

SOMERSET (4)

Main reference: Lawrence (1985); some details from Lawrence (1988)

Date: 1984 (see Lawrence, 1988, p.11)

Age-range: Y3 ? ('eight-year-olds')

Type of children: Low attainment (all with reading quotient (r.a./c.a. x 100) below 85)

N of experimental group: 94 in 8 schools (received DISTAR as AI2, plus self-esteem counselling for 45 minutes once a week, in pairs, from one of 35 non-professional counsellors)

N of control group: 78

Ns of alternative intervention (AI) groups: (AI1) 79; (AI2) 84

Total N of pupils: Ns above total 335; Lawrence (1985, p.194) says 374, Lawrence (1988, p. 11) says 372 - but these may represent total number at pre-test

N of schools: Lawrence (1985, p.194) gives total number of schools across all 4 groups as 29; Lawrence (1988, p.11) gives number of experimental schools as 8; separate numbers for other 3 groups nowhere stated

Nature of alternative interventions:
- (AI1) DISTAR as AI2, plus drama teaching designed to enhance self-esteem, for about 45 minutes once a week, in groups of 7-15, given by County Adviser for Drama
- (AI2) DISTAR only, in groups of 6-10, for one hour, 3 times per week, from teachers trained by a manager of the accredited UK providers of DISTAR training

Equivalence of groups: Not stated, but appears to have been random assignment of pupils to groups

Length of intervention in weeks: 20

Reading test: Burt Word Recognition Test (Vernon revision, 1973)

Pre- and post-test average scores and s.d's: not stated

**Gains (in *raw score),* s.d's of gains, and effect sizes calculated using differences between gains over s.d. of control group *gain:*

| | gain in raw score | | effect |
	average	(s.d.)	size
experimentals	14.3	(6.4)	0.92
controls	8.8	(6.0)	
AI1	11.8	(6.6)	0.50
AI2	10.7	(5.1)	0.32

Ratio gain: n/a

Statistical significances: experimentals and AI1 made significantly greater gains than other two groups. 'Therapeutic' conditions (experimentals = counselling plus DISTAR; AI1 = drama plus DISTAR) did not differ, and were better than DISTAR only (AI2) and no intervention (controls), which also did not differ.

18   SPELLIT

Main reference: Rack and Hatcher (2002a, in press, b, in press)

Date: 1999-2000

Age-range: Y2-4

Type of children: Low attainment ('reading and spelling in the lower 10% on standardised tests')

N of experimental group: 51

N of control group: 58

N of alternative intervention: 41

Nature of alternative intervention group: 'Home Support Programme consisting of activities and exercises to be done at home for around 15 minutes a day, for 5 days a week over a 30 week period'

Equivalence of groups: children were allocated at random, with some minor adjustment to achieve a balance of age, IQ, etc., in the three groups

Length of intervention in weeks: 30, but 39 on average between pre- and post-test

Reading test: BASWRT

Pre- and post test average r.a's and standardised scores (s.d's not stated), gains, and ratio gains:

Group	r.a. pre	r.a. post	Gain (in months of r.a.)	RG	Standard Score pre	Standard Score post	Gain
Exps.	5.77	6.69	11	1.2	82.41	84.73	2.32
Conts.	5.85	6.37	6	0.7	82.66	81.26	-1.4
AI	5.80	6.58	9	1.0	84.15	85.22	1.07

Effect sizes: were not stated and could not be calculated

Statistical significance: not stated

19   THRASS

Main reference: Matthews (1998**)**

Date: 1998

Age-range: Y3–6 (also Y2, Y7–9, not reported here)

Type of children: Low attainment

N of experimental group: 160 (for year groups, see below)

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 13

Tests: (Reading) Neale; (Spelling) Schonell

Pre- and post-test average r.a's and s.a's and s.d's: not stated

Gains (in months of r.a. / s.a.) and ratio gains:

	N	Reading accuracy		Reading comprehension		Spelling	
		Gain	RG	Gain	RG	Gain	RG
Y3	30	6.6	2.2	7.0	2.3	7.5	2.5
Y4	45	7.3	2.4	8.2	2.7	2.7	0.9
Y5	39	10.3	3.4	11.3	3.8	2.7	0.9
Y6	46	7.1	2.4	12.5	4.2	3.0	1.0

Effect sizes: were not given and could not be calculated.

Statistical significances: not stated

20   Time for Reading

Main reference: Elliott et al. (2000)

Date: not stated, but about 1995

Age-range: Reception (4-5)

Type of children: Low attainment, or at risk of it (pre- and early readers in a disadvantaged urban area)

Ns of experimental groups: (pre & post) 68; (3-year follow-up) 50

Ns of control groups: (pre & post) 72; (3-year follow-up) 49

Equivalence of groups: Random assignment of classes to conditions; pre-test data showed no significant differences

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 26 – but 3-year follow-up testing was conducted 2½ years after end of intervention

Tests: (pre & post) Specially constructed odd-one-out tests of initial phonemes and rhymes; (3 year follow-up) WORD (Wechsler Objective Reading Development) Scales

Pre- and post-test average raw scores and gains:

	Pre	**Post**	**Gain**
Initial phoneme test			
exp.	3.69	5.40	1.71
cont.	3.83	5.80	1.97
Rhyme test			
exp.	4.92	6.14	1.22
cont.	5.00	6.26	1.26

Performance on reading measures at 3-year follow-up

	N	Reading accuracy		Reading comprehension		Spelling	
		Mean	s.d.	Mean	s.d.	Mean	s.d.
Experimental	50	89.8	15.6	88.5	14.7	91.7	14.3
Control	49	90.6	16.4	89.6	13.8	93.5	11.7

Ratio gain: n/a

Effect sizes: could not be calculated but would effectively be zero

Statistical significances: all differences between experimentals and controls ns

21   AcceleRead AcceleWrite (formerly Jersey Computer Assisted Reading Development Programme)

(1)   Jersey

Main reference: Jersey Advisory Service (1993)

Date: 1993

Age-range: Y3-9 (Ns for separate years not given; average age at outset 10:3)

Type of children: Low attainment (r.a. well below c.a.)

