



Achieving the Vision

The Final Research Report of the West Dunbartonshire Literacy Initiative

Tommy MacKay

November 2007

Achieving the Vision

The Final Research Report of the West Dunbartonshire Literacy Initiative

Author:

Dr Tommy MacKay
Psychology Consultancy Services
Ardoch House
Cardross
Dunbartonshire G82 5EW

Consultant to the West Dunbartonshire Literacy Initiative

Publisher:

West Dunbartonshire Council
Educational Services
Garshake Road
Dumbarton
G83 3PU
Telephone 01389 761654

ISBN: 978-0-9537736-6-4

Foreword

‘Achieving the Vision’ is a bold title for a research report. However, it is chosen to celebrate the achievement of something that has never been done in the world before: the eradication of illiteracy from an entire education authority.

The West Dunbartonshire Literacy Initiative is unique. Its achievements have been recognised nationally and internationally at many levels, including the Centre for Policy Studies (Burkard, 2006) and the Prime Minister (Brown, 2007).

This Final Research Report provides an overview of the entire 10-year study. It complements the earlier Phase 1 Research Report (MacKay, 2006), which provided a full description of all of the studies conducted and presented the detailed results for the first six years of the initiative. The Phase 1 report was very much more extensive than this Final Report as it was written to serve a different purpose. It continues to be the authoritative source for the full coverage of the design, implementation and evaluation of the study. It sets out the entire methodology in detail, together with the assessment measures used, and is supported by an extensive literature review. The Final Report presents the whole project in summary form, together with the final results.

Achieving the vision of raising attainments and eradicating illiteracy across the whole school population is based not only on an educational imperative but also on a total commitment to psychological research. At every stage this commitment has informed the design, implementation and evaluation of the initiative. However, it has been a psychological science that is driven by values of seeking to promote social justice and human well-being. Its objectives and methodology are governed by a belief that:

‘Psychology has the potential to help bring about a significantly better world, in keeping with its ethical mandate to promote human welfare. Yet too often we settle for too little’ (Prilleltensky & Fox, 1997, p. 4).

As a research study the project addressed an ambitious agenda. It aimed to apply psychology on a large scale to endemic social and educational problems with a view to laying a foundation for major, intergenerational change in a disadvantaged population. While therefore it incorporates the formal analysis of statistics relating to many thousands of children and young people, the ultimate aim has been to achieve meaningful and positive change in individual lives. This aim is crystallised in the statement made at one of the dissemination conferences by Kathleen Duncan, a pupil at Braidfield High School, Clydebank:

'When all this started I couldn't read. I was a failure. Now I have a cupboardful of books at home. My favourite authors are Roald Dahl and J.K. Rowling. Now I am a success.'

It is statements like this, echoed many times by young people on the project, that highlight the place of values in science, and of an applied psychology agenda that addresses the needs of the most disadvantaged and vulnerable in society.

The aim of eradicating illiteracy is not only an ambitious one but one that has far-reaching implications. Each year in the UK over 100,000 young people leave school 'functionally illiterate' (Organisation for Economic Co-operation and Development, 2000; The Basic Skills Agency, 2001). In West Dunbartonshire the problem was endemic when the project began in 1997. Ten years later it has effectively been eradicated.

This project was founded on visionary goals. The achievement of these goals in 2007 is not a final resting place. While the research phase has finished, West Dunbartonshire is committed to maintaining and enhancing these achievements, and to a population in which there is zero tolerance of illiteracy. The result will be a more skilled workforce, a stronger economy and a better quality of life for thousands of vulnerable young people.

Abstract

Objectives: The aim of this study was to design, implement and evaluate a multiple-component intervention to address underachievement and illiteracy in West Dunbartonshire, taking full account of educational change processes in the context of real world research.

Method: A main study and four supporting studies were conducted. The main study involved the design and implementation through 10 years of a multiple-component intervention in 58 nurseries and primaries, using a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts of the same age. Children in the early stages (N = 3,000+ annually) were individually assessed on a baseline assessment designed for the study, while older pupils (N = 3,000+ annually) took group tests. The synthetic phonics study used a quasi-experimental design to compare two phonics programmes in 18 schools. The attitudes study was a long-term follow up of 24 children from an earlier randomised control trial. The declaration study designed, implemented and evaluated a novel strategy in 12 nurseries and primaries in another education authority (N = 565), using a quasi-experimental design. It served the purpose of informing aspects of West Dunbartonshire's intervention. The individual support study was a quasi-experimental study in secondary school (N = 24), followed by extension into 35 primaries and then into all secondary schools.

Results: In the main study, comparison of cohorts showed year-on-year gains on all tests and across all age groups, with sustained post-intervention gains in later years. In each of the four supporting studies gains were found for the experimentals, pointing to benefits in the use of synthetic versus traditional phonics, in changing attitudes to reading, in making declarations of future reading achievement and in the use of intensive individual support. The extension of the individual support study, together with the effects of the other interventions, resulted in the effective eradication of illiteracy from school leavers in the authority by summer 2007.

Conclusions: The interventions reported in this study have resulted in raised achievement, have effectively eradicated illiteracy in West Dunbartonshire and have developed a foundation for intergenerational change in attainment levels.

Contents

<i>Foreword</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
Chapter 1 Summary	1
Chapter 2 Introduction and overview	4
Chapter 3 Raising achievement: The final results of the main study	13
Chapter 4 Eradicating illiteracy: The final results of the individual support study	29
Chapter 5 Conclusions	37
Epilogue	49
References	51
Appendices	53
<i>Appendix 1 The Structure of Scottish Education</i>	
<i>Appendix 2 Contents of the Phase 1 Research Report</i>	

The West Dunbartonshire Literacy Initiative

The design, implementation and evaluation of a strategy to raise achievement and to eradicate illiteracy

Chapter 1

Summary

Relation of Final Report to Phase 1 Report

The West Dunbartonshire Literacy Initiative Phase 1 Research Report (MacKay, 2006) provided the full account of the aims, rationale, design, methodology and main results of this 10-year initiative, together with a full literature review covering the three research bases essential to the understanding of the project – the process of educational change, socio-economic disadvantage and interventions in literacy. It also covered the rationale for and design of the baseline assessment. A chapter authored by Kathy Morrison, at that time Project Leader and Head Teacher Early Intervention, described the initiative in action.

All of these aspects of the Phase 1 Report continue to be central to a full understanding of the initiative, and where further detail is required it will be found in that report. The scope of the Final Report is to provide a summary of the Phase 1 Report and to present the final results of the initiative until the end of the 10-year research study ending in 2007. In doing so it draws freely from sections of the earlier report where these are helpful in providing necessary context and detail.

The results shown in the Phase 1 Report covered the main study and the four supporting studies up to the period ending in December 2003. By that stage three of the studies had been completed and two were in process. The completed studies were: the synthetic phonics study, the attitudes study and the declaration study. The studies in process were the main study and the individual support study. Both of these provided a strong predictive base for the results shown in the Final Report.

The main study, which had the principal objective of raising literacy attainment for the whole population at all levels, had transformed standards of achievement. While scores continued to show modest year-on-year improvements it was expected that the final research results, as measured for the 10th successive year in December 2006, would broadly reflect the general outcomes shown in the Phase 1 Report. This proved to be the case. The detailed graphs and statistics provided in the Phase 1 Report continue therefore to be of relevance, but new data are provided in the Final Report in summary form to show the final outcomes.

The individual support study aimed to eradicate illiteracy from the authority's schools. The Phase 1 Report showed the results of the original quasi-experimental study in one secondary school using Toe By Toe. It also showed the results of the subsequent gains score study across the 35 primary schools. This provided a foundation for developing the initiative into all of the secondary schools in the authority and for predicting that the aim of eradicating illiteracy among school leavers by 2007 was on target. This has now been achieved as the results shown in the Final Report demonstrate.

Objectives

Socio-economic disadvantage accounts overwhelmingly for the variance in reading achievement across populations, and illiteracy and underachievement in disadvantaged children are endemic. Short-term interventions with small samples have resulted in increased reading test scores, but comprehensive, sustainable strategies for dealing with large populations are lacking. The aim of this study was to design, to implement and to evaluate the effects of, a multiple-component intervention to address underachievement and illiteracy in areas of socio-economic disadvantage, taking full account of the factors affecting educational change in the context of real world research.

