WHAT WORKS FOR SLOW READERS?

The Effectiveness of Early Intervention Schemes

Greg Brooks
Nicola Flanagan
Zenta Henkhuzens
Dougal Hutchison

nfer

WHAT WORKS FOR SLOW READERS?

The Effectiveness of Early Intervention Schemes

Greg Brooks
Nicola Flanagan
Zenta Henkhuzens
Dougal Hutchison

Third impression

nfer

Third impression, lightly revised
Published in March 1999
by the National Foundation for Educational Research
The Mere, Upton Park, Slough, Berkshire SL1 2DQ

Registered Charity No. 313392
ISBN 0 7005 1480 5

CONTENTS

LIST OF TABLES

LIST OF FIGURES

THE PROJECT TEAM

Dr Greg Brooks, Senior Research Officer and Project Director

Zenta Henkhuzens, Research Officer

Nicola Flanagan, third year undergraduate student of University of Surrey, on work experience at NFER in the academic year 1997-98

Dougal Hutchison, Principal Research Officer and Chief Statistician

Effie de Souza, Project Secretary

The first three authors contributed to the literature search. Zenta Henkhuzens sent out the requests for information and managed most of the resulting data-handling, including compiling the database. Nicola Flanagan pursued missing and unclear information, kept the database up to date, provided the first draft of chapter 3, and compiled Table 4. Dougal Hutchison vetted the statistical data. The remainder of the book was the work of Greg Brooks.

ACKNOWLEDGEMENTS

We should like to record our appreciation to the following colleagues at NFER:

- Revinder Pang, for devising the framework for the database
- Dr Wendy Keys and Dr Marian Sainsbury, for commenting on the report in draft
- Anne Nicholls for making the text more accessible
- Mary Hargreaves for laying out the text and Dr Enver Carim for overseeing publication of the report

and especially to

all the individuals and organisations who provided us with information: without their cooperation none of this project would have been possible. They are all acknowledged by name on the next two pages.

ACKNOWLEDGEMENTS OF INFORMATION PROVIDED

The information for this analysis in this book was gathered between 1995 and early 1998, partly through a literature search (including use of the *Register of Educational Research in the United Kingdom*, based at NFER), but mainly from an appeal for information which was sent out to LEAs, Institutions of Initial Teacher Education, and a number of individuals known to be active in the field. Responses to this appeal numbered 88, and these are all listed below, and gratefully acknowledged by the authors. The National Literacy Trust's database of initiatives was also searched, and we gratefully acknowledge this assistance too.

Local Education Authorities

Birmingham
Barnsley
Bradford
Buckinghamshire
Cambridgeshire
Calderdale
Cheshire
Croydon
Derbyshire
Devon
Doncaster
Dyfed
Ealing
Enfield
Essex
South Glamorgan
West Glamorgan
Greenwich
Gwent
Harrow
Hartlepool
Islington
Jersey
Kent
Kirklees
Knowsley
Lambeth
Leeds
Leicestershire
North Lincolnshire
Lothian
Manchester
Newport
Norfolk
Northamptonshire
Northumberland
Oldham
Richmond upon Thames
East Riding of Yorkshire
Rotherham
Sefton
Sheffield
Shropshire
St. Helens
Suffolk
Surrey
Tower Hamlets
Warwickshire
North Yorkshire

Institutions of Initial Teacher Education

Bretton Hall
Brighton University
Cheltenham and Gloucester College of Higher Education§
De Montfort University
Exeter University
Greenwich University
University of London Institute of Education
Leeds University
Leicester University
Manchester Metropolitan University
Newcastle University
Oxford Brookes University
Sunderland University

Individuals

Kay Alderdice	American Community Schools
Mrs Alfrey	Canterbury Christ Church College
Jean Alston	Toe by Toe
Prof. Ronald Arnold	University of Westminster
Ray Barker	Urban Learning Foundation
Liz Brooks	The Dyslexia Institute
Jennifer Chew	Strode's Sixth Form College, Egham
Una Christophers	First Steps
Suzi Clipson-Boyles	} Catch Up
Diana Bentley	
Dee Reid	
Viv Edwards	University of Reading
Prof. Elizabeth Goodacre	University of Middlesex
Angela Hobsbaum	Reading Recovery National Network
Nonie Hunter	Knowsley Teachers Centre
Jeff Hynds	Freelance consultant, Kent
Jenny Ingham	Bradford Book Flood
Christopher Jolly	Jolly Learning
Colette Kelleher	Bright Sparks
Sue McCaldon	Kickstart
Pamela Merriott	
Martin Miles	Talking Systems
Harry Muker	Optika, Maidenhead
Robert Tweed	Quality Development Service
Jo Wilson	THRASS
David Wray	EXEL

CHAPTER ONE

FOCUS AND INTENTION OF THIS BOOK

1.1 The focus

As many children as possible should learn to read satisfactorily first time, but what of those who don't? How are they to be helped? This book is about early intervention schemes that have been devised to help struggling readers, and is intended to inform schools' choices among such schemes.

More exactly, the questions this book is addressing are:

— **What intervention schemes are there which have been used in Britain in an attempt to boost the reading attainment of lower-achieving but non-dyslexic pupils in at least one of Years 1–4 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 Britain is partly intended to avoid a deluge of information on schemes used anywhere else 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 to *teachers*, in order to assist them in making their own choices of approach. Those choices have of course to 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.

The onus on schools to make such choices has been increased by the announcement of the National Literacy Strategy which is to be introduced in September 1998. That Strategy is to be based on the National Literacy Project approach, especially its Literacy Hour,

> *unless a school can demonstrate through its literacy plan, schemes of work and performance in Key Stage 2 tests, that the approach it has adopted is at least as effective.* (Literacy Task Force, 1997, p.5)

There is therefore all the more need for schools to have clear information, in order to make principled decisions about which strategy to adopt for children who do not 'crack' the task of literacy learning first time.

Table 1: Full and abbreviated names of the 20 studies

1	Basic Skills Agency's Family Literacy Demonstration Programmes	(1) BSA Family Literacy
2	Bradford Better Reading Partnership	Bradford BRP
3	* Buckinghamshire Phonological Awareness Training	Bucks PAT
4	* Catch Up *Matched Time*	Catch Up
5	* Cumbria Reading with Phonology Project *Reading-only* *Phonology-only*	Cumbria
6	Docklands Learning Acceleration Project	Docklands
7	Dyfed Improving Reading Standards in Primary Schools Project	Dyfed
8	Inference Training *Comprehension exercises* *Rapid decoding*	Inference Training
9	* Integrated Learning Systems, National Council for Educational Technology study	Integrated Learning Systems
10	Jersey Computer Assisted Reading Development Programme	Jersey Computer
11	Leeds Sustained Reading Intervention	Leeds SRI
12	Lewisham Literacy 2000	Lewisham 2000
13	* Paired Reading in Kirklees	(2) Paired Reading
14	* Parental Involvement in Haringey *Extra reading*	(3) Parental Involvement
15	Pause, Prompt and Praise *Sessions with untrained tutors*	(4) 'Pause, Prompt and Praise'
16	* Reading Recovery in London and Surrey *Phonological Intervention*	(5) Reading Recovery
17	Saint Lawrence School, Towcester, Northants	Saint Lawrence School
18	Shropshire: Raising Attainment In Shropshire Education	Shropshire RAISE
19	* Somerset Self-esteem and Reading Project *Self-esteem counselling only* *Remedial phonics only* *Remedial reading only* *Drama plus DISTAR* *DISTAR only*	(6) Somerset
20	The Handwriting, Reading and Spelling Sequence (THRASS Writetrack 1994)	(7) THRASS94

Key: * = *study with well defined no-treatment control group - see Appendix and section 2.9*
( ) – *'Notes to Table I' are on facing page*

Notes to Table 1:

(1) Because there are many other family literacy projects, references to the Basic Skills Agency's programmes are prefixed with 'BSA'.

(2) Topping and Lindsay (1992) reviewed dozens of Paired Reading schemes from all over the English-speaking world. For this book, the Kirklees scheme, which was not only British 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'.

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

(4) 'Pause, Prompt and Praise' is generally referred to in quotation marks so that in lists its internal comma does not cause confusion.

(5) Reading Recovery has been the subject of at least a dozen evaluations in Britain, 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 Britain. 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.

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

(7) The THRASS programme has gone through a number of versions, and titles. The evaluation which is quoted in this book was carried out in 1995 on the 1994 version, which had the full title 'The Handwriting, Reading and Spelling Sequence', or for copyright purposes the name 'THRASS Writetrack 1994'. To make it clear that this, and not any later version of THRASS, is the version referred to here, the abbreviation THRASS94 is used.

The project at Saint Lawrence School should be singled out for special comment. 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.

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

1.2 The need

What proportion of children do not get it right first time? An estimate (for England) can be based on the results of National Curriculum assessments. Table 2 gives the percentages of children not yet achieving level 2 in reading at the end of Key Stage 1 (age 7) in 1994-97.

Table 2: Percentage of children in England achieving below level 2 in reading in Key Stage 1 National Curriculum tests, 1994-97

Year	Percentage
1994	20%
1995	22%
1996	22%
1997	20%

Sources: GB. DfE (1994); GB. DfEE (1996a, b, 1997)

Thus roughly a fifth of children were not yet achieving level 2 in reading by age 7. This does NOT mean that they could not read at all; most children on level 1 were in fact close to achieving level 2. Nevertheless, any child who has not yet reached level 2 in reading by the end of Key Stage 1 is likely to have difficulty in coping with the steadily increasing demands of the curriculum in Key Stage 2 (and beyond).

So what can be done for children who do not get their learning to read right first time? We have identified 20 studies which are useful in attempting to answer this question.

1.3 The schemes covered

The titles of the 20 studies we have analysed, and the shorter names by which they are mainly referred to in this book, are as shown in Table 1. Several of the studies contained evaluations of more than one scheme, so in order to show the full coverage of the book, relevant studies are shown with their 'alternative treatments' (that is, the other approaches with which the main ones were compared) listed in italics below their full title. These alternative treatments bring the total of approaches evaluated up to about 30. In addition, many of the studies contained 'no treatment' ('normal schooling') control groups, and these are also analysed here: studies with a well defined no-treatment control group are marked with an asterisk.