N of experimental group: 61 in 15 primary & 4 secondary schools (62 for spelling)

N of control group: a control group of 9 pupils (10 for spelling) is mentioned but is statistically inadequate and therefore ignored here

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 4

Reading and spelling tests: British Ability Scales

Gains in months of r.a./s.a. (s.d's and follow-up r.a's and s.a's not given) and ratio gains:

	Gain	**Ratio gain**
reading	8.3	8.3
spelling	4.0	4.0

Average standardised scores at pre- and post-test and 10-week and 6-month follow-ups, gains from pre-test (s.d's not given), and effect sizes for post-test vs. pre-test only calculated using s.d. of standardisation sample:

Reading	**Average score**	**Gain**	**Effect size**
pre	92.4		
post	100.7	8.3	0.55
10-week follow-up	103.0	10.6	
6-month follow-up	105.7	13.3	

Spelling	Average score	Gain	Effect size
pre	96.0		
post	100.0	4.0	0.27
10-week follow-up	100.7	4.7	
6-month follow-up	103.8	7.8	

Statistical significances: not stated

Follow-up: The follow-up data show that the children continued to make relative gains.

21   AcceleRead AcceleWrite (formerly Jersey Computer Assisted Reading Development Programme)

(2)   Devon

Main reference: Unpublished data supplied by Martin Miles

Date: 2002

Age-range: 'Older KS2'

Type of children: Low attainment ('identified as experiencing difficulties with reading and/or spelling')

N of experimental group: 30

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 4

Tests: British Ability Scales Word Reading and Spelling

Pre- and post-test average r.a's and s.a's and s.d's: not stated

Gains in months of (r.a. / s.a.) and ratio gains:

	Gain	**RG**
reading	16.1	16.1
spelling	9.8	9.8

Effect sizes: were not given and could not be calculated.

Statistical significance: not stated

22   Integrated Learning Systems, National Council for Educational Technology study

(1)   Mainstream, A.   Phase II

Main reference: National Council for Educational Technology (1996)

Date: 1994-96

Age-range: Y3-6 ? ('Key Stage 2'; in School U, Y2 and Y6; Key Stage 3 also in project but not covered here)

Type of children: Mixed-ability (mostly; in School U, SEN)

N of experimental group: 760 in 7 primary schools in main study (NCET, 1996, p. 12); this certainly includes children involved in numeracy but not literacy - but not clear if it includes controls - only 375 experimentals traceable in details of report

N of control group: not stated

Equivalence of groups: not stated

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: not stated (26 in Schools A and U)

Reading test: Progress was measured by tests within the programs. These provided '... assessment in terms of "AVG" levels. These represent US grade equivalents and although preliminary work to reference them to UK measures seems to indicate that they are reasonably accurate, they have been treated with caution by our evaluators' (NCET, 1996, p.21).

Pre- and post-test average scores and s.d's: not stated

Gain: 'No consistent learning gains' (NCET, 1996, p.19); in School A, 30 experimentals (across full primary age range?) made average gain of 8.4 months of r.a. in 6 months - controls' average gain was 2.7 months; in School U, control group *outperformed* experimentals

Ratio gains: RGs overall and for School U were not given and could not be calculated; for School A, exps: 1.4; controls: 0.5

Effect sizes calculated using pre-test s.d's of experimental and control groups (for formula used see NCET, 1996, pp.6 & 10, footnotes), as stated in report:
    School A, 0.55 in favour of experimentals;
    School U, 0.4 in favour of controls (NCET, 1996, p.19);
    report also implies that overall effect size was close to zero and ns, because there were no consistent learning gains

Statistical significances: ns except for Schools A and U

22   Integrated Learning Systems, National Council for Educational Technology study

(1)   Mainstream, B.   Phase III

Main reference: BECTa (1998)

Date: 1996-97

Age-range: Y5

Type of children: Mixed-ability

N of experimental group: 193 in 11 schools

N of control group: 284 in 19 schools

Equivalence of groups: not matched; pre-test differences handled statistically

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 52

Reading test: not stated

Pre- and post-test average standardised scores and s.d's: not stated

Gain: not stated

Ratio gain: n/a

Effect size: -0.02 (BECTa, 1998, p.9)

Statistical significance: statistically significant in favour of control group even though difference was very small (BECTa, 1998, p.9)

22   Integrated Learning Systems, National Council for Educational Technology study

(2)   For pupils with low attainments in reading

No data reported here – see section 3.24

23   RITA (Reader's Intelligent Teaching Assistant)

The data for IA&T here are identical to those for the pre- and post-test in the entry for IA&T above.

Main reference: Nicolson et al. (1999)

Date: not stated (1997–98?)