Methods

A main study and four supporting studies were conducted. The main study involved the design and implementation of a multiple-component literacy intervention in all pre-school establishments (N = 23) and primary schools (N = 35) in West Dunbartonshire, Scotland's second most disadvantaged education authority. (The structure of pre-school, primary and secondary education in Scotland is shown in Appendix 1.) The study used a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts of the same age. The sample comprised all children in the pre-school year and in Primary 1 and Primary 2 classes (N = approximately 3,000 per year; range 2,538 in 2006 to 3,659 in 1997).

A baseline assessment (MacKay, 1999a, 2006) covering concepts of print, phonological awareness and early reading skills was designed, and was used as a baseline prior to the start of the intervention, and for evaluation in each year of the intervention. All children were tested individually. This allowed comparisons of cohorts of children at the same level each year, as well as the opportunity to trace the progress of individual pupils. Group reading tests were also conducted with all pupils in Primaries 3, 4 and 7 to allow evaluation of the progress of pupils who had moved on from the stages at which the programme was operating. The multiple-component programme consisted of 10 strands, each drawn from existing research on literacy intervention.

The four subsidiary studies were conducted to assess three developing aspects of the 10 strands of intervention, namely: 'a strong and structured phonics emphasis' (one

subsidiary study); ‘changing attitudes, values and expectations’ (two subsidiary studies); and ‘identification and support for children who are failing’ (one subsidiary study).

The phonics study had for its sample Primary 1 classes in 18 primary schools (9 experimentals, 9 controls). A ‘synthetic phonics’ programme was introduced for the experimentals, while the controls had a traditional ‘analytic phonics’ programme. The first study of attitudes and expectations was a long-term follow up of 24 children who had participated in an earlier intervention. The second involved the development of new strategies that were implemented and evaluated in a second education authority, using a sample of pupils (N = 565) in the pre-school year and Primary 1 classes in 12 establishments (6 experimentals, 6 controls). The identification of and support for children who were failing involved a three-fold approach: first, a quasi-experimental study at secondary school involving an individual support programme (12 experimentals, 12 controls); second, the extension of the individual support programme into classes at the top end of primary school in 35 schools; third, the application of the programme to every pupil beyond primary stage who had still not achieved functional literacy.

Results

In the main study, significant year-on-year gains on all baseline assessment tests and across all age groups were found throughout the 10-year intervention period, when cohorts at each age level were compared with the cohorts at the same level in the previous year. Group tests at Primaries 3 and 4 indicated that these gains were being sustained at a reduced level after the intervention was completed. In the synthetic phonics study, significant gains for experimentals on non-word reading and on word reading were found, with indications of sustained improvements into the middle years of primary schooling. In the first study of attitudes and expectations, children whose reading ages had increased more than five years previously as a result of a randomised control trial based on attitude change were reassessed. The experimentals were still reading at significantly higher levels than the controls. The importance of attitudes and expectations was further supported by the second study. The experimentals showed significant gains on baseline assessment tests and also showed significant shifts towards more positive attitudes to reading. Finally, in the individual support study, the secondary school experimentals showed very significant gains, and this pattern was reflected when the programme was introduced across primary schools. The relatively small remaining number of pupils who entered secondary school with low literacy levels were successfully supported at an individual level.

Conclusions

This study has pointed to the effectiveness of a multiple-component literacy intervention in enhancing reading achievement and addressing illiteracy in a large population of children and young people in an area of socio-economic disadvantage. It has also indicated the potential of the intervention as a foundation for planning intergenerational change in achievement levels.

Chapter 2

Introduction and Overview

Aims and rationale of the study

Building on the foundations laid by a number of preparatory studies, outlined below, the aim of this research was to design, implement and evaluate a multiple-component literacy intervention for raising achievement and addressing illiteracy in a whole population in areas of significant socio-economic disadvantage. The principal focus of the study was the early years of education, from the pre-school year in nursery education to the end of the second year of primary schooling. This covered a sample whose ages ranged from under four years at the beginning of their pre-school year to just under seven years at the end of Primary 2. Additional studies were conducted with other age groups in the later primary years and into secondary school.

Many investigations, including the randomised control trial that preceded this research (MacKay, 1999b), have demonstrated that it is possible to raise children's reading scores in small samples, in single establishments, using single intervention strategies over a short-term period. This research represented the first phases of a long-term study addressing a much more complex challenge – tackling the educational impact of socio-economic disadvantage in a large sample, across many establishments, using multiple-component interventions over a long-term period. The aim, in short, was to lay the foundations for intergenerational change in an entire population.

The anticipation of sustainable change on this scale was supported by a clear rationale with two key elements. The first was to design an intervention strategy using the existing and developing evidence base for enhancing literacy levels. By implementing this strategy primarily as an early intervention programme it was expected not only that the overall literacy levels of children in the early years of schooling would be raised, but also that the number of children experiencing reading failure would be significantly reduced. This planned reduction in the numbers failing as they entered the later primary years would create greater scope and economic feasibility for supporting these pupils with intensive individual help to overcome their difficulties. By following each of these children through with support until they had achieved functional literacy the problem of illiteracy in the school years would be eradicated.

The second element involved recognition of the importance of context as well as content in the design and implementation of large-scale educational interventions. If programmes do not take explicit account of the factors affecting educational change they are unlikely to be sustained effectively in the longer term, irrespective of the quality of their content. This study therefore incorporated a recognition of the process of change. The key context variables were articulated and utilised in the implementation of all aspects of the intervention to ensure that the delivery of the content was sustainable and successful.

Overview of the studies

The research reported here comprises five studies: the main study and four supporting studies, one of these relating to phonics, two to attitudes and expectations and one to individual support of pupils experiencing reading failure. A number of 'preparatory' studies preceded the commencement of this research, and reference to them is of particular relevance as they laid the foundations on which many of its key aspects were developed.

In addition, a baseline assessment scheme (MacKay, 1999a) was designed as part of this research. Its development is outlined in detail in the Phase 1 Research Report, together with an account of other assessment measures used.

The 'preparatory' studies

Three studies or groups of studies were of particular significance in helping to formulate several of the main features of the current research: first, a series of studies of playground behaviour (MacKay & Briggs, 1994; Briggs, MacKay & Miller, 1995); second, a study of attitudes and values among children with reading failure (MacKay, 1995a, 1999b); third, an early intervention literacy study (MacKay & Watson, 1996, 1999).

The playground studies were of importance in highlighting the complexities of the process of change in educational settings and in clarifying a number of key factors for effective interventions in this area. The implications of four such projects were outlined by MacKay and Briggs (1994). One highly successful playground intervention, using a qualitative research design incorporating several quantitative measures, highlighted the significance of the socio-economic context in relation not only to behaviour but also to difficulties in learning (Briggs, MacKay & Miller, 1995). It also emphasised the importance of factors such as self-esteem in effecting positive change. The groupwork intervention paradigm it developed, with its focus on self-esteem, was adopted in the next study, as outlined below.

The study of attitudes and values moved the focus of a groupwork intervention from playground behaviour to reading failure (MacKay, 1995a, 1999b). This was a randomised control trial in which 24 pupils in Primary 4 and Primary 5 with severe levels of reading failure were matched in triads for age, cognitive ability and reading level and randomly allocated to two experimental groups and a control group. One experimental group was the subject of a 10-week intervention based on changing attitudes and values regarding education and the relevance of reading, while the other, in addition, followed a paired reading programme at home. Both experimental groups achieved significant gains in reading scores compared with controls, together with reported improvements in other areas such as behaviour in school. Measures of attitude change of experimentals versus controls were also significant. This study led to a focus on the importance of attitudes, values and expectations in the main study reported here. It also provides the relevant context for Study 3, 'the attitudes study'.