CHAPTER TWO
SIGNPOSTS

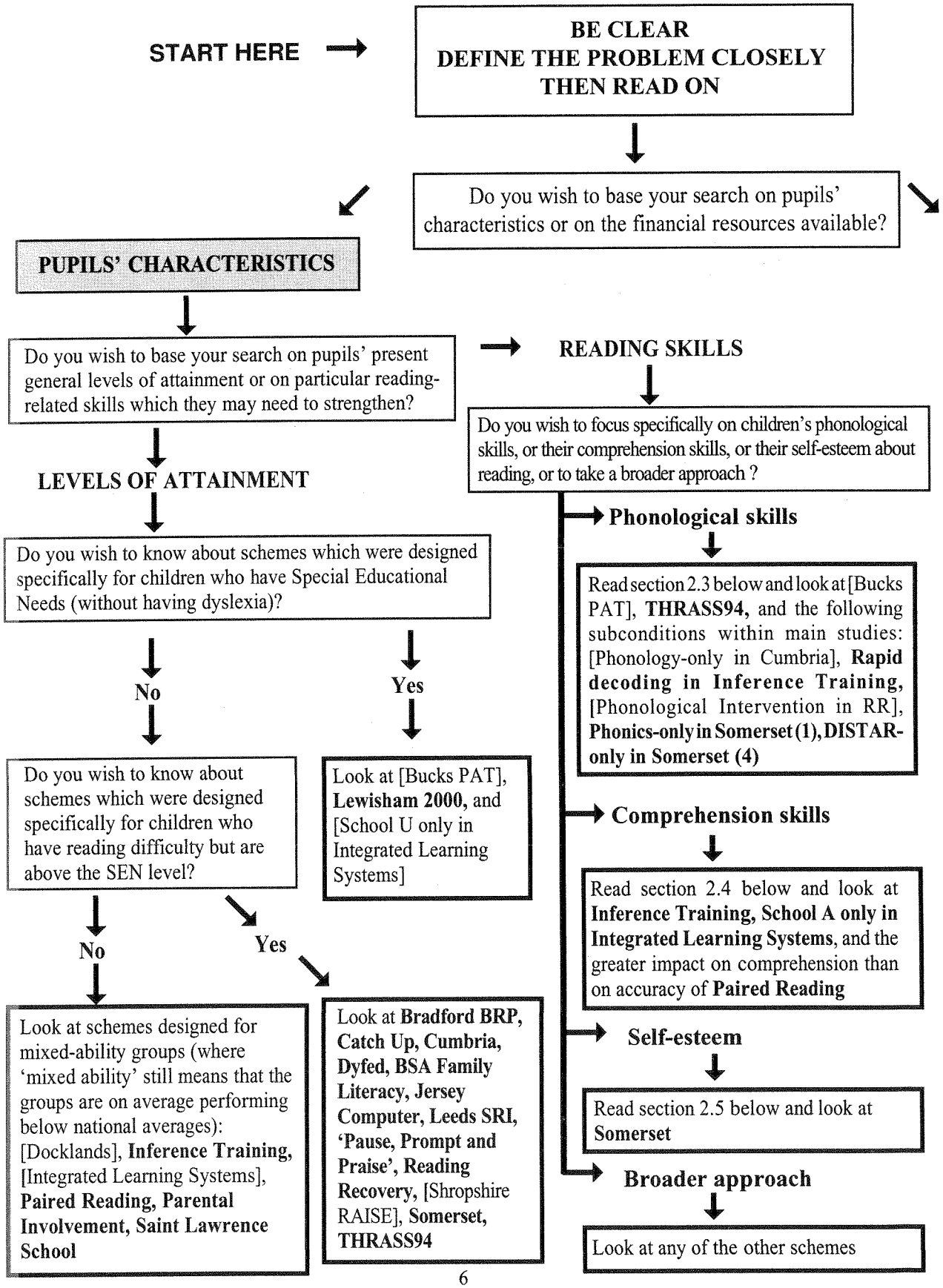
2.1 Finding your way

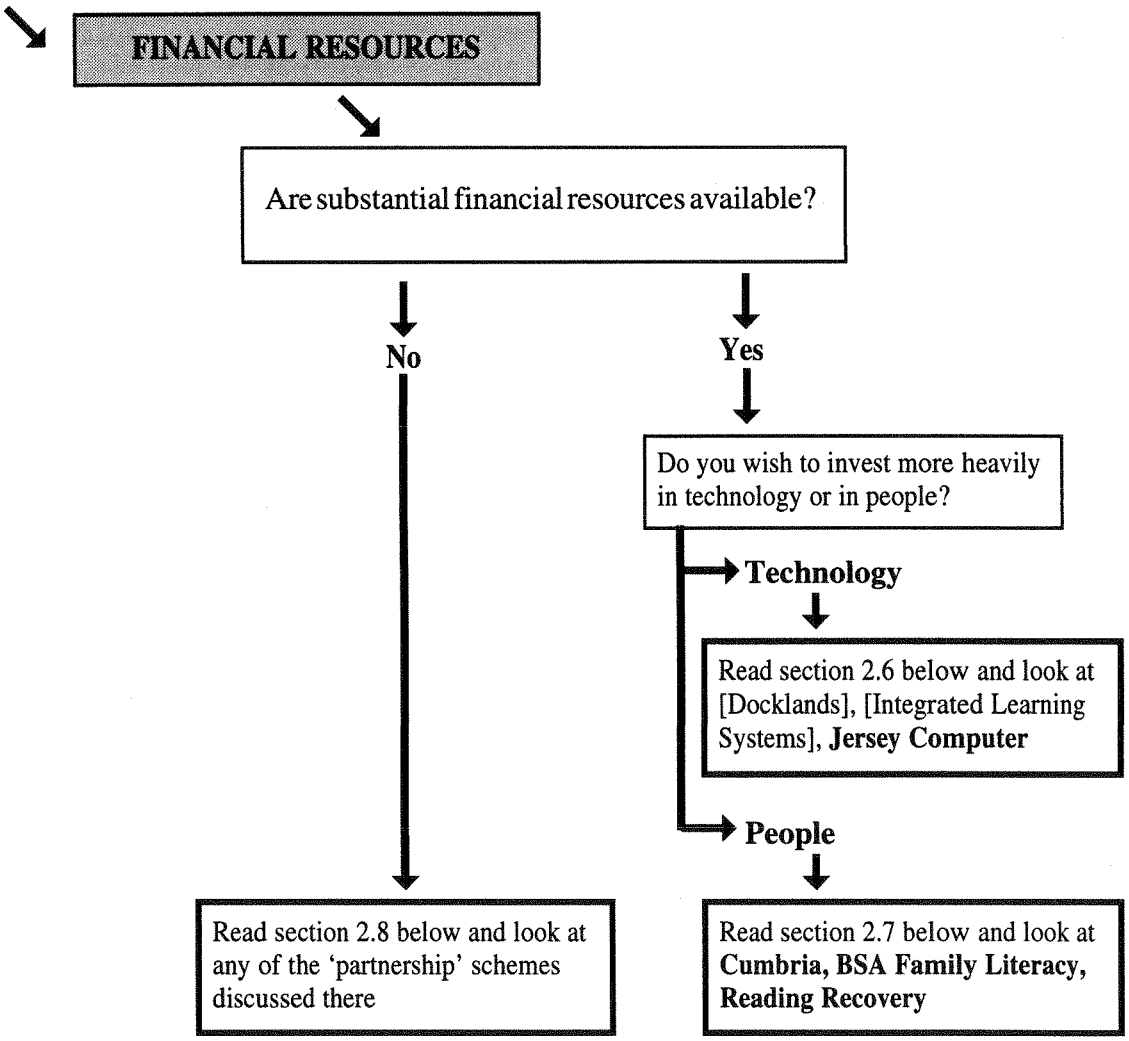
Reading research is a jungle, and quantitative evaluations of early interventions are among its densest thickets. On pages 6-7 we provide some signposts to help you find the schemes which may be most relevant to your situation, in the form of a flowchart. Each exit box in the flowchart gives the names of one or more schemes to consider. All the schemes mentioned are listed and described in alphabetical order in chapter 3 – but remember that these are only the schemes analysed for this book. Some exit boxes also direct you to later sections of this chapter.

Before using the flowchart it would be advisable to read the caveat in section 2.2 on the scale of the various evaluations.

And when using the flowchart **it is particularly important to remember that the flowchart mentions all schemes, whether effective or not**; however, as a rough guide, schemes we consider more effective are shown in the flowchart in **bold**, while those we consider less effective are shown in [square brackets].

FLOWCHART: FINDING YOUR WAY THROUGH THE SCHEMES




FINANCIAL RESOURCES
Are substantial financial resources available?
No
Yes
Do you wish to invest more heavily in technology or in people?
Technology
Read section 2.6 below and look at [Docklands], [Integrated Learning Systems], Jersey Computer
People
Read section 2.8 below and look at any of the 'partnership' schemes discussed there
Read section 2.7 below and look at Cumbria, BSA Family Literacy, Reading Recovery

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.

Table 3: Numbers of children involved in the 20 studies, in decreasing order of number in experimental group

Ref. no.	Name of study	Numbers of children			
		experimental	control	alternative treatment(s)	Total
13	Paired Reading in Kirklees	2372	446		2818
9	Integrated Learning Systems	760?	?		760+
2	Bradford BRP	510?	?		510+
11	Leeds SRI	457			457
6	Docklands	379			379
1	BSA Family Literacy	147			147
19	Somerset (sum of 4 studies)	143	91	225	459
16	Reading Recovery	89	153	91	333
4	Catch Up	74	17	14	105
18	Shropshire RAISE	67	50		117
10	Jersey Computer	61			61
20	THRASS94	59			59
12	Lewisham 2000	57	250		307
14	Parental Involvement	51	152	45	248
7	Dyfed	51	22		73
5	Cumbria	32	31	61	124
17	Saint Lawrence School	32			32
8	Inference Training	26		26	52
3	Bucks PAT	24	24		48
15	Pause, Prompt and Praise	10		10	20

Key: ? = *number not stated clearly or at all in report*
Note: *Where no number is shown, there was no control or alternative treatment group.*

Thus, numerically, Paired Reading nearly outweighs all the rest together – the reason for its huge numbers is explained in the description in chapter 3. 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 – the third smallest study (in terms of number of children in the experimental group) was meticulously designed and reported, while the reporting of some with much larger numbers was considerably less full.

2.3 Focusing on phonological skills

Paired Reading had the largest numbers of children, but phonological skills were the focus of the largest number of studies. Among those analysed here are:

- two main schemes [Bucks PAT] and THRASS94 and
- five alternative treatments [Phonology-only in Cumbria], Rapid decoding in Inference Training, [Phonological Intervention in the Reading Recovery (RR) 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 treatments 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 THRASS94, which gives explicit attention to phoneme-grapheme and grapheme-phoneme relationships within a broad framework, and was reasonably effective.

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

2.4 Focusing on comprehension skills

Most aspects of reading improvement are under-researched in Britain, 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 Britain; 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 study in this set which addressed this issue directly was Inference Training. This 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.

There are two further insights from projects with other main focuses. 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 2.6.

Also, Paired Reading appeared to bring about greater improvements in children's comprehension than in their reading accuracy.

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.

2.5 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 treatment it was very effective. And in the final study drama teaching designed to boost self-esteem plus a specific reading treatment 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.

2.6 Focusing on IT

Results from the three studies using IT as their main resource (Docklands, Integrated Learning Systems, Jersey Computer) are again mixed. These studies varied significantly in scale. Jersey Computer involved only a few children in each of several year groups. At the other extreme, Integrated Learning Systems was a national evaluation by the University of Leicester for the (then) National

Council for Educational Technology, now the British Educational Communications and Technology Agency.

Docklands and Integrated Learning Systems had one thing in common: their main results for the impact of the schemes on reading were 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).

However, there were two contrary findings. The Jersey Computer study did produce highly significant gains. This was a small study, and the number of Early Years children within the total of 61 experimentals cannot be deduced from the report. However, what is striking about the approach itself 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 Integrated Learning Systems, 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 IT approaches require, technology used to boost literacy attainment deserves to be targeted as precisely as possible.

2.7 Large-scale programmes

Three of the schemes covered here merit this description: Cumbria (the Reading and Phonology condition), BSA Family Literacy, 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.

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

2.8 Partnership approaches

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 and most researched partnership approach is Paired Reading, and the effectiveness of this approach has been fully demonstrated. Anyone interested in following this up is recommended to contact the Paired Learning Centre at the University of Dundee.

Catch Up is perhaps the most precisely designed of the partnership schemes. It is targeted specifically on children who achieve level 1 in reading at the end of Key Stage 1 – see chapter 1 of this book for the evidence that about 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.

Other schemes of this general type whose evaluations have demonstrated their effectiveness are Bradford BRP, Dyfed, Leeds SRI, Lewisham 2000, Parental Involvement, 'Pause, Prompt and Praise' and Saint Lawrence School.