Age-range: Y2-3

Type of children: Low attainment

Ns of experimental groups: (Y2) 58; (Y3) 16, in 4 schools in total

Ns of control groups: (Y2) 58; (Y3) 45, in different schools

Ns of alternative intervention groups: (Y2) 59; (Y3) 36, in same 4 schools and classes as experimentals but in previous year

Nature of alternative intervention: called 'Traditional' in Nicolson et al. (2000) but actually the experimentals in Interactive Assessment and Teaching (IA&T) – see separate entry

Equivalence of groups: matched on age and reading performance

Length of intervention in weeks: 17

Tests: Wechsler Objective Reading Dimension (WORD) reading and spelling

Ns, pre- and post-test average standardised scores and s.d's, and effect sizes calculated using pooled pre- test s.d's:

	Reading Standard Score					Spelling Standard Score				
Intervention type	*Pre-test* M	s.d.	*Post-test* M	s.d.	Effect size	*Pre-test* M	s.d.	*Post-test* M	s.d.	Effect size
Y2										
RITA	89.60	3.41	90.81	3.61	0.30	80.19	6.67	88.66	8.91	0.98
Control	89.74	4.01	89.41	5.30		84.49	8.65	86.24	9.79	
AI (IA&T)	89.03	3.51	92.76	7.46	0.94	84.26	8.97	91.67	10.57	0.95
Y3										
RITA	79.69	3.50	87.13	12.21	1.34	78.38	5.50	84.44	8.02	0.77
Control	79.49	5.55	80.53	7.01		81.19	7.87	83.22	8.22	
AI (IA&T)	79.94	3.41	83.31	3.61	0.61	82.64	6.67	88.28	8.91	0.72

Ratio gains: n/a

Statistical significances: Experimentals and AI groups in both year groups made significantly greater gains than controls; experimental and AI groups did not differ

24   Paired Writing (1)

Main reference: Sutherland and Topping (1999). Also summarised in Topping (2001), and Topping et al. (2000). Approach also described in Topping (1995).

Date: not stated (c.1997?)

Age-range: Scottish Primary 4 (=Y3) ('8-year-olds')

Type of children: Mixed-ability

Ns of experimental groups: 16 in each of two classes in 1 school; one group had helpers ('tutors') of same ability (and swopped roles at intervals), the other had helpers of different ability (and did not swop roles)

Ns of control groups: 16 in each of the same two classes

Equivalence of groups: chosen randomly (alternate children on class register allocated to different groups, than groups randomly assigned to intervention or control)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 8

Writing assessment: Scottish 5-14 National Curriculum Guidelines (SQA, 1997) which have 5 levels, A (low)-E (high), converted to numerical scale 1-5 for statistical purposes in this study

Average pre-and post-test raw scores and gains for writing, and s.d's of pre-and post-test score (s.d's of gains not stated), statistical significances, and effect sizes calculated (by GB) using this formula: difference in gains (experimental group's gain minus relevant control group's gain) divided by pre-test s.d. of relevant control group (source: Sutherland and Topping, 1999, Table 1, p.170, edited and incorporating details from text):

	Cross-ability		**Same-ability**	
Experimental	Control	Experimental	Control
		Pre-test		
1.75	1.31	1.63	1.75
(0.97)	(0.92)	(0.78)	(0.43)
ns		ns	ns	
		Post-test		
2.13	1.44	1.69	1.56
(0.99)	(0.79)	(0.92)	(0.61)
gain 0.38	0.13	0.06	-0.19
$p = 0.036$	ns	ns	ns
ns		$p = 0.049$	
effect size = 0.27		effect size = 0.58	
difference between gains of 2 experimental groups significant, $p = 0.038$			

Ratio gain: n/a

24   Paired Writing (2)

Main reference: Nixon and Topping (2001). Also summarised in Topping (2001), and Topping et al. (2000). Approach also described in Topping (1995).

Date: not stated (c.1998?)

Age-range: writers: Scottish Primary 1 (=Reception); helpers: Scottish Primary 7 (=Y6)

Type of children: writers: mixed-ability; helpers: 'weak writers….at their own level' (Topping et al., 2000, p.85)

N of experimental group: 10 writers (tutors), selected randomly from all 58 P1 children in one school; 10 helpers (tutors); all in one year at one school

N of control group: 30 other children in same year at same school

Equivalence of groups: not equivalent because not matched

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 6

Writing assessment: 17-point scale based on research literature, especially Gorman and Brooks (1996)

Average pre- and post-test raw scores and gains for writing (s.d's not stated) and statistical significances (source: Nixon and Topping, 2001, Table II, p.50, lightly edited):

	Pre-test	Post-test	Gain	Significance level (Wilcoxon)
Paired (*n* = 10)	6.67	10.17	3.50	*p* = 0.003 (one-tailed)
Unpaired (*n* = 30)	5.67	7.33	1.67	*p* < 0.001 (one-tailed)
Difference	1.00	2.84		
Significance level (Mann-Whitney)	*p* = 0.461 (two-tailed)	***p* = 0.034** (one-tailed)		

Ratio gain: n/a

Effect sizes: not stated and could not be calculated

24   Paired Writing (3)

Main references: Yarrow and Topping (2001). Also summarised in Topping (2001), and Topping et al. (2000). Approach also described in Topping (1995).

Date: not stated (c.1997?)