The above study of reading failure was also important as a preparation for an early intervention literacy study (MacKay & Watson, 1996, 1999). The success of the intervention at mid-primary school level using a single strategy (changing attitudes and values) led to examining the issue of tackling reading failure in the same establishment at school entry age using a wider range of strategies. The sample comprised an experimental group of two Primary 1 (P1) classes (N = 46) in one school and controls (N = 44) matched for age and socio-economic status (SES) in two other schools in the same neighbourhood. All participants were assessed using a pre-test procedure designed specifically for the study and including reading readiness, reading achievement and attitudinal factors. The intervention, which extended over a five month period, was multi-dimensional and utilised variables supported by research in the areas of curriculum and teaching methods, attitudinal factors and home support. Experimentals achieved significantly better post-test results than controls, particularly in phonological skills, and a number of other benefits were reported by the school. The project achieved all of its objectives and led to the introduction of strategies to modify the 'literacy environment' of P1 entrants into all of the pre-school provision in the area.

Each of these studies therefore contributed to the planning and design of the current intervention.

Study 1 ('The main study')

The main study comprised an early intervention programme for all pupils in the pre-school year, Primary 1 and Primary 2 in all of the pre-five establishments (N = 23) and primary schools (N = 35) in West Dunbartonshire. This provided a total sample each year of approximately between 3,000 and 4,000 pupils at these stages. After the first year, one third of this number left the early intervention programme annually as they progressed from Primary 2 to Primary 3, to be replaced by a similar number of new cases entering the programme in the pre-school year.

All of these pupils were individually tested at the end of each calendar year (November-December) using the baseline assessment scheme designed for the programme (MacKay, 1999a). Testing took approximately 20 minutes, and was conducted by classroom teachers and by early intervention teachers employed for the project. All testers were trained in the procedure. About 120 testers were required each year, and training was ongoing to ensure that any new staff joining the programme were trained before carrying out the assessments.

The key elements in the baseline assessment were concepts of print, phonological awareness (for example, rhyme detection and production) and early reading skills (for example, knowledge of letter sounds, blending and word reading). The baseline was carried out prior to the start of the intervention programme and the same assessment measure was used for evaluation of progress each year. This provided a measure of literacy skills for each of the three stages from pre-school to Primary 2 and allowed comparisons of cohorts of children at the same level each year, as well as yielding data charting the progress of individual pupils from year to year.

Further measures of reading ability were also obtained for every child in the higher primary stages, so that there would be a means of evaluating change in the population as the pupils on the programme moved through their primary school years. For this purpose the Norman France Reading Tests (France, 1978, 1981) were conducted in May each year for all pupils in Primaries 3, 4 and 7. These group reading tests were administered by class teachers. Although they lacked the sensitivity of the individually administered baseline assessment they nevertheless provided an overall measure of change in the cohorts at each of these stages.

In terms of content, a multiple-component intervention programme comprising '10 strands' was designed. These strands were drawn from the evidence base for literacy interventions, and recognised areas such as the importance of phonological awareness in the early stages, of a strong and structured emphasis on phonics and of attitudes, values and expectations. An early intervention team comprising a project leader (head teacher) and 10 teachers was appointed. This team constituted the key personnel both in supporting pre-school and primary staff in the delivery of the programme throughout all establishments and in providing additional classroom help.

From the start of the intervention the factors affecting educational change were articulated. A steering group was established and key context variables of 'vision, profile, ownership, commitment and declaration' were promoted as being of the highest importance. It was considered essential that the project should be marked by these five key factors: by all involved with it having the vision to believe that extraordinary results could be achieved; by being presented at all times in the highest profile as something of great importance; by everyone from the leader of the Council to the parents and the children themselves identifying with it and owning it as their own project; by everyone giving a long-term commitment to making it work effectively; and by constant and bold declarations that this initiative would have outstanding success.

These concepts were constantly and deliberately mediated at all levels of the education authority – Councillors, educational directorate, quality assurance personnel, head teachers, class teachers, other school staff, the early intervention team – indeed all who had an involvement with the project. Fidelity of implementation was maintained through a programme of constant monitoring, assessment and training. Motivation was also maintained both by according the project the highest profile in major dissemination conferences and by the phased introduction of new developments to prevent the initiative from becoming stale or routine.

Study 2 ('The synthetic phonics study')

This study aimed to develop one of the 10 strands of intervention designed for the main study – 'a strong and structured phonics emphasis'. It adopted a developing area of research and practice in phonics teaching, 'synthetic' phonics (starting with letter sounds and learning how to combine these to make words), as an alternative to the traditional or 'analytic' phonics approach normally employed in the teaching of reading (beginning at whole word level and breaking words down into letter sounds). The synthetic phonics method had shown good potential, but had not been developed

in situations where socio-economic disadvantage was a significant background factor, or where initiatives to enhance the teaching of analytic phonics were also being conducted.

Eighteen primary schools were selected for the study (9 experimentals, 9 controls). This was a quasi-experimental study as random selection of establishments would not have been a practicable possibility. The major curriculum change required by the experimental schools was such that the project could only be expected to work by asking for volunteers. These nine volunteer schools were matched in pairs with the nine schools selected as representing the nearest controls in terms of socio-economic profiles and literacy attainment levels. A synthetic phonics programme was introduced to Primary 1 classes in the nine experimental schools. The initiative was supported by a major training programme for all staff involved in implementation, together with comprehensive support arrangements and regular feedback meetings. A range of quantitative and qualitative measures was used for assessment. This included not only the pre-post baseline assessments for the first year of the study but also follow-up evaluations using the same assessment measure as the sample moved through Primary 2, and group reading tests as they moved through Primaries 3 and 4.

Study 3 ('The attitudes study')

This study served to develop another of the 10 strands of intervention designed for the main study – ‘enhancing attitudes, values and expectations’. Of the 24 children who had taken part in the randomised control trial on reading failure (outlined above in the preparatory studies), 20 were traced to their secondary schools in 1999, five and a half years after the initial study had taken place. At the time of the original study they were around the middle years of their primary schooling, and at follow up they were around the middle years of secondary schooling. All were individually assessed for their level of literacy skills, allowing new experimental v. control comparisons to be made.

Study 4 ('The declaration study')

This study also served to develop the strand of intervention relating to ‘enhancing attitudes, values and expectations’. The study was carried out in 12 primary and nursery schools (six experimental, six control) in East Renfrewshire during Session 1999-2000. The aim of the study was to change children’s expectations regarding their achievement in literacy, and to assess the impact of such change on actual reading scores. A total of 565 pupils participated in the six experimental establishments – 320 at pre-school level and 245 in Primary 1, with 27 teachers plus school management involved in implementation. Schools were matched for SES and included establishments with high and low levels of disadvantage.

Staff were trained in a novel intervention strategy based on changing expectations through declarations by pupils regarding future achievement, and this was implemented daily throughout a period of approximately nine weeks. A systematic sample of 60 children, five from each experimental and control school, was assessed individually before and after the intervention, using the baseline assessment designed

for the main study. The sample comprised each n th child from the register, where n equalled one-fifth of the number in the primary class or nursery pre-school year group. Pre-post measures of attitudes to reading were also obtained. These quantitative measures were supported by qualitative indicators obtained from children who were assessed individually and also from staff in relation to all of the participating pupils in the experimental establishments.

Study 5 ('The individual support study')

The individual support study recognised that two factors relating to the later years of schooling required to be addressed. First, there were many pupils in the upper primary years and in secondary school who were already experiencing reading failure and who were not going to benefit from a literacy intervention focused on the early years. Second, a strategy was needed to identify and support of children who were still failing even after they had been through the early intervention programme.

This study was carried out in three phases. First, a quasi-experimental study was conducted in one West Dunbartonshire secondary school, with 24 pupils referred for learning support because of low reading levels. Of these pupils, 12 controls were assigned to the normal learning support programme while 12 experimentals were enrolled in an intensive individual support programme. The programme selected as meeting the criteria determined by the researchers was 'Toe By Toe' (Cowling & Cowling, 1993). The allocation of cases to the two conditions was not random, as it had to be subject to the normal timetabling and other constraints encountered in a large secondary school. However, the two samples were matched as closely as possible. Experimentals received individual tuition for 20 minutes a day, for a period of approximately three months. Pre-post assessments were conducted at the start and finish of a 12-month period.

The second phase of the study involved the identification of pupils at upper primary level who were experiencing significant reading difficulties. These children were initially identified by staff in the 35 primary schools in the authority, and following individual testing 104 were selected from 32 schools as meeting support criteria. Approximately 120 individual support workers were trained in the use of the programme by the author and the learning support teacher who delivered the secondary school intervention. These were drawn from a wide range of personnel – teachers, classroom assistants and volunteers. Monitoring and support structures were put in place to ensure effective implementation.