On the other hand, there was one partnership scheme which appeared to be less effective, namely Shropshire RAISE. Here the problem may have been that no particular technique was taught or recommended – there was general encouragement to work on reading but the general impression from the report is of a diffuse approach.

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. Such schemes work at least across the whole of Early Years, and certainly also up to

Y6, as the *Knowsley Reading Project* (Brooks, Cato *et al.*, 1996) showed. 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, partnership approaches deserve close consideration.

2.9 The impact of normal classroom teaching

As shown in Table 1, eight of the 20 studies provided evidence on well-defined control groups who received no extra treatment, in other words normal classroom teaching (Bucks PAT, Catch Up, Cumbria, Integrated Learning Systems, Paired Reading, Parental Involvement, Reading Recovery, Somerset (1 and 4)). As shown in the Appendix, three further studies (Bradford BRP, Dyfed, Shropshire RAISE) provided ratio gains for such groups which could not be relied upon in comparisons with the experimental groups' RGs but which were valid in their own right. Data on the impact of normal schooling was therefore available from about half the studies, and the total number of children in the relevant groups was at least 986 (and probably considerably larger, since the control group numbers in Bradford BRP and Integrated Learning Systems were not known).

As the Appendix again shows, most control groups made normal progress. This finding is, however, circular: children receiving normal schooling mostly made the progress to be expected of children receiving normal schooling. What is more interesting is that some control groups made *better* than expected progress despite, apparently, receiving no extra treatment. This finding came from just two studies: Paired Reading, and the Welsh-medium schools in Dyfed. What might these schemes have had in common? It seems that in Kirklees (the LEA where the Paired Reading study was conducted) the experimental treatment affected a significant proportion of schools. Similarly, the Welsh-medium schools in the former county of Dyfed are few in number and form a tightly-knit community.

So it may be that these initiatives affected a high proportion of the schools in the areas in which they took place, and the experimental schools were observed by others. This may have influenced non-participating schools to 'raise their game', and provide 'normal teaching' of a higher effectiveness than usual. Density of coverage 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 normal 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 normal schooling therefore proves the need for early intervention schemes: in general, normal schooling does not enable slow readers to catch up.

2.10 Lasting effects

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, seven of those covered in this book did follow children up at least once. The details are given at the end of the Appendix, and may be summarised by saying that

— **in almost every case the children at least maintained the improvements, and**

— **in some cases the children went on making relative gains.**

2.11 Conclusions

- Normal schooling ('no treatment') does not enable slow readers to catch up.
- Work on phonological skills should be embedded within a broad approach.
- Children's comprehension skills can be improved if directly targeted.
- Working on children's self-esteem and reading in parallel has definite potential.
- IT approaches only work if they are precisely targeted.
- Large-scale schemes, like BSA Family Literacy and Reading Recovery, though expensive, can give good value for money.
- Where reading partners are available and can be given appropriate training, partnership approaches can be very effective.
- Most of the schemes which incorporate follow-up studies continued to show gains.

CHAPTER 3
THE SCHEMES AND THEIR EVALUATIONS

This chapter describes the 20 schemes, 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 20 schemes are summarised in Table 4 on page 16.

Table 4: General characteristics of the schemes

Ref no.	Programme	Age range (see note)	Y1	Y2	Y3	Y4	Duration	Number of sessions for each child in experimental group
1	BSA Family Literacy	3-6 yrs	✔	✔			12 weeks	8 hours a week
2	Bradford BRP	5-12 yrs	✔	✔	✔	✔	10 weeks	3 x 15 mins a week
3	Bucks PAT	5-12 yrs				✔	20 weeks	10 mins daily
4	Catch Up	7-8 yrs			✔		(see p.21)	10 mins a week indiv. + 15 mins a week group
5	Cumbria	6-7 yrs		✔			20 weeks	2 x 30 mins a week
6	Docklands	7-9 yrs			✔	✔	52 weeks	35 mins daily
7	Dyfed	5-6 yrs	✔				38 weeks	3 x 20 mins a week
8	Inference Training	7-8 yrs			✔		4 weeks	2 x 20-45 mins a week
9	Integrated Learning Systems	6-14 yrs		✔	✔	✔	26 weeks	variable
10	Jersey Computer	7-14 yrs			✔	✔	4 weeks	20 mins daily
11	Leeds SRI	5-6 yrs, 11-12 yrs	✔				26 weeks	not stated
12	Lewisham 2000	4-6 yrs 7-8 yrs 11-12 yrs			✔		30 weeks	40 mins a week
13	Paired Reading	5-16 yrs	✔	✔	✔	✔	9 weeks	variable
14	Parental Involvement	6-8 yrs		✔	✔		52 weeks	variable
15	Pause, Prompt & Praise	9-10 yrs				✔	9 weeks	15 mins daily
16	Reading Recovery	6-7 yrs		✔			20 weeks	30 mins daily
17	St Lawrence School	7-8 yrs			✔		39 weeks	15 mins daily reading + 2 x 20 mins a week literacy games
18	Shropshire RAISE	6-9 yrs			✔		20 weeks	4 x 25 mins a week
19	Somerset (1)	8-9 yrs				✔	26 weeks	20 mins a week/indiv + 30 mins a week/group
	Somerset (2)	8-9 yrs				✔	18 weeks	30 mins a week/indiv. + remedial
	Somerset (3)	7-9 yrs			✔	✔	17 weeks	30 mins a week/indiv. + remedial
	Somerset (4)	8 yrs			✔		20 weeks	45 mins a week/pair + 3 x lhr a week/group
20	THRASS94	6-9 yrs, 10-11 yrs		✔	✔	✔	22 weeks	30 mins daily

Note: The age-range shown is that for which the scheme was designed; the year-groups shown are those within Years 1-4 for which the scheme has been evaluated and analysed for this report.

3.1 Basic Skills Agency's Family Literacy 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, and stemmed from the fact that children whose parents have low literacy are themselves more likely to experience literacy difficulties, thus creating a cycle of deprivation. The Basic Skills Agency devised the Family Literacy 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 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 was commissioned to evaluate the initiative in the four school terms from Summer 1994 to Summer 1995. For the purposes of this book, 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 both at entry and on completion of the course. Data were available from this test on 147 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.

References

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

3.2 Bradford Better Reading Partnership

The Better Reading Partnership aims to help children to become better readers by providing explicit training for adults. The 15-minute sessions, which occur three times a week, follow a common structure and focus on the development of independent reading strategies.

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 is a vital component in this partnership. They recruit adults from existing curriculum support staff and parent volunteers already helping in the school. The two-day training course includes a direct observation using a one-way viewing facility. The ongoing training is supported by a project co-ordinator who meets with the partners to discuss the development and progress pupils are making and consider new aspects of the reading process. In 1997, the partners have already worked with 1649 children. The forecast for the full project is that it will be provided for about 4000 children.

The partners work with the pupils for ten weeks. They read together for 15 minutes, three times a week. Each reading session follows 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 is on the development of independent learning strategies. The reading partners are encouraged to discuss the text with the pupil and are trained to prompt the pupil to problem solve difficulties and to develop reading behaviours that will 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)

Contact

Bradford Inspection and Consultancy Services
T.F. Davies Professional Development Centre
Rosemount
Clifton Villas
Bradford
BD8 7BY

3.3 Buckinghamshire Phonological Awareness Training

In developing the PAT programme (Wilson, 1993) 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 of the Code of Practice; and 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 were blind to the children's treatment groups. The programme ran for 20 weeks. Groups of six met four days a week for twenty minutes.

The results were not clear-cut. The experimental children did make significantly more progress than the controls; but the experimentals made scarcely any more progress than would have been expected from normal schooling and development.

References

Wilson (1993); Wilson and Frederickson (1995)

3.4 Catch Up

Catch Up targets Year 3 children who obtained Level 1 in reading in the national tests at the end of Key Stage 1. It is recognised that for these children to meet National Curriculum requirements, an intensive yet manageable programme is of critical importance. Catch Up is a 10-minute structured teaching programme that is carried out once a week with individual children by the class teacher.

Scheme

Catch Up was developed as a result of a study undertaken by the project consultants. 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 session is divided into three discrete sections. The first two minutes can be called the *prepared reading approach*. A book 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. This is where the teacher acts upon the information gleaned and decides which skill should be worked on.

Individual sessions are followed by a group reading session lasting 15 minutes, once a week. This ties in with the National Literacy Strategy. The aim is to enable the children to read with fluency and intonation. In addition Catch Up has produced a CD ROM game to enable the children to practice literacy skills, and there is a pupil record book intended to make record keeping simple, a summative assessment of progress which is carried out at the end of each term, and a Catch Up telephone hotline to offer support to teachers. The whole programme is designed to last a full school year.

Evaluation

The full Catch Up programme is intended to last a whole school year. However, the only evaluation to date is of 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 (1997)

Contact

School of Education
Oxford Brookes University
Wheatley Campus
Oxford
OX33 1HX

3.5 Cumbria Reading with Phonology Project

This 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 were assigned to one of three sets undergoing teaching. Group one received training in phonological skills and help in learning to read. Pupils in the second set received teaching in reading alone. The third set 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 'blind' to 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 treatment was Reading Recovery.

Reference

Hatcher *et al.* (1994)

3.6 Docklands Learning Acceleration Project

The London Docklands Development Corporation came together, in 1995, with the National Literacy Association to fund a two-year project involving over 600 pupils in 15 schools. Daily sessions of English and Maths using the latest multi-media instructional software were designed to enhance and reinforce learning as well as improving attendance, behaviour and enthusiasm. Each child was given an initial assessment which set the most suited level. Teachers were able to access data from the computers in order to monitor individual progression.

Scheme

In 1995 when the project began the children were reading on average 10 months behind their chronological age. It was feared that, if they continued at this rate until they were due to leave primary school, they would be more than two years behind in their literacy acquisition. The main objective of the project was to double the rate of acquisition for the 600 children taking part.

The project encompassed 15 schools, aiming to improve the literacy of 600 Year 3 and 4 children. The project, run in collaboration with The National Literacy Association, used the latest in multi-media to motivate the children. Pupils were equipped with Acorn Pocketbook Computers, which are a modified version of the Psion Organiser. In some schools 90 per cent of the parents attended training days to help their children get the most out of the latest technology.

The project sought to carry out its aims by ensuring daily sessions of Integrated Learning Systems English and Maths, so that a foundation would be achieved. Open Integrated Learning Systems is a structured learning and recording package which manages children's learning. The child first sits an initial assessment which sets an appropriate level of work. If the child encounters difficulties the machine explains the principles. Teachers can access data from the computers to assess the progress of their pupils and create individual workplans.

Evaluation

The report available for the compilation of this book covered the evaluation of the first year of the project (1995-96); the evaluation was carried out by researchers from the London Institute of Education. Data were collected about spelling attainment, reading attainment, sex, free school meals, ethnicity and English as a first or second language. Even allowing for background factors, progress was not impressive. The best result was on the Suffolk reading test,

where children made on average 13 months' gain in 12 months (the other results were poorer). The evaluators nevertheless argued that this was a good result, given that usually children in such schools make only a few months' progress in a year. However, the result still means that these children made no more progress than would be expected on average nationally, despite the extra technology and resources.

Reference

Scott *et al.* (1997)

Contact

The Urban Learning Foundation
56 East India Dock Road
London
E14 6JE

3.7 Dyfed Improving Reading Standards in Primary Schools Project

Dyfed Education Authority developed this project to examine the efficacy of thinking skills in helping to improve reading in Year 1 children. In order to assist the transfer of thinking skills into the reading domain, a reading activity lasting 20 minutes was devised for the class teacher to deliver three times a week to a small group. The scheme has been published in both English and Welsh, and consists of an explanation of the approach, a set of booklets giving suggestions to teachers, and a video to complement the material.