Age-range: Scottish Primary 6 (=Y5) ('10- and 11-year-olds')

Type of children: 'A problematic mixed-ability class'

N of experimental group: 13 (14 at pre-test but one pair lost), all in one class in one school

N of control group: 13 (14 at pre-test but one pair lost), all in same class

Equivalence of groups: Children matched in pairs on basis of gender and pre-test writing scores and allocated to groups; groups then allocated randomly to be experimentals or controls. Each group then divided at median score – lower half of experimentals became writers (tutees); lower half of controls became their control group; upper half of experimentals became helpers (tutors); upper half of controls became their control group. However, here all experimentals are treated as one group and all controls as another because Ns would otherwise be too small.

N of alternative intervention groups: (no alternative intervention group)

Nature of alternative intervention: n/a

Length of intervention in weeks: 6 (8 weeks between pre-and post-test)

Writing assessment: as Sutherland and Topping (1999) but using 35 sub-criteria to create 35-point scale. The writing was marked by people who were unaware of which group the children belonged to.

Pre-and post-test average raw scores and gains for writing, s.d's of post-test and gain scores (s.d's of pre-test scores not stated), and effect size:

	Interaction (exps)	No Interaction (conts)
pre-test	11.10	11.16
post-test	16.15	13.54
(s.d.)	(4.06)	(4.89)
gain	5.08	2.38
(s.d.)	(2.33)	(3.52)

Statistical significance: $p = 0.016$

Effect size: 0.53

Ratio gain: n/a

25   Family Literacy

Main references: Brooks et al. (1996a, 1997); Gorman and Brooks (1996)
Date: Summer 1994-Summer 1995 (1 cohort of children in each term)

Age-range: pre-school to Y2 (ages 3-6)

Type of children: Low attainment, or at risk of it

N of experimental group: 362 at outset, on about 20 sites. Smaller numbers at post-test and at each of three follow-ups (see below) because calculations based only on children with complete data ('returners')

N of control group: (no control group)

N of alternative intervention group: (no alternative intervention group)

Length of intervention in weeks: 12

Writing assessment: on 7-point scale derived empirically from analysis of about 900 of the scripts, (see Gorman and Brooks, 1996a), later extended to 12-point scale (see Brooks *et al*., 1997)

Sample sizes, average raw scores and s.d's at pre- and post-test and 12-week, 9-month and long-term follow-ups, and gains from pre-test:

	N	average score	(s.d.)	Gain
pre-test vs post-test	279	3.5 4.1	(1.6) (1.7)	0.6
pre-test vs 12 week follow-up	179	3.7 4.6	(1.6) (1.4)	0.9
pre-test vs 9 month follow-up	91	4.0 5.4	(1.5) (1.3)	1.4
pre-test vs long-term follow-up	175	3.4 8.0	(1.6) (1.7)	4.6

Ratio gain: n/a

Effect size: not stated and could not be calculated

Statistical significances: p<0.05 for all differences from pre-test, and evaluators judged the progress to be above what would have been expected

A.3　Comparisons between schemes

To provide a basis for comparing the interventions, including alterative intervention and control groups, the two forms of impact measure (RGs and effect sizes) have been put into rank orders, first for reading, then for spelling, then for writing - see Tables A.4-8 below. In Tables A.4-5, where measures for both reading accuracy and reading comprehension were available, both have been listed.

In Tables A.4 and A.6,

- (Table A.4 only) all the blanks under 'comprehension' mean that only accuracy data were available for those groups;

- where there was a satisfactory control group, the significance of the difference in gains has been indicated between the two RGs;

- where the control group was non-equivalent, the significance of the difference in gains has always been shown as uncertain – but the control group's RG is still valid in its own right, even though shown in brackets;

- where there was no control group, 'controls' and 'significance of the difference' columns are left blank.

In eight of the evaluations studied here (The Catch Up Project (pilot and national studies), Inference Training, MTSR (initial study), Parental Involvement, Reading Intervention (original), Reading Recovery, Somerset, RITA) different interventions were compared within one study. In the case of Reading Intervention (original) and Reading Recovery, two of the interventions were effectively the same, namely Reading Recovery itself or its Cumbria derivative on the one hand, and a phonological approach on the other (Phonological Training in Reading Intervention (original), Phonological Intervention in the RR study). Most of these studies provided useful comparative quantitative data, with statistical tests of the differences between approaches - these are included in the descriptions above, and form part of the basis for the judgments reported in chapter 2. However, it proved impossible to indicate the statistical significance of differences between experimental and alternative intervention groups clearly in Tables A.4-8, and this information is therefore provided in Table A.9. In the case of Inference Training, the differences include those between the two experimental groups.

Relative to the first edition, there were fewer studies with control, comparison or alternative intervention groups.

Table A.4: List of reading studies in decreasing order of ratio gain (RG) for accuracy

RG of 1.0	=	Exactly standard progress
RG of 1.4 or above	=	Impact of educational significance
RG of less than 1.4	=	Impact of doubtful educational significance

Study	Year group	Taught by	RG, accuracy		RG, comprehension		Follow-up
			Exps	Conts	Exps	Conts	
AcceleRead AcceleWrite, Devon	Y5-6	computer & supervising adult, 1-1	16.1				
AcceleRead AcceleWrite, Jersey	Y3-9	computer & supervising adult, 1-1	8.3				Continued to gain up to 10 months later
Phono-Graphix™ in Bristol	Y2-6	teacher & other adult, 1-1	8.3		8.3		
Phono-Graphix™ in Bristol	Y4-6	teacher & other adult, 1-1	5.8		4.3		
Inference Training, AI1 (comprehension exercises)	Y3	teacher, group	5.0		9.6		
Family Literacy for New Groups	Y4	other adults, group	4.7				
BRP in Derbyshire	Y1	other adults, 1-1	4.6				
Phono-Graphix™ in Surrey	Y4	teacher, 1-1	4.5				
MTSR, pilot, exps	Y2	teacher, group	4.5				
Inference Training, exps 1 (less skilled comprehenders)	Y3	other adults, group	4.3		17.4		
Inference Training, exps 2 (skilled comprehenders)	Y3	other adults, group	3.9		5.9		