The final phase was the extension of the programme into all secondary schools to address the needs of the relatively small number of pupils who were still experiencing significant difficulties in literacy at this stage. This was a key part of the final strategy in ensuring that illiteracy would be eradicated at school leaving age.

Overview of methodology

Table 2-1 provides an overview of the methodology of the studies in terms of design, sample, assessment measures and analysis.

Table 2-1 Overview of methodology

	Design	Sample	Assessment measures	Intervention	Analysis
Study 1: Main study	Long-term; a cross-lagged design in which pre-intervention population cohorts were controls for subsequent same age intervention cohorts	Total N=60,808 of whom: Individual assessments: 30,903 (pre-test 3,659; post-test: 8,167 pre-school; 9,365 P1; 9,712 P2) Group assessments: 29,905 (9,758 P3; 9,876 P4; 10,271 P7) in 58 establishments (23 pre-school, 35 primary)	Individual: Baseline assessment scheme (MacKay, 1999a, designed for study – Appendix 2) Group: Norman France Reading Tests (France, 1978, 1981). All assessments administered annually	A multiple-component literacy intervention with 10 strands; designed for study but embedded within the curriculum as determined by national and local authority guidance	Independent two-sample <i>t</i> tests with effect sizes calculated on the standard deviations of the pre-test cohorts
Study 2: Synthetic phonics	Quasi-experimental, triangulated by a range of qualitative measures	18 primary schools (9 experimental, 9 control). Total N=590 (315 experimental, 275 control) Targeted N=180 (90 experimental, 90 control)	Baseline assessment scheme (MacKay, 1999a); spelling test (designed for study – Phase 1 Report, Appendix 3); qualitative measures (Phase 1 Report, Appendix 4)	A synthetic phonics programme (based on Lloyd, 1992)	Total sample: independent two-sample <i>t</i> tests with effect sizes calculated on the standard deviations of the pre-test cohorts Targeted sample: related two-sample <i>t</i> tests
Study 3: Attitudes	A comparison study: long-term follow-up of randomised control trial, comparing original experimental and control groups after 5½ years	N=19 (11 experimental, 8 control) from sample of 24 in original RCT	Neale Analysis of Reading Ability, Revised British Edition (Neale, 1989)	Not applicable. Original intervention described in MacKay (1995a, 1999b)	Independent two-sample <i>t</i> tests, with effect sizes calculated on the standard deviations of the standard assessment measure
Study 4: Declaration	Quasi-experimental, triangulated by a range of qualitative measures	12 schools (nursery and primary, 6 experimental, 6 control) Total sample: N=565 (320 nursery, 245 primary) Targeted sample: N=60 (30 experimental, 30 control)	Baseline assessment scheme (MacKay, 1999a); attitudes test, designed for study; qualitative measures	All children in experimental schools and nurseries made declarations regarding future literacy achievements (as in Phase 1 Report, Chapter 13)	Baseline assessment: independent two-sample <i>t</i> tests, with effect sizes calculated as for main study Attitude change: chi-square test
Study 5: Individual support	A quasi-experimental study in secondary school, supported by a gains-score study in primary school	Secondary: one school, N=24 (12 experimental, 12 control) Primary: 35 schools, N=104	Secondary: Gapadol Reading Comprehension Test (McLeod & Anderson, 1972) Primary: Neale Analysis of Reading Ability, Revised British Edition (Neale, 1989)	An individual support package for developing basic reading skills (Cowling & Cowling, 1993)	Secondary: Independent two-sample <i>t</i> tests with effect sizes shown; effect sizes were calculated on the standard deviations of the standardised assessment measure Primary: inspection of individual gain scores

Overview of results

The preparatory studies had already provided a basis from which to plan the design of a long-term, population-wide intervention in literacy using multiple-component strategies. They had highlighted not only the extent of underachievement in literacy and difficulties in related areas in disadvantaged populations, but also the importance of variables other than the content of the curriculum, most particularly attitudes, values and expectations. They had also pointed to the value of a broad-based approach to early intervention.

In the main study, a consistent pattern of higher achievement levels was found for all baseline assessment tests across each of the pre-school, Primary 1 and Primary 2 cohorts. These enhanced results were obtained year-on-year throughout a 10-year intervention period when each cohort was compared with the cohort at the same age level in the previous year. Children at all levels of achievement benefited from the intervention. The proportion of pupils obtaining high scores rose significantly, while at the other end of the scale those with very low scores reduced dramatically in numbers. The group reading tests at Primary 3 and Primary 4, although very much less sensitive than the baseline assessments in their scoring, indicated that the gains made as a result of early intervention were being reflected in reading scores in the years following the intervention. The practical effect of these improvements was very apparent. Indeed, the class teachers in Primary 3 noted that the new reading levels of children entering their classes since the programme began were challenging the delivery of the normal P3 curriculum, which was requiring to be re-appraised.

In the synthetic phonics study it was hypothesised that the baseline assessment tests that should show the effects of the programme were those requiring word attack skills, namely, the non-word reading test and the word reading test. It was on these two tests that significant gains for pupils in the nine experimental schools were found, together with overall improvements in reading performance on group tests at Primary 4. Extensive qualitative data were also obtained from the staff in these schools. This provided very strong support for the effectiveness of the programme, since teachers were virtually unanimous in asserting that their pupils were working at higher levels of skill than had ever been known before. This view was expressed so universally that the nine volunteer schools were joined within a year or so by virtually every other primary school, so that the area became, in effect, a 'synthetic phonics' authority.

The attitudes study and the declaration study strengthened the base for ensuring that addressing attitudes, values and expectations should be built into any literacy intervention programme as a variable which could affect outcomes but which was essentially separate from the literacy content of the programme itself. Both studies addressed this variable in its own right, while keeping the content of the reading curriculum constant. The attitudes study indicated that over five years after a brief intervention to raise literacy scores by changing attitudes and values, the experimentals were still reading at a significantly higher level than the controls, even though they had received no differential treatment during these intervening years.

Again, this study was of importance in planning the main long-term intervention with a view to intergenerational change.

The declaration study resulted not only in significantly higher scores on early literacy skills for the experimentals but also in significant shifts towards more positive attitudes and expectations regarding reading. As a study of children in their pre-school year and Primary 1 it was particularly relevant to informing a large-scale early intervention.

In the individual support study dramatic mean gains were made by the experimentals in their reading scores, while the controls progressed only at the expected rate – that is, they made considerably less than one year's gain over the 12-month period between pre-test and post-test. This result supported the pilot work carried out in preparation for the study, during which very high gain scores were reported for pupils receiving the intervention. This provided a firm basis for setting up the training and implementation arrangements by which the study was extended throughout the upper classes in the primary sector. Following the established success of the programme at secondary using quasi-experimental methods, the primary project was carried out without selecting further controls. A high level of gain scores was achieved across the sample during a five-month intervention period. Finally, all pupils in secondary who had still not achieved functional literacy levels were enrolled to the programme. The result was that by June 2007 there were only three pupils left at school leaving stage in secondary schools in the authority with literacy levels which had not met the target of functional literacy.

These results have consistently supported the two main objectives of the whole initiative – significantly raising the achievement levels of this entire population and providing a basis for the eradication of illiteracy.

Chapter 3: Raising Achievement

The Final Results of the Main Study

Relation of Final Report to Phase 1 Report

Chapter 9 of the Phase 1 Report presented a detailed breakdown of all of the results of the main study covering the period 1997 to 2003. By that stage the pattern of raised achievement across the entire population in the pre-school year, Primary 1 and Primary 2 had been established. The detailed analysis in the Phase 1 Report therefore continues to reflect the main results of the study. However, it was necessary to demonstrate that the pattern established by 2003 had continued to be maintained. This chapter serves that purpose by providing a summary showing that the final 2006 results have not only maintained earlier gains but have demonstrated continuing modest year-on-year gains.

Aims and hypotheses

The aims of the main study were:

- to raise the literacy levels of all children in the pre-school year, Primary 1 and Primary 2 in all schools throughout the authority using a multiple-component intervention
- to provide a basis for long-term improvements in literacy levels in the later years of schooling
- to reduce the numbers of children experiencing reading failure as a basis for eradicating illiteracy throughout the authority.