Scheme

Dyfed recognised that the ability to think and solve problems is essential for full participation in the National Curriculum. The aim of this project is to evaluate the efficacy of early reading initiatives, based upon thinking skills developed in the prevention of different reading difficulties. The study comprised two approaches. One experimental group received both extra reading activities and a thinking skills programme, while the other received only the reading activities. Group three was a control group receiving no additional intervention. There were English-medium and Welsh-medium versions of all three groups. Those receiving only reading activities were children identified as being at risk of reading difficulty.

The class teacher delivered a reading activity, lasting 20 minutes, three times a week, to a group of backward readers. The children read a short book together. After the initial reading pupils took it in turns to read. The other pupils were expected to follow the text being read by finger pointing. If a difficulty with reading a word was encountered the teacher encouraged the pupils to help each other tackle the problem.

Evaluation

The evaluation was carried out by Dyfed LEA. The Dyfed Early Reading Check, which consists of testing five developmental areas (language, learning style, memory, number and perceptual motor skills), was devised and administered to all pupils in order to identify areas of difficulty. The Reading Check also had the purpose of evaluating the efficacy of the scheme.

Six schools delivered thinking skills plus reading and six schools delivered the reading-only approach The progress made by pupils in the two intervention groups was similar, and not significantly better than the control group.

Reference
Dyfed County Council (1994)

Contact
Dyfed County Council/Cyngor Sir Dyfed
Education Department/Adran Addysg
Pibwrlwyd
Carmarthen/Caerfyrddin, Dyfed
SA31 2NH

3.8 Inference Training

This scheme focuses upon the band of children who fall within the normal IQ range 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 treatment: 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 treatment 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 20-45 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)

3.9 Integrated Learning Systems, National Council for Educational Technology study

Integrated Learning Systems is the general name for a number of initiatives (including Docklands, see above). The study summarised here is the one mounted by the National Council for Educational Technology (NCET) and evaluated by a team led by staff of the University of Leicester.

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 numeracy and literacy. 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.

Evaluation

As stated, the project was evaluated by a team led by staff of the University of Leicester. Specifically, this account is based on the report of Phase II of that evaluation (1994-96). (Phase I was smaller, though its conclusion for literacy was much the same; Phase III (1996-97) did not involve any children in Years 1-4.) Though the results of Phase II for numeracy were reasonable, those for literacy were unimpressive: no overall benefit compared to control groups. Two exceptions were noted: a Special School where controls outperformed experimentals, and a mainstream primary school where the ILS group did significantly better than controls, 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 *et al.* (1994); Underwood (1997); Underwood and Brown (1997)

3.10 Jersey Computer Assisted Reading Development Programme

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 experimental children (and from nine controls - 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 changes in pupil behaviour (which were not the prime focus of the project) were also reported.

Reference

Jersey Advisory Service (1993)

Contacts

Jersey Advisory Service
P.O. Box 142
Highlands
St. Saviour
Jersey
JE4 8QJ

Talking Systems
22 Heavitree Road
Exeter
Devon
EX1 2LQ

3.11 Leeds Sustained Reading Intervention

This three-year project aimed to detect children who may escape the notice of a teacher in a large classroom, often the quiet ones who were not poor enough for statements of Special Educational Need or grave concern but who found it difficult to keep up with the rest of the class and would have steadily fallen behind throughout the year. By increasing teachers' knowledge of the reading process, techniques of assessment, materials that can be used to promote active and reflective reading, and the development of a home-school partnership, reading standards were successfully raised.

Scheme

Twenty-eight primary schools and seven secondary schools were involved in the intervention. The programme ran in year blocks for three years. It was developed in two phases, training and teaching. Teachers were nominated from each school and trained as a Co-ordinator. The training gave guidance in relating theories of reading to classroom practice, techniques of assessment, materials to promote active and reflective reading, and ways of involving parents and issues relating to school policy were covered.

Details of the teaching approach are not given in the report.

Evaluation

Leeds LEA carried out its own evaluation. The Carver Word Recognition Test and the Salford Reading Test were used to select the pupils who were to be involved in the study. These pupils were re-tested each year for the duration of the scheme.

The project seemed to be most effective for primary school pupils. In the first year of the project, 58 per cent of pupils gained over 24 months of reading age, and 28 per cent gained over 36 months. In the second year, 68 per cent of pupils gained over 24 months, and 27 per cent gained over 36 months. The qualitative improvements were substantial: increase in confidence, increased motivation and enthusiasm as well as improved social skills with peers.

Reference

Birch (1995)

Contact

Elmete Professional Development Centre
Elmete Lane
Leeds
LS8 2LJ

3.12 Lewisham Literacy 2000

Education achievement levels in Lewisham were well below the national average. In 1994, Lewisham was 103rd out of the 115 LEAs in the national league tables. In order to improve upon this the Literacy 2000 project was set up in 1995, to run until the year 2000. Participating schools are asked to target specific pupils who they deem to be at risk of under achievement. These pupils are set targets in terms of the increase in literacy attainment.

Scheme

Literacy 2000 aims to raise the standards of literacy by involving parents and enhancing teachers' skills to bring about effective change. The project is run in collaboration with national agencies funding various activities throughout the course of the project: the Basic Skills Agency, Reading is Fundamental and the Department for Education and Employment are involved. Due to the fact that the literacy project will run over five years it has been broken up into four phases of development. Phase 1 involves a central training programme across the curriculum, promoting a home-school partnership. Phase 2 focuses on the school's resources and how they can be incorporated into training. Phase 3 embeds the project in the School Management Plan. Phase 4 allows the project to become part of the 'norm' of the school culture.

The project is managed by a co-ordinator who receives training in management skills. Target teachers receive training in the teaching of literacy for the first year, and from the second year onwards each school receives school-based training for all staff. The training programme focuses on many areas, including language acquisition, phonological awareness, strategies to involve parents, and homework activities. Classroom assistants are trained to support pupils' reading development. A Literacy 2000 family liaison worker is involved in activities such as homework projects. By introducing a library project, which enables libraries to target their stock to support the target year groups, library membership has increased by 33 per cent. Weekend and holiday events are organised to encourage family involvement in libraries.

The project provides general guidance to schools on the targeting of pupils. There are no fixed rules about the size of the target group. Usually the pupils receive one 40-minute session per week, during which they are withdrawn from class. Such withdrawals are planned to take place at different times in order to minimise the time lost by any one subject.

Baseline information was collected on factors that could be said to affect achievement such as gender, free school meal status, special educational needs stage, library membership and ethnicity. Schools were asked to submit a

project plan giving details of project activities including evidence of inclusion in the School Management Plan, description of activities, roles and responsibilities, assessment, monitoring and evaluation.

Evaluation

The University of Exeter was commissioned to carry out an evaluation of the first year of the project (1995-96). Though the further stages of the project are known to cover all children in a number of Lewisham primary schools, the experimental group for the first year evaluation consisted entirely of children at Stages 1-2 of the Code of Practice. No statistical test was reported, but from the effect size it seems clear that the experimentals made significantly more progress than the controls, who were the remaining children in the same classes in the same schools.

Reference

Wray and Rogers (1997)

Contact

Quality Assurance and Development
Lewisham Professional Development Centre
Kilmorie Road
London
SE23 2SP

3.13 Paired Reading in Kirklees

This is one of the simplest schemes yet devised, and the subject of by far the largest evaluation.

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). 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.

References

Morgan (1976, 1986); Topping (1988, 1990); Topping and Lindsay (1992); Topping and Wolfendale (1985); Wolfendale and Topping (1996)

Contact

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

Tel: 01382 344000
Fax: 01382 229993

3.14 Parental Involvement in Haringey

'Learning at mother'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 Britain and beyond. For the purposes of this book, the first well-known and -researched scheme of this sort in Britain, 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.

References

Hewison (1985); Tizard *et al.* (1982); Topping and Wolfendale (1985); Wolfendale and Topping (1996)

3.15 Pause, Prompt and Praise

This approach requires that tutors should: pause (for up to five seconds) when the reader hesitates or makes a mistake; prompt after the pause, if the pupil has not correctly identified the word; and praise the child for progress and self-corrections.

Scheme

Research at the University of Auckland, New Zealand, inspired by the earlier work of Clay on Reading Recovery, resulted in the development by Glynn *et al.* (1979) of a behavioural method for tutoring oral reading now known as 'Pause, Prompt and Praise'.

'Pause, Prompt and Praise' has several requirements. Adequate reading material must be provided. Tutors must provide appropriate feedback as they listen to the child read. Mistakes are to be expected; they are an important part of the learning process. 'Self-correction is a part of the good reading process. When the child makes an error it is important to pause for at least five seconds, giving ample time to self-correct.'

Given the constraints of teachers in dividing their time between all pupils, others such as parents and more advanced peers can be trained in the appropriate tutoring skills.

Evaluation

Wheldall has carried out several evaluations of the technique in Britain, and they mostly show that it is effective. The only report found which focused on poorer readers of an age relevant to this book was a very brief description of what the authors call an unsuccessful study (see Wheldall and Colmar, 1990, p.130). However, a fresh look at the findings suggests that it was not entirely unsuccessful: poorer readers in Y4 did make progress in one-to-one sessions with 13-year-old tutors, but an equal amount whether the tutors were trained in Pause, Prompt and Praise or not. Both results reinforce the importance of partnership approaches.

References

Glynn *et al.* (1979); Wheldall and Colmar (1990)

3.16 Reading Recovery in London and Surrey

Reading Recovery arose out of an extensive research project carried out by Clay in New Zealand. It identifies children who are having difficulty 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 an early intervention programme devised by Marie Clay in New Zealand between 1976 and 1979. It 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. Pupils 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 discontinued are referred for more detailed assessment and specialist help.

The first LEA in Britain 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 two Education and Library Boards in Northern Ireland.

Evaluation

The original request for information for this book produced more replies about Reading Recovery than about any other initiative, and these constituted about a quarter of all of the information received. 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). However, none of these reports provided either an impact measure or data from which such a measure could be calculated. The only report which did provide

such information was the London Institute of Education study of Reading Recovery in six London LEAs and Surrey (Sylva and Hurry, 1995). Since that study also included both carefully-chosen control groups and an alternative treatment condition (Phonological Intervention), it has been taken to represent Reading Recovery for the purposes of this analysis.

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

The Sylva and Hurry study showed that Reading Recovery children made significantly greater progress than either Phonological Intervention children or controls; and that Phonological Intervention children did not make significantly greater progress than controls.

References

Clay (1979, 1985, 1993); Gardner *et al.* (1997); Hurry and Sylva (1998);Prance (1992); Sylva and Hurry (1995); Wright (1992)

3.17 Saint Lawrence School, Towcester, Northants

Saint Lawrence School wanted to improve the literacy skills of Y3 pupils with lower reading ages than chronological ages. The main focus of the project was that all children would take part in reading activities for 15 minutes a day with a reading partner - preferably an adult. Peers were also used twice a week before school to listen to others reading.

Scheme

Saint Lawrence School was awarded a sum of money from their LEA's 'Raising Achievement in Literacy and Numeracy at KS2' fund. The main aims were to raise achievement of all Y3 pupils, with a focus on children with a reading age six months or more below their chronological age.

The project began by raising awareness and understanding. A letter was sent out to parents giving a description of the broad aims of the project, an intended outline, and arrangements for a meeting where queries could be clarified. Targets to achieve these aims included reading conferences to set 'parent friendly' targets. By encouraging parents to play an active role in their child's development, the emphasis on a home-school partnership had to foster aspects of pleasure and enjoyment rather than being seen as an effort.