Study	Year group	Taught by	RG, accuracy			RG, comprehension			Follow-up
			Exps		Conts	Exps		Conts	
The Catch Up Project in Norfolk, county, 6 schs	Y2-3	other adults, 1-1	3.7						
MTSR in Bolton	Y2	teacher, group	3.7						
BRP in Derbyshire	Y2	other adults, 1-1	3.7						
BRP in Derbyshire	Y4	other adults, 1-1	3.5						
Reciprocal Teaching	Y6	teacher, group	3.5			6.4			
The Catch Up Project, pilot (exps in matched schools)	Y3	teacher, 1-1	3.4	?	0.4				
THRASS	Y5	teacher, group	3.4			3.8			
Paired Reading, exps in control-group designs	Y1-11	other adults/pupils, 1-1	3.4	*	2.0	4.6	*	2.5	Continued to gain for 17 weeks and more
Paired Reading, all exps	Y1-11	other adults/pupils, 1-1	3.3			4.3			
The Catch Up Project in Norfolk, county, 12 schs	Y2-3	other adults, 1-1	3.3						
BRP in Derbyshire	Y3	other adults, 1-1	3.3						
BRP in Derbyshire	Y5	other adults, 1-1	3.2						
BRP in Redcar and Cleveland	Y1-6	other adults, 1-1	3.2						
BRP in Derbyshire	Y6	other adults, 1-1	3.1						
Family Literacy for New Groups, linguistic minorities	Nursery-Y2	other adults, group	3.0						
Inference Training, AI2 (rapid decoding)	Y3	other adults, group	3.0			8.2			
Somerset (3), exps (counselling plus remedial)	Y3-4	other adults, 1-1	3.0						
Reading intervention for children with dyslexia	Y2-10	teachers	2.9	ns	3.0				
Reading Recovery in Bristol	Y1-2	teacher, 1-1	2.9						
The Catch Up Project, pilot (all	Y3	teache, 1-1	2.6	?	(0.4)				

Study	Year group	Taught by	RG, accuracy			RG, comprehension			Follow-up
			Exps		Conts	Exps		Conts	
experimentals)									
The Catch Up Project in Norfolk, King's Lynn	Y3-6	other adults, 1-1	2.5						
BRP in Bradford, middle schs	Y4-7	other adults, 1-1	2.5	?	(0.8)				
BRP in Bradford, first schs	Y1-3	other adults, 1-1	2.4	?	(0.8)				
THRASS	Y6	teacher, group	2.4			4.2			
THRASS	Y4	teacher, group	2.4			2.7			
THRASS	Y3	teacher, group	2.2			2.3			
Phono-Graphix™ in Bristol	Y1	teacher, group	2.2						
MTSR, pilot, exps	Y5	teacher, group	2.2						
Somerset (1), AI1 (counselling only)	Y4	other adults, 1-1	2.2	*	0.8				
Somerset (2) exps (counselling plus remedial)	Y4	other adults, 1-1	2.0						
Reading intervention, general use	Y2-10	teacher, group	2.0						
The Catch Up Project in Norfolk, Thetford	Y2-4	other adults, 1-1	2.0						
The Catch Up Project in Cornwall	Y3	teacher, 1-1	1.9						
Reading Recovery (L&S), exps	Y2	teacher, 1-1	1.9	*	0.9				Some gains maintained, some lost over 3 years
Reading intervention, orig., exps, Neale	Y2	teacher, group	1.8	*	0.9	1.9	*	0.8	1 year on, exps still ahead relatively, but all groups making less than standard progress
The Catch Up Project in Wakefield	Y3-4	teacher, 1-1	1.8						
BRP in Durham	Y1	other adults	1.8			0.4			
Individual Spelling, group 2	Y2-3	teacher, group	1.8						
Individual Spelling, group 1	Y2-3	teacher, group	1.7	*	0.8				Continued to gain up to 5 months on

Study	Year group	Taught by	RG, accuracy			RG, comprehension			Follow-up
			Exps		Conts	Exps		Conts	
MTSR, pilot, AI2	Y5	teacher, group	1.7						
Knowsley	Y6	other adults	1.7						
Somerset (1), exps (counselling plus phonies)	Y4	other adults	1.7	?	0.8				
Somerset (2), AI (remedial only)	Y4	other adults	1.7						
Reading intervention, orig., exps, BASWRT	Y2		1.5	*	1.1				
Somerset (1), AI2 (phonics only)	Y4	other adults	1.5	?	0.8				
BRP in Worcs, phase 1	Y1-6	other adults	1.5	*	0.9				Gain maintained over a further term
Reading intervention for children with MLD	Y2-10	teachers	1.4	ns	1.7				
The Catch Up Project, pilot (matched time)	Y3	teacher, group	1.4	?	(0.4)				
Integrated Learning Systems, School A	Y3-6	computer & supervising adult, 1-1	1.4	*	0.5				
The Catch Up Project in Norfolk, Gt Yarmouth	Y6	other adults	1.4						
The Catch Up Project, national, exps.	Y3	teacher, 1-1	1.4	?	(1.0)				