The following hypotheses were proposed:

- 1 that the year groups of children receiving the intervention programme would have higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level
- 2 that these gains would be sustained after the children left the programme at the end of Primary 2
- 3 that both high and low achievers would show gains, but that in particular there would be significant reductions in the numbers of children experiencing reading failure in the early years.

The '10 strands' of intervention

From the research literature on effective interventions in literacy, 10 strands were selected as the basis for the programme (Box 3-1). All of them were of central importance to the intervention. Of these 10 areas, seven were identified as 'key strands' – that is, they were planned and structured in a formal way as the basis on which the programme would operate. These were: phonological awareness and the

alphabet; a strong and structured phonics emphasis; extra classroom help in the early years; raising teacher awareness through focused assessment; increased time spent on key aspects of reading; identification of and support for children who are failing; and home support for encouraging literacy. The other three strands, while also being viewed as crucial to successful intervention, were promoted in a less formal and structured way, and were identified as being ‘supporting strands’. These were: fostering a ‘literacy environment’ in school and community; lessons from research in interactive learning; and changing attitudes, values and expectations.

The rationale for each of these strands is covered in detail in Chapter 8 of the Phase 1 Report.

Box 3-1 The ‘10 strands of intervention’

Strand 1:	Phonological awareness and the alphabet
Strand 2:	A strong and structured phonics emphasis
Strand 3:	Extra classroom help in the early years
Strand 4:	Fostering a ‘literacy environment’ in school and community
Strand 5:	Raising teacher awareness through focused assessment
Strand 6:	Increased time spent on key aspects of reading
Strand 7:	Identification of and support for children who are failing
Strand 8:	Lessons from research in interactive learning
Strand 9:	Home support for encouraging literacy
Strand 10:	Changing attitudes, values and expectations

Method

Design

This was a long-term, multiple-component intervention study, using a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts at the same age levels.

Sample

The sample was every child in the pre-school year in all nurseries (N = 23 establishments) and in Primary 1 and Primary 2 in all schools (N = 35 schools) throughout West Dunbartonshire. This provided an intervention sample of 27,244 from 1997 to the end of 2006, plus a pre-intervention control population of 3,659, making a total of 30,903, all tested individually. The breakdown of the sample is shown in Table 3-1.

Table 3-1 Early intervention sample: individual assessments

	Pre-school	Primary 1	Primary 2	Total
1997 Pre-intervention controls	1083	1307	1269	3659
1998	1177	1185	1260	3622
1999	1039	1160	1140	3339
2000	1021	1090	1152	3263
2001	986	1100	1097	3183
2002	893	1054	1127	3074
2003	798	1001	1047	2846
2004	768	959	1014	2741
2005	750	928	960	2638
2006	735	888	915	2538
Total	9250	10672	10981	30903

To assess the effects of the intervention into the later years of primary school, the pre-intervention populations in Primaries 3, 4 and 7 were assessed each year on group tests, allowing comparison of same-age cohorts when the early intervention sample progressed through to these stages. This provided a further sample of 29,905 from 1998 to 2006, of whom 8,948 were pre-intervention controls and 20,957 had received all or part of the intervention programme in earlier years. The breakdown in Table 3-2 shows which of the sample had been exposed to one, two or three years of the early intervention programme.

Table 3-2 Group reading tests: Primaries 3, 4 and 7

	Primary 3	Primary 4	Primary 7	Total
1998	1078	1026	1022	3126
1999	*1139	1108	1187	3434
2000	**1221	*1247	1267	3735
2001	***1068	**1128	1057	3253
2002	***1130	***1128	1203	3461
2003	***1047	***1115	*1213	3375
2004	***1048	***1034	**1115	3197
2005	***1025	***1046	***1116	3187
2006	***1002	***1044	***1091	3137
Total	9758	9876	10271	29905

Early intervention sample: * one year ** two years *** three years

The cut-off point for reporting outcomes was December 2006 for individual baseline assessments, and May 2006 for group tests. This allowed year-on-year comparisons for the early intervention. It also allowed comparisons in P3, P4 and P7 for children who had been through the entire early intervention programme from their pre-school year onwards.

The total number of tests conducted for the main study from pre-intervention baseline to the cut-off dates for reporting was as follows: 30,903 individual tests and 29,905 group tests were carried out, making a grand total of 60,808 tests. Not all of this vast data set needed to be utilised in testing the hypotheses for the main study. However, it allowed general trends to be checked year on year, calibrating and confirming the data reported for the key years that have been selected for detailed analysis. It also served other purposes as reported throughout the Phase 1 Report.

‘Significance’ and meaning

The main study was the centre piece of this research. All of the other studies – synthetic phonics, attitudes, declaration and individual support – were designed to support it, and to provide strategies for strengthening its implementation. It was therefore of the most crucial importance to the intervention that it should not only have significant results, but also that these results should have meaning beyond statistical significance in and of itself. It was necessary that the results should be seen as making an impact, and as providing a basis for confidence that long-term and meaningful changes could be effected and sustained, and that these would have actual importance in the lives of the population served by the project.

It was for this reason that a decision was made at the start of the intervention to report results not only in terms of statistical significance using probability values, but also in terms of effect sizes. This decision was guided by the over-arching principles governing the entire research project. It was a project based on a commitment to the scientific method of enquiry, but a science that enshrined in its methods and its priorities a commitment to values. It was a piece of research involving a large amount of public funding applied to vulnerable children and young people most of whom lived in areas of significant socio-economic disadvantage. Statistical significance in and of itself might have served the purposes of a researcher, but might have made no real and lasting impact on the lives of those who were participants in the research.

The governing principles that led to the adoption of a values framework dictated that statistical significance must be viewed in terms of wider questions that were primarily social, cultural and political rather than scientific – questions about whether lives were being changed as a result of the intervention; questions about whether children would leave school with the skills needed for a successful career in a knowledge society; questions about whether ‘significant’ results actually meant significant to the participants in the research or only to the researcher.

It was this main study more than any other that brought the matter back to the fundamental issue of the values framework outlined in Chapter 2 of the Phase 1 Report. It was absolutely necessary that an intervention in literacy should meet orthodox scientific criteria in determining the validity of its outcomes, but for the declared purposes of this research these criteria must be compatible with promoting health, caring and compassion, self-determination and participation, human diversity and social justice (Prilleltensky & Nelson, 1997).

Scope of results reported

The overview of data presented here has been selected to illustrate the results in relation to the three hypotheses for the main study. First, to answer the question of whether the intervention was effective, the baseline assessment results for each year group throughout the intervention period were compared with the results achieved by the same year group cohorts prior to intervention. The pre-intervention cohorts therefore served as the controls for the subsequent cohorts. Second, to answer the question of whether the effects lasted, the group tests given each year to children in the later primary years were analysed to show comparisons of cohort scores before and after intervention. Third, to answer the question of whether all groups benefited, separate comparisons were made for the children with the lowest and highest scores. The tests used throughout these analyses were independent two-sample *t* tests, calculated using the data analysis tools on Microsoft Excel Version 8.0.

Effectiveness of the intervention: the baseline assessment results

The first key question was whether the intervention was effective. It addressed the hypothesis that the year groups of children receiving the intervention programme would have higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level. Throughout the 10-year intervention period the baseline assessment results showed a marked and consistent trend, which may be summarised as demonstrating a systematic enhancement of scores on virtually every test for every group and across every year. On the more elementary tests, such as concepts of print and nursery rhymes, especially in the older age groups, overall enhancement levels were largely dictated by the ceiling of the tests. On tests with a high ceiling – in particular the early literacy skills tests in the younger age groups and the word reading test in the older groups – the results continued to show an upward trend. The implications of floor and ceiling effects on a number of the baseline assessment tests have been recognised and are considered in the discussion section below.