All Y3 pupils and their parents were invited into school 20 minutes early on Tuesday and Thursday mornings to play board games such as Dominoes and Snakes and Ladders. The help of Y6 pupils was enlisted to carry out this part of the project. Such games were not only enjoyed by the children but encouraged the use of language and communication skills.

The profile of reading was raised by introducing guest speakers and story tellers. Training for adult volunteers was provided by the LEA English team. The use of volunteers enabled children's reading to be heard daily. Pairing a child with an adult for a 15-minute session at the same time every day was set up to give consistency.

Evaluation

The children's reading ages were calculated using the Salford reading test, both at the beginning of the project, in September 1996, and at the end, in June 1997. On average, the children had made about twice as much progress as would be expected in that time.

Reference
Irwin (1997)

Contact
Saint Lawrence School
Islington Road
Towcester
Northants
NN12 6AU

3.18 Shropshire: Raising Attainment In Shropshire Education

RAISE involved a group of Shropshire schools working together with an advisory teacher to assess attitudes and achievement, implement strategies to raise attainment, and then evaluate these strategies. Pupils who may not have been achieving the average standards for their age or capability were targeted. In collaboration with the University of Sheffield, teachers were trained in classroom organisation and the variety of approaches to teaching reading.

Scheme

Raising Achievement In Shropshire Education focused on the reading capabilities of Year 3 pupils. Reading Recovery was deemed too expensive to use in terms of funding and human resources. Six schools took part in piloting RAISE. Teachers were trained on classroom organisation and the variety of approaches to teaching reading. Participating schools received sufficient finance to employ supply teachers for three days of training and £100 towards the purchase of some recommended books.

It was agreed that pupils should be taught in groups of four to six for 25 minutes on four occasions a week. In all instances teaching was to be undertaken by the class teachers who had undergone the training. There was unanimous agreement that the project should remain targeted to the groups of pupils selected originally and not refocus on different groups.

Evaluation

Shropshire LEA undertook its own evaluation of the initiative. The 1996 report from the Authority gives data on three cohorts of children, who had taken part in October 1994-March 1995, October 1995-March 1996, and March-July 1996. Unfortunately, the data for the second and third cohorts are given in a form which does not permit the calculation of an impact measure. Judgment on this scheme is therefore based solely on data from the first cohort, and within that on the largest year group involved, Y3. The data show that the RAISE children made greater progress than the controls; but also that the RAISE children made scarcely more progress than would be expected on average from normal schooling (the controls, meanwhile, fell behind 'normal' progress).

Reference

Jones (1996)

Contact

Shropshire County Council Education Services
The Shirehall
Abbey Foregate
Shrewsbury
SY2 6ND

3.19 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 then moved to Western Australia). 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:

> *It is curious that despite the fact that the retarded reader's most outstanding characteristic appears to be his poor emotional adjustment, remedial reading so often takes the form of a direct attack on the mechanics of reading.* (Lawrence, 1973, p.11)

Schemes

Like Inference Training, 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 experimental group received some form of Rogerian self-esteem counselling *plus* a specific reading treatment. 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 treatment was the remedial teaching already provided within the school, which was mainly phonics. A control group received no extra treatment, and there were two alternative treatment 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 experimental groups received counselling plus remedial teaching, while the controls received only remedial teaching (and were therefore alternative treatment, rather than no-treatment 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-treatment control group, and three groups who all received DISTAR. One group received only that treatment, while the other two received in addition one of two 'therapeutic' treatments 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 experimental group receiving the first of the 'therapeutic' treatments 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 headteachers 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 treatment was an accepting and non-judgmental relationship between counsellor and children.

The group receiving the second of the 'therapeutic' treatments 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 treatment 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 treatment, which also did not differ.

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

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

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

The specific reading treatment alone was ineffective in study 3, reasonably effective in study 4 (though no more so than no treatment, and significantly less so than the therapeutic treatments), 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 treatment can be effective, and that the boost to self-esteem can also (and perhaps more cost-efficiently) be delivered through appropriate drama teaching. We believe his approach has potential, has been undeservedly neglected, and should be revived. The need for motivational factors in poor reading to be re-explored is heightened by the current anxiety over boys' low achievement compared to girls, and the possibility that part of the reason may be boys' negative attitudes to reading.

References

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

3.20 The Handwriting, Reading and Spelling Sequence (THRASS Writetrack 1994)

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 acronym originally stood for 'The Handwriting, Reading and Spelling Sequence', but more recently the title has become 'Teaching Handwriting, Reading and Spelling Skills'. The programme itself has also been continuously developed and revised (see Davies and Ritchie, 1996), 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.

In the version which has been most fully evaluated, the scheme took place over two terms in 12 urban schools which were selected from a variety of areas in Sheffield. The children varied from Y2 to Y6, giving the study the opportunity to gauge the effectiveness of the programme at different ages. Teachers were 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

Later versions of THRASS seem not yet to have been formally evaluated in Britain. The information used for this analysis is therefore taken from the evaluation conducted by Johnson (1995), a Reading Recovery tutor in Sheffield, for the City of Sheffield Education Department, on the version known as THRASS Writetrack 1994. In general, the results showed good impact on children's comprehension scores, and less, though still some, impact on accuracy.

References

Davies and Ritchie (1996); Johnson (1995)

Contacts

THRASS Training International
11A Kilmorey Park
Chester
CH2 3QS

City of Sheffield Education Department
Service for Pupils with Learning Difficulties
East Hill Secondary School
East Bank Road
Sheffield
S2 3PX

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 very neatly designed and meticulously reported project on computer-generated speech feedback (Davidson and Noyes, 1995) was nevertheless too small to be considered significant (five experimental and five control children). However, its entirely non-significant result (the two groups made exactly equal gains, p=0.997) was consistent with those from the two large-scale IT studies (Docklands, Integrated Learning Systems);

— A further study in the Somerset self-esteem series (Lawrence and Blagg, 1974) was also too small. Six experimental children received counselling from non-professionals plus board games designed to enhance reading; six in a first alternative treatment (AT) group received counselling from non-professionals only; six in a second AT group received the board games only; and six in a third AT group played non-reading games (there was no no-treatment 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 young poorer readers;

— The Knowsley Reading Project (Brooks, Cato *et al.*, 1996) and the government's Summer Literacy Schools (Sainsbury *et al.*, 1997) did not count as early interventions because they were for children in Y6;

— Literacy Acceleration (Lingard, 1993, 1997) has gathered copious data but so far exclusively at secondary level;

— Several schemes on which quantitative research has been carried out with children of relevant ages in Britain 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, Literacy for All, which is the British version of the US programme 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 Britain.

4.2 Pupil numbers

Some of the evaluations with small pupil numbers cited here were not designed to be large-scale studies, and their small numbers are therefore appropriate. However, this study includes several schemes which are well-known and in widespread use, yet still have evaluations based on small pupil numbers. The various reasons are as follows:

— DISTAR: This has a large evaluation literature, but mainly from the USA. The only evaluation located in which the subjects were lower-attaining British readers of a relevant age was embedded within the Somerset (4) study by Lawrence (1985). In this, the number of DISTAR- only subjects was 84, though there were two other experimental groups, of 94 and 79, who also received DISTAR, plus a treatment designed to boost self-esteem (counselling and drama, respectively);

— '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 treatment group;

— Reading Recovery: The literature on this scheme is voluminous, and the information on it from Britain constituted about a quarter of all the information received for this report. Several thousand children have now passed through Reading Recovery in the UK alone, and many thousands have done so in New Zealand, Australia and the USA. But almost all the British reports received gave data only as raw scores without

standard deviations, or in terms internal to the scheme and meaningful only within it, in particular 'percentage of pupils successfully discontinued'. Only two of the British studies, Sylva and Hurry (1995) on London and Surrey, and Gardner *et al.* (1997) on Northern Ireland, reported data from tests yielding reading ages or standardised scores; and only the London and Surrey study had a control group (it also had an alternative treatment, Phonological Intervention). It was therefore appropriate to treat the London and Surrey study as the exemplar for Reading Recovery - but that study had only 89 children in its main experimental group, the group which had a matched control group in different schools. Hence the apparently low sample size for this evaluation of a widely-used scheme;

— THRASS: This was developed in Britain, but has been evaluated more systematically in Western Australia than here. Apart from some tiny informal studies, the only evaluation in Britain was carried out in 1995 on the original version, THRASS Writetrack 1994, and the total number of children in that study was 59, across three school years.

4.3 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 this book varied from the meticulous to the appalling. The number of studies excluded from the analysis (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.).

This is not to say that those studies which we have included necessarily told us everything we needed to know in order to tabulate just the most salient characteristics. 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 20 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 age data and not standardised scores. Though apparently easier to interpret, reading ages are statistically unsatisfactory – for example, establishing whether a gain in test scores is statistically significant is more difficult for reading ages than for standardised scores. Reading 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 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 20. 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 treatments. But the total is still not impressive, and it is to be hoped that anyone currently devising an intervention scheme will automatically consider the necessity for, and commission, an evaluation. The government is setting an example here, with its evaluations of the Summer Literacy Schools and of the National Literacy Project and Strategy.

4.4 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-age tests.

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 treatment groups
- how children were assigned to the different groups, for example randomly or by matching
- the nature of any alternative treatment
- the exact length of the intervention
- the reading 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
- 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 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 treatments (including no treatment) 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

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: United Kingdom Reading Association.

BIRCH, J. (1995). GEST 'Raising Standards in Inner City School' Leeds Project – Sustained Reading Intervention. Leeds: Elmete Professional Development Centre (mimeographed).

BROOKS, G., CATO, V., FERNANDES, C. and TREGENZA, A. (1996). *The Knowsley Reading Project: Using Trained Reading Helpers Effectively.* Slough: NFER.

BROOKS, G., GORMAN, T.P., HARMAN, J., HUTCHISON, D. and WILKIN, A. (1996). *Family Literacy Works: the NFER Evaluation of the Basic Skills Agency's Family Literacy Demonstration Programmes.* London: Basic Skills Agency.

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.

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. (1998). Statistical information supplied on Autumn 1997 pilot evaluation of Catch Up.

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

DAVIDSON, J. and NOYES,, P. (1995). 'Computer-generated speech feedback as support for learning instruction.' *Support for Learning*, **10**, 1, 35-9.

DAVIES, A. and RITCHIE, D. (1996). *THRASS*. Glasgow: Collins.

DYFED COUNTY COUNCIL (1994). Improving Reading Standards in Primary Schools Project. Final Report, GEST 1993-94. Carmarthen: Dyfed County Council Education Department (mimeographed).

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

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.

GLYNN, T., McNAUGHTON, S.S., ROBINSON, V. and QUINN, M. (1979). *Remedial Reading at Home: Helping You to Help Your Child*. Wellington: New Zealand Council for Educational Research.

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

GREAT BRITAIN. DEPARTMENT FOR EDUCATION (1994). *Testing 7 Year Olds in 1994: Results of the National Curriculum Assessments in England*. London: DFE.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION AND EMPLOYMENT (1996a). *Results of the 1995 National Curriculum Assessments of 7 Year Olds in England*. London: DFEE.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION AND EMPLOYMENT (1996b). *Results of the 1996 National Curriculum Assessments of 7 Year Olds in England*. London: DFEE.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION AND EMPLOYMENT (1997). *Results of the 1997 National Curriculum Assessments of 7 Year Olds in England: a Summary for Heads and Governors*. London: DFEE.