The Catch Up Project, national, matched time	Y3	teacher, 1-1	1.3						
Reading intervention, orig., AI1 (reading only) Neale	Y2	teacher, 1-1	1.3	ns	0.9	1.2	ns	0.8	
RR (L&S), AI (Phonological Intervention)	Y2	teacher, 1-1	1.2	ns	0.9				
BRP in Worcs, phase 2	Y1-6	other adults, 1-1	1.2	*	0.8				
SPELLIT, exps.	Y2-4	teacher, ?	1.2	?	(0.7)				

BRP in Durham	Y2	other adults	1.1			0.8		
Reading intervention, orig., AI1, BASWRT	Y2	teacher, 1-1	1.1	ns	1.1			
Reading intervention, orig., AI2, BASWRT	Y2	teacher, 1-1	1.1	ns	1.1			
Reading intervention, orig., AI2, Neale	Y2	teacher, 1-1	1.1	ns	0.9	0.9	ns	0.8
PAT	Y4-7	teacher, 1-1	1.1	*	0.9			
SPELLIT, AI	Y2-4	other adults, 1-1	1.0					
BRP in Durham	Y5	other adults, 1-1	1.0					
BRP in Durham	Y3	other adults, 1-1	0.9					
Somerset (3), AI (remedial only)	Y3-4	other adults, 1-1	0.8					
BRP in Durham	Y4	other adults, 1-1	0.6					

Key:

Conts	=	Controls
Exps	=	Experimentals
( )	=	RG based on non-equivalent control group but valid in its own right
*****	=	Difference in gains is statistically significant
ns	=	Difference in gains is statistically non-significant
?	=	Significance of difference in gains was not stated or was unreliable

Table A.5: List of reading studies in decreasing order of effect size for accuracy

Key:

Effect Size of 0.25 or above	=	Impact of educational significance
Effect size of less than 0.25	=	Impact of doubtful educational significance

Study	Year Group	Taught By	Effect Size		Follow-up
			Acc	Comp	
Family Literacy in Hampshire	R	other adults, group	1.91		
RITA, exps	Y3	teachers, group	1.34		
Reading intervention, orig., exps (reading & phonology), Neale	Y2	teachers, group	1.60	1.29	1 year on, exps still ahead relatively, but all groups making less than standard progress
IA&T (also within RITA)	Y2	teachers, group	0.94		Gain was lost
Somerset (4), exps (counselling plus DISTAR)	Y4	other adults, group	0.92		
Paired Reading	Y1-11	other pupils, 1-1	0.87	0.77	Gain was maintained up to 17 weeks on
Parental Involvement, experimentals	Y2-3	parents, 1-1	0.84		Gain was maintained up to 3 years on
Reading Recovery (L&S), experimentals, Neale	Y2	teachers, group	0.82	ns	Some gains maintained, some lost over 3 years
The Catch Up Project, pilot, experimentals in matched schools	Y3	teachers, 1-1	0.78		
Family Literacy for New Groups, linguistic minorities	nursery-Y2	other adults, group	0.72		
Reading Recovery (L&S), exps., BASWRT	Y2	teachers, group	0.70		
IA&T (also within RITA)	Y3	teachers, group	0.61		Gain was maintained 6 months on
Family Literacy for New Groups	Y4	other adults, group	0.58		
AcceleRead AcceleWrite, Jersey	Y3-9	computer & teacher, 1-1	0.55		
Integrated Learning Systems, School A	Y3-6	computer & teacher, 1-1	0.55		
Somerset (1), AI 1 (counselling only)	Y4	other adults, group	0.53		
Somerset (4), AI I (drama plus DISTAR)	Y4	other adults, group	0.50		
Reading intervention, orig., exps (reading & phonology), BASWRT	Y2	teachers, group	0.45		
Knowsley	Y2	other adults, 1-1	0.35		
BRP in Worcs, phase 2	Y1-6	other adults, 1-1	0.34		
Somerset (1), experimentals (counselling plus phonics)	Y4	other adults, 1-1	0.33		
Somerset (4), AI2 (DISTAR only)	Y4	teachers, group	0.32		
RITA, exps.	Y2	teachers, group	0.30		
Family Literacy Demonstration Programmes	nursery-Y2	other adults, group	0.29		Some further gains up to 12 weeks, then maintained up to 3 years
Somerset (1), AI2 (phonics only)	Y4	teachers, group	0.27		
Reading intervention, orig., AI2 (phonology only), Neale	Y2	teachers, group	0.27	0.02	
The Catch Up Project, pilot, AI (matched time)	Y2	teachers, 1-1	0.25		

PAT	Y4-7	teachers, 1-1	0.16		
BRP in Worcs, phase 1	Y1-6	other adults, 1-1	0.13		Gain was maintained up to one term on
RR (L&S), AI (Phonological Intervention), BASWRT	Y2	teachers, group	0.11		
Parental Involvement, AI (extra teaching)	Y2-3	teachers, group	0.09		
Reading intervention, orig., AI1 (reading only), BASWRT	Y2	teachers, group	0.04		
Reading intervention, orig., AI2 (phonology only), BASWRT	Y2	teachers, group	0.02		
Time for Reading	R	volunteers, 1-1	ns		
Integrated Learning Systems, overall	Y3-6	computer & teacher, 1-1	ns		
Integrated Learning Systems, phase III	Y5	computer & teacher, 1-1	-0.02		
Integrated Learning Systems, Sch. U	Y2&6	computer & teacher, 1-1	-0.4		