Table 3-3 summarises the results for the key tests applicable to each year group. The tests of word reading and other more formal literacy skills were not applicable to the pre-school sample as far as overall comparisons are concerned, since most children were not able to score on these, although every child had the opportunity to attempt them if able to do so. Similarly, the early phonological tests were not useful for overall comparisons for the older groups, as most children by the early primary years had reached the ceiling on items such as concepts of print. The summary shows comparisons with the pre-intervention baseline of 1997 in relation to three years, namely, 1998, the first year of intervention, 2003, the sixth year of intervention as

reported in the Phase 1 Report, and 2006, the 10th and final year of intervention. By the end of the first year all results had shown a significant increase except the alphabet for pre-school, and letter names for Primary 1. By the end of the sixth year all results had risen significantly. By the end of the 10th year the results had not only been maintained but showed further modest gains. For the years 2003 and 2006 all results were significant at the p<0.001 level.

Table 3-3 Summary of results on key baseline assessment tests

Test	Pre-school			Primary 1			Primary 2		
	1998 N = 1177	2003 N = 798	2006 N = 735	1998 N = 1185	2003 N = 1001	2006 N = 888	1998 N = 1260	2003 N = 1047	2006 N = 915
Concepts of print	***	***	***						
Nursery rhymes	***	***	***						
Initial letter sounds	*	***	***						
Rhyme detection	***	***	***	***	***	***			
Rhyme production	***	***	***	***	***	***	***	***	***
The alphabet	ns	***	***	***	***	***	***	***	***
Lower case letter sounds				***	***	***			
Letter names				ns	***	***	***	***	***
Non-word reading test				***	***	***	***	***	***
Word reading test				***	***	***	***	***	***

Significance (compared with 1997 pre-intervention baseline, one-tailed tests):

* p<0.05 ** p<0.01 *** p<0.001

The detailed breakdown of the results from 1997-2003 provided in the Phase 1 Report illustrated the raw scores on all tests in a series of figures, together with *t* values, significance levels and effect sizes. These are tabulated here with the addition

of the final year of the study, 2006. Tables 3-4, 3-5 and 3-6 show raw scores and effect sizes.

Table 3-4 Baseline assessment 1997-2006: raw scores and effect sizes – preschool

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Concepts of print	4.3	5.4	6.0	6.4	0.58	0.92	1.10
Nursery rhymes	9.2	13.5	16.0	16.6	0.89	1.42	1.53
Initial letter sounds	1.7	1.8	2.4	2.7	0.07	0.41	0.59
Rhyme detection	2.4	3.4	4.3	4.8	0.49	0.93	1.18
Rhyme production	0.8	2.4	3.4	3.9	0.93	1.54	1.80
The alphabet	1.3	1.4	2.2	1.9	0.03	1.11	0.74

Table 3-5 Baseline assessment 1997-2006: raw scores and effect sizes – Primary 1

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Rhyme detection	4.3	5.1	5.6	5.7	0.43	0.75	0.77
Rhyme production	2.1	4.3	5.4	5.4	0.89	1.37	1.37
The alphabet	1.7	2.5	3.7	3.9	0.69	1.76	1.95
Lower case letter sounds	13.5	17.7	23.8	24.6	0.55	1.29	1.39
Letter names	4.0	3.9	11.9	15.9	-0.01	1.08	1.62
Non-word reading test	2.3	3.8	11.1	12.1	0.29	1.76	1.98
Word reading test	5.5	6.6	13.6	16.0	0.15	1.20	1.55

Table 3-6 Baseline assessment 1997-2006: raw scores and effect sizes – Primary 2

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Rhyme production	3.3	5.2	5.8	5.8	0.79	1.04	1.04
The alphabet	2.8	3.5	3.9	4.0	0.54	0.89	0.94
Letter names	13.4	17.1	23.1	24.7	0.39	1.01	1.19
Non-word reading test	12.8	14.2	17.0	17.4	0.23	0.65	0.76
Word reading test	24.2	26.5	35.8	40.4	0.16	0.83	1.16

The baseline assessments were normed on the pre-intervention cohort (N = 3,659). As a non-intervention population, this cohort served as the control group for the intervention samples in subsequent years. This provided the standard deviations that served as the basis for the calculation of effect sizes throughout the period of intervention. The effect sizes therefore show the raw score gains expressed as a proportion of the pre-intervention standard deviation for each test. The effect sizes for the three intervention years in the comparison (1998, 2003, 2006) are shown on exactly the same basis, that is, they represent a comparison with the pre-intervention cohort. This was to answer the question, ‘How does the intervention sample in any given year compare with the data available for the same age group without intervention?’

These effect sizes show a strong effect for the intervention. In particular, the comparison from pre-intervention in 1997 to the final year in 2006 shows substantial effect sizes on every test. The average effect size shown across all tests was 0.45 in 1998, 1.11 in 2003 and 1.26 in 2006.

Tables 3-7 and 3-8 provide an overall summary for each year group in relation to the four years for which results are shown. A combined score for ‘phonological awareness’ has been calculated by adding the score for concepts of print to the combined scores for the actual phonological tests – nursery rhymes, initial letter sounds, rhyme detection and rhyme production – while a combined score for ‘early reading skills’ has been obtained by adding the scores for the alphabet, lower case letter sounds, letter names, non-word reading and word reading. Phonological awareness scores are shown for all age groups, but ceiling effects became apparent in P1 and more so in P2, where many pupils, particularly following the intervention, passed all tests of this type because the skills were fully established. The ceiling for all phonological tests combined was a score of 44, and it will be noted that as the intervention progressed the mean score came ever closer to that figure in P1 and P2.

In P1 63% of pupils had a ceiling score of 44 and this rose to 82% in P2. Early reading skills scores are not shown for the pre-school children as these more formal skills are marked by very substantial floor effects at this age.

Table 3-7 Baseline assessment 1997-2006: combined raw scores and effect sizes for phonological awareness

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Pre-school	18.5	26.5	32.1	34.3	0.92	1.57	1.82
Primary 1	29.2	37.2	41.6	42.1	1.00	1.55	1.61
Primary 2	33.9	39.7	42.8	43.4	0.9	1.37	1.48
.							

Table 3-8 Baseline assessment 1997-2006: combined raw scores and effect sizes for early reading skills

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Primary 1	27.0	34.5	64.1	72.5	0.36	1.78	2.18
Primary 2	77.3	86.4	105.5	112.5	0.31	0.97	1.20
.							

It will be noted that in terms of phonological awareness, the average pupil in the pre-school year in 2006 was scoring higher than the average pupil in P2 in 1997. In terms of early reading skills, the average P1 pupil in 2006 was scoring almost at the level of the average P2 pupil in 1997. Given that the baseline tests are conducted in November to December each year, this means that pupils were making very fast progress on entering school or nursery. For example, a P1 pupil who had been in school for only about four months was doing about as well in 2006 as the average pupil in 1997 after one year and four months.

Lasting effects of intervention

The second key question was whether the intervention had lasting effects. It addressed the hypothesis that gains would be sustained after the children left the

programme at the end of Primary 2. For this purpose the data collected from the start of the intervention for children in the later primary stages were utilised. The results of the Norman France group reading tests in Primaries 3, 4 and 7 were analysed.

The 1998 Norman France tests were all conducted on pupils who had not been subject at any time to the intervention. This first year of group testing was therefore taken as a pre-intervention baseline for the cohorts at Primaries 3, 4 and 7. By the time of the testing conducted in 2006, all children had received at some stage the full three years of intervention. The general trend of these results may be summarised as follows. Pupils who had been on the intervention showed a significant increase in group reading test scores compared with the baseline cohorts at the same stages, with the effect being strongest for those who had received intervention most recently – that is, the P3 pupils. The results are shown in Table 3-9.

Table 3-9 Norman France group tests compared with 1998 baseline

	Reading age 1998	N	Reading age 2006	N	Effect size	Significance
Primary 3	7y 0m	1078	7y 8m	1002	0.34	p<0.001 (<i>t</i> = 8.038)
Primary 4	8y 1m	1026	8y 4m	1044	0.14	p<0.001 (<i>t</i> = 3.461)
Primary 7	10y 6m	1022	11y 0m	1091	0.20	p<0.001 (<i>t</i> = 4.743)

While all of the results showed significant pre-post differences, the effect sizes for the Norman France tests were much more modest than for the baseline assessments. However, the limitations of these comparisons are considered in the discussion section below.