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.

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.

IRWIN, C. (1997). Report to Parents on Literacy Project. Towcester: Saint Lawrence School (mimeographed).

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 (mimeographed).

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 (mimeographed).

JONES, M. (1996). Raising Standards of Reading. Shrewsbury: Shropshire County Council (mimeographed).

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, 48-51.

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.

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.

LITERACY TASK FORCE (1997). *The Implementation of the National Literacy Strategy*. London: DFEE.

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.

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.

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

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

SAINSBURY, M., MASON, K., WHETTON, C. and SCHAGEN, I. (1997). *Evaluation of Summer Literacy Schools*. London: QCA.

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.

SCOTT, D., HURRY, J., HEY, V. and SMITH, M. (1997). The Report of the Evaluation of the National Literacy Association (NLA) Docklands Learning Acceleration Project. London: NLA Docklands Learning Acceleration Project (mimeographed).

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

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. (1988). *Paired Reading Training Pack.* Third edn. Huddersfield: Kirklees Paired Learning Project.

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

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

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.

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.

WRAY, D. and ROGERS, S. (1997). 'The Lewisham Literacy 2000 Project External Evaluation: Interim Report.' In: LEWISHAM EDUCATION *Project Management Plan and External Evaluation: Interim Report*. Lewisham: Lewisham Education Quality Assurance and Development.

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

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 20 studies, in alphabetical order. Under each evaluation are listed the salient statistical and related data used in the analysis reported in this book. Before the log, the organisation of the entries is described; and that description is followed by a number of notes of clarification.

After the log of the 20 studies, the basis on which schemes have been compared is explained. The final section gives brief details of follow-up data from seven studies.

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 treatment group, where there was one	
Nature of alternative treatment	
For each group, number of schools, where known	
Whether groups were equivalent	2)
Length of intervention in weeks	
Reading test(s) 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 treatment groups, where known	9)

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) attainment, which may in some 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**: Seven of the studies (Catch Up, Cumbria, Inference Training, Parental Involvement, 'Pause, Prompt and Praise', Reading Recovery, Somerset) had alternative treatment 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.

   Thirteen studies had a no-treatment control group – but in five of these cases (Bradford BRP, Dyfed, Jersey Computer, Lewisham 2000, Shropshire RAISE) 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). The Jersey Computer control group was so small (9 children) 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 seven of the eight remaining control-group studies (Bucks PAT, Catch Up, Cumbria, Paired Reading, Parental Involvement, Reading Recovery, Somerset (1 and 4)), the control groups were set up in ways which did make their results reliable as comparisons for the experimental groups.

   This leaves Integrated Learning Systems. 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 test results have been analysed here, on the grounds that the main focus of this enquiry is interventions designed to boost reading attainment. Some tests yield more than one reading 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 tests cited have been classified as giving measures of reading accuracy.

4) **Range of tests used**: In all, at least 19 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.

Table A.2: Reading tests used in the 20 studies

Test	Number of studies
British Ability Scales Word Reading Test (BASWRT), various forms	6
Burt, various editions	2
Carver Word Recognition Test	1
Customised tests within computer programs, yielding US grade levels	1
Hodder and Stoughton Literacy Baseline	1
Holborn	1
Macmillan Individual Reading Analysis	1
Neale Analysis, various editions and forms	6
New Macmillan Reading Analysis	1
NFER Reading Test A	1
NFER-Nelson Group Reading Test 6-12	1
Peabody Individual Achievement Test	1
Prawf Darllen Clwyd	1
Primary Reading Test	1
Salford, various forms	4
Schonell	4
Standard (Daniels & Diack) Reading Tests, 2 levels	1
Suffolk	2
Widespan	1
(not stated)	1

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

5) The units in which average scores and s.d's are stated are almost always either reading ages or standardised scores, never both. Raw scores were used in just one case, Somerset (4). Where units are r.a's, scores are given in years and months – this has in some cases required conversion from months or from whole and decimal years.

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

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

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 only in the Paired Reading report. However, the r.a. data required for calculating RGs were given in the following reports: Bradford BRP, Bucks PAT, Catch Up Cumbria, Docklands, Dyfed, Inference Training, Leeds SRI, 'Pause, Prompt and Praise', Reading Recovery, Saint Lawrence School, Shropshire RAISE, Somerset (1-3), THRASS94. In addition, for Integrated Learning Systems it was possible to calculate the RG for one school out of seven.

The following reports did not use tests which yielded r.a's, and therefore RGs could not be calculated for them: BSA Family Literacy, Integrated Learning Systems (for six schools out of seven), Jersey Computer, Lewisham 2000, Parental Involvement, Somerset (4).

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, 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 no attempt has been made to calculate an effect size from its data.

Even where the control group was properly set up the choice of s.d. to use as the divisor in the formula is still problematice, 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 Integrated Learning Systems 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 *pre*-test s.d. may be the appropriate divisor.

In practice, for this book, much of the debate over which s.d. to use became irrelevant, since only three 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'. This applied in 11 cases (Bradford BRP, Docklands, Dyfed, Inference Training, Leeds SRI, 'Pause, Prompt and Praise', St Lawrence School, Shropshire RAISE, Somerset (2 and 4), THRASS94);
- where an evaluation report stated an effect size it was accepted, however calculated. This applied in three cases (Integrated Learning Systems (for two schools out of seven), Paired Reading, Reading Recovery);
- 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. In more detail:

⇒ for Parental Involvement the *post*-test s.d. was used because that was the only control group s.d. given

⇒ for Somerset (4) the s.d. of the *gain* was used because that was the only control group s.d. given (and despite the reservation about this method of calculation mentioned above)

⇒ in 4 cases the control group *pre*-test s.d. was used:

- in 1 of these cases (Somerset 1) the pre-test s.d. was the only control group s.d. given
- in the other 3 cases more than one control group s.d. was given (Bucks PAT: pre-test and gain; Catch Up: pre- and post-test and gain; Cumbria: pre- and post-test); here the pre-test s.d. was used on statistical advice.

There were two 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 treatment group was not significantly greater than that made by the control group. Where this was the case (Integrated Learning Systems [overall], Reading Recovery [four measures out of six, three of them for the alternative treatment group]), it has been stated and the information has been taken as reliable.

As can be seen by comparing the list just given with that given above for RGs, seven reports stated, or allowed the calculation and/or deduction of, both RGs and effect sizes: Bucks PAT, Catch Up Cumbria, Integrated Learning Systems, Paired Reading, Reading Recovery, Somerset (1).

Three reports (BSA Family Literacy, Jersey Computer, Lewisham 2000) have not yet been covered. BSA Family Literacy had no (explicit) control group, and the control groups in Jersey Computer and Lewisham 2000 were inadequate for providing the data to calculate effect sizes as defined above.

Fortunately, all three studies 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)

This method of calculating effect sizes was used for the 3 studies just mentioned (BSA Family Literacy, Jersey Computer, Lewisham 2000).

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, *only two* reports (BSA Family Literacy, Docklands) 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. This was the case in six studies (Bradford BRP, Dyfed, Leeds SRI, Saint Lawrence School, Shropshire RAISE, THRASS94).

Statistical significances of the differences between gains were given (or implied) in nine cases (Bucks PAT, Cumbria, Inference Training, Integrated Learning Systems, Paired Reading, Parental Involvement, 'Pause, Prompt and Praise', Reading Recovery, Somerset), but not in Catch Up, Jersey Computer or Lewisham 2000.

A.2 Data on evaluated schemes, in alphabetical order

Abbreviations:

AT	=	alternative treatment
BASWRT	=	British Ability Scales Word Reading Test
c.a.	=	chronological age
conts	=	controls
exps	=	experimentals
N	=	sample size
n/a	=	not applicable
ns	=	non-significant
r.a.	=	reading age
s.d.	=	standard deviation
stand.	=	standardised
RG	=	ratio gain

Key to school years:

Label of school year			Age of pupils (in years)
in England and Wales	in Scotland and Northern Ireland	in North America	
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

1 Basic Skills Agency's Family Literacy Demonstration Programmes

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

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

Age-range: pre-school 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

N of control group: (no control group)

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

Length of intervention in weeks: 12

Reading test: Reading Recognition subtest of Peabody Individual Achievement Tests

Pre- and post-test average standardised scores and s.d's:

	average stand. score	**(s.d.)**
pre-test	84.1	(17.0)
post-test	88.5	(17.9)

Gain (standardised score points): 4.4

Ratio gain: n/a

Effect size calculated using s.d. of standardisation sample: 0.29

Statistical significance: $p<0.05$

2 Bradford Better Reading Partnership

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 treatment group: (no alternative treatment group)

Length of intervention in weeks: 10

Reading test: Suffolk

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

3 Buckinghamshire 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 treatment group: (no alternative treatment 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.:

	pre-test		post-test	gain		RG	effect size
	average r.a. (yrs & months)	(s.d.)	average r.a. (yrs & months)	ave. (months)	(s.d.)		
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, $t=1.73$, $p<0.05$. Controls made slightly less than standard progress; therefore, though experimentals made only slightly more than standard progress, this could be seen as a satisfactory result for these SEN children.

4 Catch Up

Main reference: Clipson-Boyles (1998)

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 treatment group: 14

Nature of alternative treatment: 'teachers were asked to spend time equivalent to Catch Up 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 treatment 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		post-test		gain		RG	effect size
	ave.	(s.d.)	ave.	(s.d.)	ave.	(s.d.)		
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

5 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 100) 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 treatment (AT) groups: (AT1) 31; (AT2) 30

Nature of alternative treatments: (AT1) reading programme only (similar to Reading Recovery);
(AT2) 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)

Reading tests: BASWRT form A, Neale revised form 1

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

		pre-test		post-test		gain	RG	effect size
test	group	r.a. (yrs & months)	(s.d.)	r.a. (yrs & months)	(s.d.)	(months)		
BASWRT	exps	5:10	(0:6)	6:9	(0:10)	11	1.5	0.45
	conts	6:0	(0:6)	6:7	(0:8)	7	1.1	
	AT1	5:11	(0:6)	6:7	(0:5)	8	1.1	0.04
	AT2	5:11	(0:7)	6:7	(0:8)	8	1.1	0.02
Neale, accuracy	exps	5:1	(0:3)	6:2	(1:0)	13	1.8	1.60
	conts	5:1	(0:4)	5:6	(0:10)	5	0.9	
	AT1	5:0	(0:2)	5:9	(0:6)	9	1.3	0.63
	AT2	5:2	(0:5)	5:10	(0:11)	8	1.1	0.27
Neale, compre.	exps	5:2	(0:4)	6:5	(0:11)	13	1.9	1.29
	conts	5:5	(0:6)	5:11	(0:9)	6	0.8	
	AT1	5:4	(0:4)	6:0	(1:0)	8	1.2	0.43
	AT2	5:5	(0:6)	5:11	(0:10)	6	0.9	0.02

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

6 Docklands Learning Acceleration Project

Main reference: Scott *et al.* (1997)

Date: 1995-96

Age-range: Y3-Y4 (average age at start 7:6, at end 8:6)

Type of children: Mixed-ability

N of experimental group: 379 in 15 schools, including 1 Special school

N of control group: (no control group)

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

Length of intervention in weeks: 52

Reading test: Suffolk

Pre- and post-test average scores (r.a's in years and months) (s.d's not given):

pre: 6:8; post: 7:9

Gain (months of r.a.): 13

Ratio gain: 1.1

Effect size: was not given and could not be calculated

Statistical significances: gain was ns when age allowed for, but authors argued that standing still was a good result for these children, on grounds that without the intervention they would have fallen even further behind

7 Dyfed Improving Reading Standards in Primary Schools Project

Main reference: Dyfed County Council (1994)

Date: 1993-94

Age-range: Y1

Type of children: Low attainment (screened on Bury Infant Check and Dyfed Early Reading Check, and failed to score on Suffolk at pre-test)

Ns of experimental groups: (1, thinking skills plus extra reading) 21
(10 English-speaking, 11 Welsh-speaking);
(2, extra reading only) 30
(14 English-speaking, 16 Welsh-speaking)

N of control group: 22 (8 English-speaking, 14 Welsh-speaking);
2 further mixed-ability control groups totalling 128 mentioned, but no data given

Equivalence of groups: Children most in need chosen as experimentals; controls were other children in same schools, but not matched

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

Length of intervention in weeks: 38

Reading tests: BASWRT, Prawf Darllen Clwyd

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

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

	English		Welsh	
	gain	**RG**	**gain**	**RG**
experimentals 1	17	1.9	14	1.6
experimentals 2	13	1.4	14	1.6
controls	8	0.9	15	1.7

Effect sizes: were not given and could not be calculated

Statistical significances: All differences in gain between experimentals and controls were said to be ns – but this is ignored here because the control groups were not properly equivalent to the experimentals, and the statistical significance of the differences is treated as unknown.