Key to Symbols:

ns	=	Figure was not given but was stated or implied to be close to zero and statistically non-significant

Table A.6: List of spelling studies in decreasing order of ratio gain

Key:

RG of 1.0	=	Exactly standard progress
RG of 1.4 or above	=	Impact of educational significance
RG of less than 1.4	=	Impact of doubtful educational significance

Study	Year Group	Taught By	Exps		Conts	Follow-up
MTSR in Bolton	Y2	teachers, group	13.4			
AcceleRead AcceleWrite, Devon	Y5-6	computer & teacher, 1-1	9.8			
Cued Spelling (1), exps.	Y4	parents, 1-1	4.1	?	(1.6)	
AcceleRead AcceleWrite, Jersey	Y3-9	computer & teacher, 1-1	4.0			Some further gain up to 6 months on
MTSR, pilot, AI2	Y5	teachers, group	3.4			
Phono-Graphix™ in Bristol	Y2-6	teachers, group	3.3			
Reading Intervention, general use	Y2-10	teachers, group	2.6			
THRASS	Y3	teachers, group	2.5			
Reading Intervention for children with MLD	Y2-10	teachers, group	2.4	ns	3.0	
Reading Intervention for children with dyslexia	Y2-10	teachers, group	2.1	*	3.2	
MTSR, pilot, exps	Y2	teachers, group	2.1			
Cued Spelling (2)	P5 (=Y4)	parents, 1-1	2.1			
Reading Intervention, orig., exps.	Y2	teachers, group	1.7	*	1.3	1 year on, exps still ahead relatively, but all groups making less than standard progress

Reading Intervention, orig, AI1	Y2	teachers, group	1.3			
Reading Intervention, orig, AI2	Y2	teachers, group	1.3			
THRASS	Y6	teachers, group	1.0			
THRASS	Y4	teachers, group	0.9			
THRASS	Y5	teachers, group	0.9			
The Catch Up Project in Norfolk, county, 12 schs	Y2-3	other adults, 1-1	0.5			
MTSR, pilot, exps	Y5	teachers, group	-3.6			

Key:

Exps	=	Experimentals
*	=	Difference in gains is statistically significant
?	=	Significance of difference in gains was not stated or was unreliable

Table A.7: List of spelling studies in decreasing order of effect size

Key:

Effect Size of 0.25 or above	=	Impact of educational significance
Effect size of less than 0.25	=	Impact of doubtful educational significance

Study	Year Group	Taught By	Effect Size	Follow-up
RITA, exps.	Y2	computer & teacher, 1-1	0.98	
IA&T, (also within RITA)	Y2	teachers, group	0.95	Gain was maintained up to 6 months on
RITA, exps.	Y3	computer & teacher, 1-1	0.77	
IA&T, (also within RITA)	Y3	teachers, group	0.72	Gain was partly lost 6 months on
Reading Intervention, orig, exps.	Y2	teachers, group	0.49	
AcceleRead AcceleWrite, Jersey	Y3-9	computer & teacher, 1-1	0.27	

Reading Intervention, orig, AI2	Y2	teachers, group	0.02	
Reading Intervention, orig, AI1	Y2	teachers, group	-0.02	

Key to Symbols:

Exps	=	Experimentals

Table A.8: List of Paired Writing studies in decreasing order of effect size

Key:

Effect Size of 0.25 or above	=	Impact of educational significance
Effect size of less than 0.25	=	Impact of doubtful educational significance

Study	Year Group	Taught By	Effect Size	Follow-up
Paired Writing (1), cross-ability	P4 (=Y3)	other pupils, 1-1	0.58	
Paired Writing (3)	P1 (=R)	other pupils, 1-1	0.53	
Paired Writing (1), same ability	P6 (=Y5)	other pupils, 1-1	0.27	

Table A.9: **Statistical comparisons between experimental and alternative intervention (AI) groups**

The Catch Up Project,
pilot and national
studies: not stated

Inference Training: - On accuracy, all differences in gains among the two experimental and two AI groups were non-significant

- On comprehension, Inference Training was more effective for less skilled comprehenders than for skilled comprehenders; Inference Training was more effective than rapid decoding (AI2) for less skilled comprehenders; BUT comprehension exercises (AI1) were just as effective as Inference Training

Multi-Sensory Teaching
System for Reading
(initial study):   not stated

Parental Involvement: The experimental and AI groups could not be compared at post-test because they differed significantly at pre-test

Reader's Intelligent
Teaching Assistant
(RITA):   The 'alternative intervention' here was Interactive Assessment and Teaching, the experimental intervention from an earlier experiment by the same authors. RITA and IA&T did not differ on any measure

Reading Intervention
(original):   The experimental intervention (Reading with Phonology) was significantly better than both AIs (reading-only, phonology-only) on all three measures

Reading Recovery: No information was given on statistical significance of differences between experimental (Reading Recovery) and AI (Phonological Intervention) groups

Somerset (1):   Professional counselling plus remedial phonics was no better than counselling only (AI1) or remedial phonics only (AI2); but professional counselling only was better than remedial phonics only or no intervention, and equal to professional counselling plus remedial phonics.