Results for children with lowest and highest scores

The third key question was whether the intervention was of benefit to all groups – to those with the lowest scores and to the high achievers. It addressed the hypothesis that both high and low achievers would show gains, but that in particular there would be significant reductions in the numbers of children experiencing reading failure in the early years. Tables 3-10 and 3-11 show the lowest 10% of scores for the four years under consideration – the 1997 baseline, the first year of intervention in 1998, the sixth year of intervention in 2003 and the final year in 2006. The data shown are the combined scores for phonological awareness and for early reading skills. Again, early reading skills are not shown for the pre-school children because of the floor effects at this age when most children are not yet learning the more formal skills.

Table 3-10 Lowest scores: raw score at 10th percentile – phonological awareness

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Pre-school	4.0	8.5	20.9	24.0	0.52	1.94	2.30
Primary 1	14.0	23.1	28.8	38.1	1.13	1.84	2.99
Primary 2	20.9	30.6	36.2	42.0	1.54	2.43	3.35

Table 3-11 Lowest scores: raw score at 10th percentile – early reading skills

Test	Raw scores				Effect sizes		
	1997	1998	2003	2006	1997-1998	1997-2003	1997-2006
Primary 1	5.0	8.92	22.49	34.8	0.19	0.84	1.43
Primary 2	23.13	32.06	62.38	93.3	0.31	1.34	2.40

As there was a particular interest in reducing the numbers of children who were failing to acquire basic literacy skills, some illustrative data are provided in Table 3-12 on the impact of the intervention on some of the key skills it was set up to address at the various stages – concepts of print for pre-school children, letter sounds in P1 and word reading in P2.

Table 3-12 High scores on key skills

Test	Percent			
	1997	1998	2003	2006
Pre-school – concepts of print: score 7+	11.9	28.7	43.2	50.1
Primary 1 – letter sounds: score 20+	28.5	47.7	90.0	92.2
Primary 2 – word reading: score 30+	31.5	37.6	66.9	77.8

The figures shown in Table 3-12 are the percentage of the pre-school year passing all or all but one of the items on concepts of print, the percentage at P1 with a score of 20 or more out of 26 on letter sounds and the percentage at P2 correctly reading at least 30 words.

Discussion

The results of this study indicate that it has achieved its aims. These were to raise the literacy levels of all children in the pre-school year, Primary 1 and Primary 2 in all schools throughout the authority using a multiple-component intervention; to provide a basis for long-term improvements in literacy levels in the later years of schooling; and to reduce the numbers of children experiencing reading failure as a basis for eradicating illiteracy throughout the authority.

Support has been found for each of the three hypotheses. First, the year groups of children receiving the intervention programme had significantly higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level. Second, gains continued after the children left the programme at the end of Primary 2. Third, both high and low achievers showed gains, and in particular there were significant reductions in the numbers of children experiencing reading failure in the early years.

Limitations and cautionary comments

The limitations in the comparisons of test scores for this study, both statistically and in terms of broader aspects of interpretation of results, have been acknowledged and are discussed here in relation to two areas: first, the comparisons made on the baseline assessment tests, and second, the comparisons made on the Norman France tests administered in the later years of primary school.

First, in terms of statistical limitations it is recognised that it has not been possible with this type of cross-lagged design, as it would have been with some other designs, to use regression analysis to adjust for regression towards the mean when there are significant floor effects. The entire baseline cohort for each test formed the controls for the following cohorts of children at the same stage, and where the floor effects were marked they could do little other than go up in subsequent years if they were going to change at all.

However, this limitation may be viewed in the wider context of the clear and significant trend of the results overall. In general the effects resulting from this study were not subtle. They were large enough that the purpose of the statistics was not so much to determine whether an intervention effect could be identified but rather to confirm its extent. The results were *obvious*. They were clearly discernible in the simplest of descriptive tables or charts used to communicate what was happening. Unmistakable increases in scores were plain across virtually every single baseline test at every age and for every year, including all the in-between years not separately shown for the purposes of this analysis.

Further comment may be made about the statistical limitations acknowledged in the analysis. The tests were not only at times positively skewed because of floor effects, making it easier to achieve increases, but also at times negatively skewed because of ceiling effects, making it harder to achieve increases. This is clearly illustrated in the phonological awareness tests for the children at P2 stage, many of whom reached the ceiling of the tests with ease because there were no phonological skills left to master. This therefore worked against showing an intervention effect. Nevertheless, it was felt useful to illustrate these results, to conduct the *t* tests and to show the effect sizes, but to acknowledge skewness for the reasons indicated. Large effect sizes have been found throughout, not only where there are floor effects but also where there are ceiling effects, and on many other tests which were not skewed in either direction.

There is therefore an overall robustness in the consistency of the wide range of results shown. Indeed, it is the robustness inherent in the size and consistency of the gains achieved across all tests and age groups that overtakes the general range of statistical limitations associated with different methods of calculating the basis for effect sizes (Glass, McGaw & Smith, 1981) and gives confidence in the outcomes of the intervention. In keeping with the hypotheses, the comparisons of scores between baseline and the selected intervention years were one-tailed. However, in almost every case, the gains (whether calculated as one-tailed or two-tailed) were far in excess of any level at which their actual significance values would be reported.

Second, there are limitations inherent in the comparisons made using the Norman France group reading tests at Primaries 3, 4 and 7. Group tests are subject to various factors other than the ability being tested that may affect scores. They involve multiple choice approaches where several abilities other than pure reading ability are required, such as good comprehension of verbal instructions provided at group level, independent working and sustained attention. They are also more prone to children copying from someone else close to them. In addition, they have a greater focus on reading comprehension, and indeed on general comprehension. While therefore they are useful in providing a measure of real reading, they are not ideal for sensitively tapping into the abilities most fostered by the early intervention programme. Whereas the baseline assessment test designed for the study was viewed by the schools as being useful to the extent that there was overwhelming support for continuing to use it after the research was completed, it was agreed to discontinue the Norman France group testing as being unhelpful following the last assessment in May 2006.

Broader aspects of interpretation of results also affected the Norman France comparisons. The further children were removed from the intervention itself – for example by P4 and especially by P7 – it was difficult to know just what range of factors might be likely to affect cohort scores either positively or negatively. Two of these factors, working in conflicting directions, are noted here.

First, as children progressed beyond Primary 2 the support they received for developing their literacy diminished. Although the overall literacy initiative incorporated a vision to tackle literacy ‘from the cradle to the grave’, the actual level of intervention was much more intense in the age groups up to P2 covered in the

main study than in the stages from P3 onwards. Children therefore finished the early intervention and at times went back into settings that did not support their enhanced literacy in the same way. This issue was being addressed in terms of new initiatives for P3 onwards as the research proceeded.

Second, and working in the opposite direction, there were factors likely to affect scores positively over the years in the later primary stages. One of these was the overall impact of this research. It was designed to touch every part of society throughout an entire Council area. The result, consistent with the key context variables frequently outlined in this work, was that the profile of reading went up, not just where the interventions were formal and intense but also elsewhere. The intensive focus on reading in the early stages, together with the central involvement of all the head teachers, meant that there was a greater ‘buzz’ everywhere about everything to do with reading. This not only influenced class teachers in the later primary stages – who had to adjust curricula to the higher levels of children coming through the intervention – but it also influenced families. Young children excited by the reading programme, not just in the main study but in all the interesting things happening with, for example, the synthetic phonics study, had siblings and parents. Although work in supporting parents is not separately reported on for the purposes of this study, many of the parents commented on the children’s enthusiasm and high achievement, and many examples were cited of older siblings having a heightened interest in reading. Teachers trained on the early intervention programme also at times moved to upper classes, taking with them new strategies for enhancing reading.

A final observation is relevant under the general heading of limitations and cautionary comments. Most of the baseline assessments were carried out by the class teachers, and they clearly had an interest in seeing their pupils performing well. Might they therefore in some way have ‘massaged’ the results? This is one of the ‘pitfalls in human research’ (Barber, 1976). Some discussion of this topic has already been raised in Chapter 7 of the Phase 1 Report regarding the features of the baseline assessment, and the question of when it is either appropriate or inappropriate to ‘teach to the test’. The overall consistency in the trend of results across the years and across so many schools and nurseries, with hundreds of people involved in testing, militates against this view, as does the consistent qualitative testimony to enhanced literacy from diverse sources, as observed below.