8 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
} all in 5 schools

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

Ns of alternative treatment (AT) groups: (AT1) 14; (AT2) 12, all in same 5 schools

Nature of alternative treatments: (AT1) comprehension exercises
(AT2) rapid decoding practice

Equivalence of groups: Groups matched on age, reading accuracy and vcoabulary; 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		post-test	gain	RG
		ave. r.a. (yrs & months)	(s.d.)	ave. r.a. (months)		
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
	AT 1	8:5	(1:2)	8:10	5.0	5.0
	AT 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
	AT 1	8:1	(1:0)	8:11	9.6	9.6
	AT 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-treatment 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 (AT2) for less skilled comprehenders; BUT Inference Training NOT more effective than comprehension exercises (AT1)

9 Integrated Learning Systems, National Council for Educational Technology study

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) – but not clear if this includes controls – only 375 experimentals traceable in details of report

N of control group: not stated

Equivalence of groups: not stated

N of alternative treatment group: (no alternative treatment 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 gain: 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

10 Jersey Computer Assisted Reading Development Programme

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

N of control group: a control group of 9 pupils is mentioned but is statistically inadequate and therefore ignored here

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

Length of intervention in weeks: 4

Reading test: BASWRT

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

pre-test	post-test	gain
92.4	100.7	8.3

Ratio gain: n/a

Effect size calculated using s.d. of standardisation sample: 0.55.

Statistical significance: not stated

11 Leeds Sustained Reading Intervention

Main reference: Birch (1995)

Date: April 1992-March 1995

Age-range: Y1 (also Y7, not analysed here)

Type of children: Low attainment ('after 1 year in school identified as not making progress equivalent to peer group')

N of experimental group:		
	1992/93	122 in 22 schools
	1993/94	167 in 32 schools (10 new)
	1994/95	168 in same 32 schools
	Total	457 in 32 schools

N of control group: (no control group)

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

Length of intervention in weeks: 26

Reading tests: Carver Word Recognition Test, Salford

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

Gain (months of r.a.): 24, for all groups

Ratio gain: 4.0

Effect sizes: were not given and could not be calculated

Statistical significances: not stated

12 Lewisham Literacy 2000

Main reference: Wray and Rogers (1997)

Date: Autumn 1995-Summer 1996 (scheduled to run till 2000)

Age-range: Nursery, Y1, Y3 and Y7; only Y3 considered here (Y1 ignored because only raw scores given, without s.d's or any control group data)

Type of children: SEN (all at Stage 1 or 2 of Code)

N of experimental group: 57 in 11 schools

N of control group: 250 (rest of year group) in same schools

Equivalence of groups: Not matched – controls were simply the rest of the year group

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

Length of intervention in weeks: 30

Reading test: NFER-Nelson Group Reading Test 6-12

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

	pre-test	post-test	gain
experimentals	88.3	92.0	3.7
controls	96.2	98.4	2.2

Ratio gain: n/a

Effect size for experimental group calculated using s.d. of standardisation sample: 0.25.

Statistical significances: not stated

13 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 (Yl-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 treatment group: (some projects had alternative treatment groups, but too numerous and disparate to report here)

Nature of alternative treatments: (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

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

14 Parental Involvement in Haringey

Main reference: Tizard, Schofield and Hewison (1982)

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 treatment (AT) group: 45 in 2 schools

Nature of alternative treatment: extra teacher help with reading

N of control group for alternative treatment: 66 in same schools

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

Length of intervention in weeks: 52

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)
AT	99.3	(16.6)
controls for AT	98.1	(13.7)

Gain: not stated

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
AT group	0.09

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

15 Pause, Prompt and Praise

N.B. The original report treated the group tutored by untrained 13-year-olds as a control group. However, this was clearly not a 'no treatment' control, and is more appropriately classified (as here) as an alternative treatment group.

Main reference: Wheldall and Colmar (1990), p.130

Date: between 1985 and 1990

Age-range: Y4 (age 9 to 10)

Type of children: Low attainment (r.a. on average 21 months below c.a.)

N of experimental group: 10 (with trained 13-year-old tutors)

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

N of alternative treatment group: 10 (with untrained 13-year-old tutors)

Nature of alternative treatment: sessions with untrained 13-year-old tutors

Equivalence of groups: Pupils were assigned at random to experimental and alternative treatment groups

Length of intervention in weeks: 9

Reading test: Neale

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

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

	accuracy		**comprehension**	
	gain	**RG**	**gain**	**RG**
experimentals	7	3.5	12	6.0
AT group	7	3.5	9	4.5

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

Statistical significances: not stated, but it is implied that both experimental vs alternative treatment differences were ns

16 Reading Recovery in London and Surrey

Main references: Sylva and Hurry (1995), 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 treatment schools

N of alternative treatment (AT) group: 91 in 23 schools in same LEAs

Nature of alternative treatment: Phonological Intervention

Ns of control groups for alternative treatment: (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 treatment 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))

Reading tests: BASWRT, Neale

Pre- and post-test r.a's on BASWRT (s.d's not given), gains (months of r.a.) and ratio gains (N.B. parallel data for Neale not given):

	pre-test (years & months)	post-test (years & months)	gain (months)	RG
experimentals	5:0	6:4	16	1.9
main controls	5:0	5:8	8	0.9
AT	5:1	5:11	10	1.2
controls for AT	5:1	5:9	8	0.9

Effect sizes for experimentals vs main controls, as stated in report (Sylva and Hurry, 1995, pp.XXIV-XXVI) – method of calculation not fully explained:

BASWRT	0.75
Neale accuracy	0.82
Neale comprehension	not given, but said to be ns

Effect sizes for AT group vs their controls: not given, but statistical findings (see below) imply these effect sizes would be close to zero and ns on all 3 measures

Statistical significances: experimentals made significantly greater progress than main controls on BASWRT and Neale accuracy (but not Neale comprehension); alternative treatment group and their controls did not differ in progress on these tests

17 Saint Lawrence School, Towcester, Northants

Main reference: Irwin (1997)

Date: September 1996-June 1997

Age-range: Y3

Type of children: Mixed-ability

N of experimental group: 32 in 1 school

N of control group: (no control group)

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

Length of intervention in weeks: 39

Reading test: Salford, used at pre-test (September 1996) and 2 post-tests, one during intervention (January 1997), other at end (June 1997); data for first post-test ignored here.

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

pre-test (years and months)	post-test 2 (years and months)	gain (months)	RG
7:1	8:6	17	1.9

Effect size: was not given and could not be calculated

Statistical significance: not stated

18 Shropshire: Raising Attainment In Shropshire Education

Main reference: Jones (1996)

Date: October 1994-March 1995

Age-range: mainly Y3 with some Y2 & Y4 (Ns for different years not stated; but since numbers in Y2 & Y4 appear to have been small, only Y3 considered here)

Type of children: Low attainment (all experimentals had r.a. below c.a., but did not have severe learning difficulties or were statemented)

N of experimental group: 67 in 6 schools (across the 3 year groups)

N of control group: 50 in 3 schools (across the 3 year groups)

Equivalence of groups: Not matched – controls seem to have been pupils who were available

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

Length of intervention in weeks: 20

Reading test: Salford, versions A, B and C, used in turn at pre-test (Oct 94) and 2 post-tests, one during intervention (Dec 94), other at end (Mar 95); data from first post-test ignored here

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

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

	gain	RG
experimentals	6.1	1.2
controls	1.6	0.3

Effect size: was not given and could not be calculated.

Statistical significance: not stated

19 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 treatment group (AT1), 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 treatments 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 8:11)

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 treatment (AT) groups: (AT1) 12 in 1 school
(AT2) 12 in 1 school – but 11 at post-test

Nature of alternative treatments: (AT1) self-esteem counselling alone (20 minutes' individual interview per week, with same psychologist as experimentals)
(AT2) 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		post-test	gain	RG	effect
	average	(s.d.)	average	(months)		size
exps	6:11	(1:2)	7:9	10	1.7	0.33
conts	6:6	(1:3)	6:11	5	0.8	
AT1	6:10	(1:1)	7:11	13	2.2	0.53
AT2	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		AT1		AT2	
	U	p	U	p	U	p
exps	?	?	50	ns	34	ns
conts			12	0.001	?	?
AT1					30	0.01

No reason given for not stating control vs experimental and control vs AT2 (remedial phonics only) values; professional counselling plus remedial phonics was no better than counselling only (AT1) or remedial phonics only (AT2); but professional counselling only was better than remedial phonics only or no treatment, and equal to professional counselling plus remedial phonics.

SOMERSET (2)

Main reference: Lawrence (1973, pp.56-65)—some details of treatments 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-treatment control group)

N of alternative treatment 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		**post-test**	**gain**	**RG**
	average	**(s.d.)**	**average**		
	(yrs & months)			**(months)**	
experimentals	7:0	(1:9)	8:0	12	2.0
AT	6:10	(1:11)	7:8	10	1.7

Effect size: was not given and could not be calculated because there was no no-treatment 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 treatments derived from Lawrence (1972)

Date: not stated (1971?)

Age-range: Y3-4 (average ages of the 8 pupil groups at beginning ranged from 7:10-9:7)

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-treatment control group)

N of alternative treatment 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-test s.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		**post-test**	**gain**	**RG**
	average	**(s.d.)**	**average**		
	(yrs & months)			**(months)**	
experimentals	6:11	(1:10)	7:11	12	3.0
AT	6:10	(2:0)	7:1	3	0.8

Effect size: was not given and could not be calculated because there was no no-treatment 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 AT2, 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 treatment (AT) groups: (AT1) 79; (AT2) 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 treatments: (AT1) DISTAR as AT2, 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
(AT2) 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 size**
	average	**(s.d.)**	
experimentals	14.3	(6.4)	0.92
controls	8.8	(6.0)	
AT1	11.8	(6.6)	0.50
AT2	10.7	(5.1)	0.32

Ratio gain: n/a

Statistical significances: experimentals and AT1 made significantly greater gains than other two groups. 'Therapeutic' conditions (experimentals = counselling plus DISTAR; AT1 = drama plus DISTAR) did not differ, and were better than DISTAR only (AT2) and no treatment (controls), which also did not differ.

20 The Handwriting, Reading and Spelling Sequence (THRASS Writetrack 1994)

Main reference: Johnson (1995)

Date: October 1994-April 1995

Age-range: Y2-4 (also Y6, not analysed here)

Type of children: Low attainment (r.a. average 18-24 months below c.a.)