Somerset (2):   Counselling by non-professionals plus remedial teaching was no better than remedial teaching alone

Somerset (3):   Counselling by non-professionals plus remedial teaching *was* better than remedial teaching alone

Somerset (4): 'Therapeutic' conditions (experimentals = counselling plus DISTAR; AI1 = drama plus DISTAR) made significantly greater gains than other two groups (AI2 = DISTAR only; controls = no intervention). The two therapeutic conditions did not differ significantly, and the other two groups also did not differ significantly

How should this mass of comparative detail on impact measures be interpreted?

The first thing to be said is that, given the uneven quality of the description, analysis and reporting of the studies, interpretation needs to be cautious and tentative. It is not the case that some schemes have been proven effective, and others ineffective, without qualification. High RGs and effect sizes do show that the relevant approaches have worked for some children in some circumstances, and *may* work for others, if implemented with similar care in similar circumstances. Low RGs and effect sizes show only that the relevant approaches have not worked for some children in some circumstances, and have no implications for the future, but they *might* work for other children in different circumstances.

That said, from inspection of the data and from the wider literature, it has been deduced that

- RGs of exactly 1.0 represent standard progress, or 'holding one's own'. Anything above this represents better than standard progress (but see the next point), while anything less means that the children are slipping (further) behind;

- RGs below 1.4, and effect sizes below 0.25, represent an impact that does not seem educationally significant. Pupils in these schemes did not just stay where they were, and did make some progress, in absolute terms; but it was slow, and they made little or no relative progress compared to control groups receiving no special intervention. Thus schemes (or conditions within schemes) with impact measures of this order did not seem to produce any impact over and above ordinary teaching, unless it is argued that 'holding their own' was a good result for such children - in other words, that without the intervention they would have fallen even further behind. Schemes in this group may be considered to have been 'less effective';

- all RGs above 1.4, and almost all effect sizes above 0.25, represent impact that is at least satisfactory, and in some cases excellent. Schemes in this group may be considered to have been 'more effective'.

Given this broad distinction, there are a few discrepancies between the RG and effect size lists for reading (not for spelling). Within Reading Intervention (original), the reading-only condition (AI1) had low RGs for accuracy and comprehension, and the phonology-only condition (AI2) had a low RG for accuracy, while the effect sizes seemed satisfactory or even high. Since the statistical analyses in the original report showed that neither AI produced greater gains than the control condition, that finding and the RGs are taken here to be the more accurate. Similarly, in BRP in Worcestershire the phase 1 effect size was low but the RG was satisfactory, and again the RG is taken to be more accurate.

The RG list for reading contains few values below 1.0 ('normal progress'), and all but a few of those RGs arose from control groups. This finding is, however, circular: children receiving ordinary teaching mostly made the progress to be expected of children receiving ordinary teaching. What is more interesting is that one control group (in Paired Reading) had RGs *above* 1.4, in fact above 2.0, and these children were therefore making better than expected progress despite, apparently, receiving no extra intervention. Perhaps Paired Reading affected a high proportion of the schools in the area in which it took place, and therefore the experimental schools were observed by others, and influenced non-participating schools to raise their game too. If this is true, it would be an argument for implementing initiatives at a fairly high density (though it would play even more havoc with evaluation statistics).

Neither spelling nor writing data were analysed in the first edition, and even now that such data have been trawled for, there is considerable less information than on reading, especially at the level of writing whole texts, where only the three Paired Writing studies and Family Literacy hint at what might be done. These studies cannot sustain any generalisations.

All the generalisations that the data seem to warrant are stated in chapter 2.

A.4   Follow-up studies

In many cases the impact observed during educational interventions is found to diminish or even vanish afterwards. Was this true of the schemes analysed here? Of the 25 main studies analysed, 15 provided no follow-up data, but 10 did provide information on re-tests of participating children at some point after the end of the intervention. These were: BRP in Worcestershire, Family Literacy Demonstration Programmes, Individual Spelling, IA&T, Paired Reading, Parental Involvement, Reading Intervention (original), Reading Recovery in London and Surrey, Time for Reading, and AcceleRead AcceleWrite in Jersey. For details, see the entries in this Appendix.

Conclusion on follow-up studies
The picture is uneven, but broadly positive. In four cases (Family Literacy Demonstration Programmes, Individual Spelling, Paired Reading, AcceleRead AcceleWrite in Jersey), children continued to make relative gains at least in the period immediately after the intervention (in Individual Spelling and Paired Reading there was only one follow-up; in Family Literacy Demonstration Programmes the children 'plateaued' after the first follow-upbut did not slip back; in AcceleRead AcceleWrite in Jersey even the second follow-up showed further relative gains).

In two studies (BRP in Worcestershire, Parental Involvement) the children maintained their gains.

In Time for Reading, there were no gains during the intervention, and no gains at follow-up three years on either.

Only in Reading Recovery in London and Surrey was there definite evidence of 'wash-out', that is, of children losing the gains they had made during the intervention - though it must be recognised that such findings are less likely to be reported. And even in the three-year follow-up to Reading Recovery, wash-out was not universal - children receiving free meals and those who had been non-readers at age six continued to benefit. (See also the much more positive findings from outside the UK noted in this Appendix under Reading Recovery and in section 3.15.) There was partial wash-out in IA&T – Y2's gain in reading was completely lost, but other evidence was more positive. In Reading Intervention (original), the experimental group maintained their relative position – but both they and the other groups had made less than standard progress, so that the gains were in fact only maintained to an extent.

Though it is therefore still possible to say that most gains were maintained, it seems impossible to generalise about factors which made the difference between wash-out and sustaining gains.