N of experimental group: Y2 11 in 2 schools
Y3 31 in 6 schools
Y4 17 in 3 schools
Total: 59 in 11 schools

N of control group: (no control group)

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

Length of intervention in weeks: 22

Reading test: Macmillan Individual Reading Analysis

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

	average r.a's			
	pre-test	**post-test**	**gain**	**RG**
Y2 accuracy	5:0	5:8	8	1.6
comprehension	5:1	6:5	16	3.2
Y3 accuracy	5:3	6:0	9	1.8
comprehension	5:5	6:5	12	2.4
Y4 accuracy	5:8	6:8	12	2.4
comprehension	6:1	6:10	9	1.8

Effect sizes: were not given and could not be calculated

Statistical significances: not stated

A.3 Comparisons between schemes

To provide a basis for comparing the interventions, including alterative treatment and control groups, the two forms of impact measure (RGs and effect sizes) have been put into rank orders – see Tables A.3-4 below. Where measures for both accuracy and comprehension were available, both have been listed.

In Table A.3,

- 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, a dash is shown where the score would be, and the 'significance of the difference' column is left blank.

In seven of the evaluations studied here (Catch Up, Cumbria, Inference Training, Parental Involvement, 'Pause, Prompt and Praise', Reading Recovery, Somerset) different interventions were compared within one study. In the case of Cumbria 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 Cumbria, 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 alterative treatment groups clearly in Tables A3-4, and this information is therefore provided in Table A.5. In the case of Inference Training, the differences include those between the two experimental groups.

Table A.3: List of studies in decreasing order of ratio gain for accuracy

	R G, accuracy			RG, comprehension		
	exps		conts	exps		conts
Inference Training, AT1 (comprehension exercises)	5.0		–	9.6		–
Inference Training, exps 1 (less skilled comprehenders)	4.3		–	17.4		–
Leeds SR	4.0		–			
Inference Training, exps 2 (skilled comprehenders)	3.9		–	5.9		–
Pause, Prompt & Praise, exps	3.5		–	6.0		–
Pause, Prompt & Praise, AT (untrained tutors)	3.5		–	4.5		–
Catch Up (experimentals in matched schools)	3.4	?	0.4			
Paired Reading, exps in control – group designs	3.4	*	2.0	4.6	*	2.5
Paired Reading, all exps	3.3	*	–	4.3	*	–
Inference Training, AT2 (rapid decoding)	3.0		–	8.2		–
Somerset (3), exps (counselling plus remedial)	3.0		–			
Catch Up (all experimentals)	2.6	?	(0.4)			
Bradford BRP, middle schs	2.5	?	(0.8)			
Bradford BRP, first schs	2.4	?	(0.8)			
THRASS94, Y4	2.4		–	1.8		–
Somerset (1), AT1 (counselling only)	2.2	*	0.8			
Somerset (2) exps (counselling plus remedial)	2.0		–			
Dyfed, English experimentals 1 (thinking & reading)	1.9	?	(0.9)			
Reading Recovery, exps	1.9	*	0.9			
St Lawrence School	1.9		–			
Cumbria, exps (reading & phonology), Neale	1.8	*	0.9	1.9	*	0.8
THRASS94, Y3	1.8		–	2.4		–
Somerset (1), exps (counselling plus phonics)	1.7	?	0.8			
Somerset (2), AT (remedial only)	1.7		–			
Dyfed, Welsh exps 1 (thinking & reading)	1.6	?	(1.7)			
Dyfed, Welsh exps 2 (reading only)	1.6	?	(1.7)			
THRASS94, Y2	1.6		–	3.2		–
Cumbria, exps (reading & phonology), BASWRT	1.5	*	1.1			
Somerset (1), AT2 (phonics only)	1.5	?	0.8			
Catch Up (matched time)	1.4	?	0.4			
Dyfed English exps 2 (reading only)	1.4	?	(0.9)			
Integrated Learning Systems, School A	1.4	*	0.5			

Table A.3: List of studies in decreasing order of ratio gain for accuracy *(continued)*

	RG, accuracy			RG, comprehension		
	exps		conts	exps		conts
Cumbria, AT1 (reading only), Neale	1.3	ns	0.9	1.2	ns	0.8
RR, AT (Phonological Intervention)	1.2	ns	0.9			
Shropshire RAISE	1.2	?	(0.3)			
Bucks PAT	1.1	*	0.9			
Cumbria, AT1 (reading only), BASWRT	1.1	ns	1.1			
Cumbria, AT2 (phonology only), BASWRT	1.1	ns	1.1			
Cumbria AT2, (phonology only), Neale	1.1	ns	0.9	0.9	ns	0.8
Docklands	1.1		–			
Somerset (3), AT (remedial only)	0.8		–			

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
– = no control group

Table A.4: List of studies in decreasing order of effect size for accuracy

	effect size	
	accuracy	**comprehension**
Cumbria, experimentals (reading & phonology), Neale	1.60	1.29
Somerset (4), experimentals (counselling plus DISTAR)	0.92	
Paired Reading	0.87	0.77
Parental Involvement, experimentals	0.84	
Reading Recovery, experimentals, Neale	0.82	ns
Catch Up, experimentals in matched schools	0.78	
Reading Recovery, BASWRT	0.75	
Cumbria, AT1 (reading only), Neale	0.63	0.43
Jersey Computer	0.55	
Integrated Learning Systems, School A	0.55	
Somerset (1), AT1 (counselling only)	0.53	
Somerset (4), AT1 (drama plus DISTAR)	0.50	
Cumbria, experimentals (reading & phonology), BASWRT	0.45	
Somerset (1), experimentals (counselling plus phonics)	0.33	
Somerset (4), AT2 (DISTAR only)	0.32	
BSA Family Literacy	0.29	
Somerset (1), AT2 (phonics only)	0.27	
Cumbria, AT2 (phonology only), Neale	0.27	0.02
Catch Up, AT (matched time)	0.25	
Lewisham 2000	0.25	
Bucks PAT	0.16	
Parental Involvement, AT (extra teaching)	0.09	
Cumbria, AT1 (reading only), BASWRT	0.04	
Cumbria, AT2 (phonology only), BASWRT	0.02	
RR, AT (Phonological Intervention), Neale	ns	ns
RR, AT (Phon. Intervention), BASWRT	ns	
Integrated Learning Systems, overall	ns	
Integrated Learning Systems, School U	-0.4	

Key: ns = figure was not given but was stated or implied to be close to zero and statistically non-significant

Table A.5: Statistical comparisons between experimental and alternative treatment (AT) groups

Cumbria:	The experimental treatment (Reading with Phonology) was significantly better than both ATs (reading-only, phonology-only) on all three measures
Inference Training:	– On accuracy, all differences in gains among the two experimental and two AT 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 (AT2) for less skilled comprehenders; BUT comprehension exercises (AT1) were just as effective as Inference Training
Parental Involvement:	The experimental and AT groups could not be compared at post-test because they differed significantly at pre-test
Pause, Prompt and Praise:	Not stated, but it was implied that both experimental vs alternative treatment differences were non-significant; that is, children with untrained tutors made equivalent gains to those with trained tutors
Reading Recovery:	No information was given on statistical significance of differences between experimental (Reading Recovery) and AT (Phonological Intervention) groups
Somerset (1):	Professional counselling plus remedial phonics was no better than counselling only (AT1) or remedial phonics only (AT2); but professional counselling only was better than remedial phonics only or no treatment, 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; AT1 = drama plus DISTAR) made significantly greater gains than other two groups (AT2 = DISTAR only; controls = no treatment). 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, because 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 below 1.4, and effect sizes below 0.25, represent 'holding one's own' – 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 treatment. Thus schemes (or conditions within schemes) with impact measures of this order did not seem to produce any impact over and above normal schooling, 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 (for this argument, see especially Bucks PAT and Docklands). 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. Within Cumbria, the reading-only condition (AT1) had low RGs for accuracy and comprehension, and the phonology-only condition (AT2) 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 AT produced greater gains than the control condition, that finding and the RGs are taken here to be the more accurate.

There is also a discrepancy within the effect size list between the Neale accuracy result (high) and Neale comprehension result (low) for Reading Recovery experimentals. Since the BASWRT effect size for the same children was high, it is assumed here that Reading Recovery was shown to be effective.

The RG list contains few values below 1.0 ('normal progress'), and all but one of those RGs arose from control groups. (The exception was Somerset (3), the remedial-only AT group.) This finding is, however, circular: children receiving normal schooling

mostly made the progress to be expected of children receiving normal schooling. What is more interesting is that some control groups had RGs *above* 1.4, and were therefore making better than expected progress despite, apparently, receiving no extra treatment. These RGs came from just two studies: Paired Reading, and the Welsh-medium schools in Dyfed. What might these schemes have had in common? Perhaps they affected a high proportion of the schools in the areas in which they 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).

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? Thirteen of the studies analysed provided no follow-up data (in a few cases, especially Catch Up, this was because they were too recent), but seven did provide information on re-tests of participating children at some point after the end of the intervention:

BSA Family Literacy

Varying numbers of children were re-tested 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. 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 the gains made.

Bradford BRP

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.

Cumbria

All children were re-tested in January 1991, 9 months after intervention had ceased and after post-test, using Neale (but not BASWRT). No further relative gains were made, but, as at post-test, on both accuracy and comprehension, experimentals' scores were significantly better than other 3 groups', and those groups did not differ significantly.

Jersey Computer

All children were re-tested at two points, 10 weeks and 6 months after the end of the intervention. The experimental children continued to make relative gains over the implicit control group provided by the standardisation sample.

Leeds SRI

First cohort (1992-93) were re-tested one year after end of intervention; 65% had r.a. above c.a., 40% had r.a. more than 6 months above c.a. (both figures zero at outset).

Paired Reading

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.

Reading Recovery

A follow-up one year after the end of the intervention (May-July 1994) was reported in Sylva and Hurry (1995), and a further follow-up after three years (Summer 1996) in Hurry and Sylva (1998). Both used essentially the same samples as the original evaluation. 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 significantly better on Neale comprehension, which they had not been at the end of the intervention. And by this point the Phonological Intervention group were significantly better than their controls on the BASWRT and Neale accuracy, which they had not been at the end of the intervention (but still not significantly better on Neale comprehension). The differences in r.a. on the BASWRT were 6 months for the Reading Recovery children, 3 months for the Phonological Intervention group.

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.'

Conclusion on follow-up studies

Only one of the follow-up studies showed 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 receving free meals and those who had been non-readers at six continued to benefit. All seven studies showed that most of the children at least held their own, and in some studies they continued to make relative gains.

nfer

WHAT WORKS FOR SLOW READERS?

The Effectiveness of Early Intervention Schemes

As many children as possible should learn to read satisfactorily first time, but what of those who don't? How are they to be helped? This book is about intervention schemes that have been devised to address this need, and is intended to inform schools' choices among such schemes.

The onus on schools to make such choices has been increased by the announcement of the National Literacy Strategy which is to be introduced in September 1998.

More exactly, the questions this book addresses are:

- Which approaches have been used in Britain in an attempt to boost the reading attainment of lower-achieving but non-dyslexic pupils in Years 1-4?
- Which of those approaches have been quantitatively evaluated?
- What are those schemes like, and how effective are they?

The book analyses 20 studies which provide evidence on the effectiveness of about 30 approaches, ranging from parental involvement to computer-assisted methods, and taking in Catch Up, DISTAR, Family Literacy, Paired Reading, Reading Recovery and THRASS.

ISBN 0 7005 1480 5

£8.00