However, it is also answered by the very close monitoring procedures established for the research. Often the assessments were undertaken by members of the early intervention team instead of the class teachers, or else they were split between them. This took place not just for the initial purposes of establishing the reliability of the baseline, but throughout the intervention. Most particularly, if any results looked as if they needed closer monitoring, steps were taken to address this. Schools with scores that seemed inconsistently high or low compared with the trend of results or with their own known performance were visited for clarification, usually by the head of early intervention and a quality improvement officer. The early intervention team knew very precisely throughout the year which schools were performing at what levels, and were a valuable source in clarifying results. Only two schools were identified where inappropriate use of baseline test materials was suspected. To

address this an alternative form of two key tests was designed – the non-word reading test and the word reading test. This was issued without prior notice to these schools and to several other schools for comparison at the time of the baseline, and one establishment was advised on its practice as a result.

Raising achievement and eradicating literacy

Against the background of the limitations and cautionary comments that have been discussed, the main study in achieving its aims has significantly advanced the overall vision underlying the research, that of addressing endemic social and educational problems by raising achievement and seeking to eradicate illiteracy in socially disadvantaged populations. The question of whether change might have taken place anyway, even without the intervention, in this type of cross-lagged design has been considered in the conclusions. However, the size of the gains in the key areas of literacy addressed by the study, and the absence of a basis for these changes occurring for other reasons, points to the success of the programme.

In essence, the changes were marked enough and specific enough that everybody knew that patterns of achievement had been definitively changed. It hardly required even the baseline assessments, far less the statistical analysis, to inform every part of the education system that reading standards were changing. The main study did not gather systematic qualitative commentary from schools on the effects of the intervention – but it was certainly proffered. Head teachers, class teachers, the various types of support staff, parents and in many cases the children themselves knew that new levels of success had been established.

These levels of success are probably demonstrated most clearly in simple terms that hardly need the support of inferential tests. For example, in many cases children were scoring more than a year higher than their pre-intervention cohorts – that is, pupils in the pre-school year were scoring above the P1 controls, and P1 were scoring above the P2 controls. Instances of this were found in tests such as non-word reading and word reading, and for the lowest scoring children on phonological awareness and early reading skills. From the first to the last of the years reported here, the changes were very meaningful in practical terms at classroom level. This is seen, for example, in the rise from 12% to 50% in pre-school pupils attaining a perfect or almost perfect score for concepts of print; in the rise from 28% to 92% in P1 pupils able to recognise 20 or more letter sounds; and in the rise from 31% to 78% in P2 pupils achieving a score of 30 or more on the word reading test. These results were particularly encouraging as the baselines were conducted not at the end of the session for these three stages but half-way through the session.

One of the features of the results over the years was that they continued to rise year on year. Certainly a large rise was anticipated and was obtained in 1998 after the first year of intervention. However, these gains continued to increase each year till the last reported results in 2006, even if the annual rate of increase was not as large as time progressed. The four selected years reported here – baseline in 1997, the first year of intervention in 1998, the sixth year in 2003 and the 10th and final year in 2006 – were individual points in a steadily rising graph. Since a new group of pre-intervention

children arrived at the pre-school stage every year to begin the programme, it meant that constantly higher outcomes were being attained from a similar starting point each time. Not only therefore were the children doing better, they were doing ‘more better’ each year. This reflects the build up of the programme over time, and the installation or ‘institutionalisation’ of the longer-term processes of educational change as discussed in Chapter 2 of the Phase 1 Report.

The results reported here achieved everything hoped for by way of preparation for the vision of eradicating illiteracy from the authority. The numbers of children experiencing reading failure were systematically reduced. Pupils who were scoring at or near the tail end on the baseline tests in 2006 would in many cases have been viewed as average pupils scoring around the midpoint in 1997. In some instances low scores, once highly prevalent, almost disappeared. For example, approximately 40% of P1 pupils could recognise only 10 or fewer letter sounds in 1997. In 2003 this had reduced to just over 1.7%, and in 2006 to 1%. In P2 10% of pupils scored five or less on the word reading test in 1997. By 2003 this was down to 1%, and by 2006 to one-third of 1% – a mere 5 pupils across the 35 schools. This reduction in the numbers failing as they enter the later primary years created greater scope and economic feasibility for supporting these pupils with intensive individual help to overcome their difficulties.

SUMMARY

This chapter provides a detailed breakdown of the main results of the study. It does so by focusing on baseline assessment comparisons between pre-test cohort controls and subsequent population cohorts at the same school stage following intervention. This is done by providing data for four key years: 1997 (pre-test scores), 1998 (scores after first year of intervention), 2003 (scores after six years of intervention) and 2006 (scores for the 10th and final year of the study). The results support the three key hypotheses: first, the intervention was effective, with large improvements in performance in baseline test scores; second, there is evidence that the effects of the intervention have had a lasting impact after children left the main programme; third, all groups, both high achievers and low achievers, benefited from the intervention. In particular, the number of children obtaining low scores in key literacy attainments was greatly reduced.

Chapter 4: Eradicating Illiteracy

The Final Results of the Individual Support Study

Relation of Final Report to Phase 1 Report

The individual support study developed another of the 10 strands of literacy intervention described in the main study, strand 7, 'identification of and support for children who are failing'. This was a key strategy for addressing one of the main long-term aims of the study, the eradication of illiteracy throughout the entire school-age population. While the baseline assessments in the pre-school year and Primaries 1 and 2 provided an excellent basis for early identification of reading failure, it was also recognised that two factors relating to the later years of schooling required to be addressed. First, there were many pupils in the upper primary years and in secondary school who were already experiencing reading failure and who were not going to benefit from a literacy intervention focused on the early years. Second, a strategy was needed for the identification and support of children who were still failing even after they had been through the early intervention programme. It was in relation to tackling these factors that the individual support study was developed.

Chapter 15 of the Phase 1 Report provided a full rationale for intensive individual support for pupils who fail to develop adequate literacy skills after the middle years of primary schooling. It also outlined the basis on which one particular programme, Toe By Toe (Cowling & Cowling, 1993), was selected for the study, together with a description of that programme. It involved the provision of structured individual teaching in basic literacy skills for 20 minutes each day. In addition, support for the use of Toe By Toe was provided in an account of published and unpublished research studies.

This report provides an overview of the key results of the study that had been completed at the time the Phase 1 Report was published. In addition, it brings the individual support study to a completion by providing the final results of the intervention in effectively eradicating illiteracy from the population of West Dunbartonshire school leavers in the summer of 2007.

Aims

The aims of this study were:

- to carry out an effective intervention in a secondary school for acquiring key literacy skills based on individual support
- to extend this intervention effectively to the children at P7 stage throughout the authority with the most significant reading difficulties
- to extend the intervention to cover every pupil in secondary school who had not achieved functional literacy levels.

Hypotheses

The following hypotheses were proposed:

- 1 that the experimental pupils in the secondary study would achieve higher reading scores than the controls
- 2 that the children on the programme in the primaries would show large gain scores
- 3 that illiteracy would be eradicated at school leaving stage by application of the programme to any pupils who had not yet achieved functional literacy levels.

Method

This study was carried out in three phases. First, a quasi-experimental study was conducted in one West Dunbartonshire secondary school, with 24 pupils referred for learning support because of low reading levels. Of these pupils, 12 controls were assigned to the normal learning support programme while 12 experimentals were enrolled in the intensive individual support programme using Toe By Toe. The allocation of cases to the two conditions was not random, as it had to be subject to the normal timetabling and other constraints encountered in a large secondary school. However, the two samples were matched as closely as possible.

Normal learning support for the controls comprised two one-hour tutorial sessions, one concentrating on the development of basic punctuation and comprehension skills and the other developing phonic skills using standard phonic workbooks. Time was also allocated to individualised spelling and paired reading programmes which took place within the mainstream English class, with support for learning staff assisting in a co-operative teaching capacity.

All experimental pupils received individual tuition for 20 minutes a day, and the programme lasted approximately three months. Pre-post assessments were conducted at the start and finish of a 12-month period.

Following further piloting of the programme in secondary and primary schools in the authority, the second phase of the study involved the identification of pupils at upper primary level (mainly Primary 7) who were experiencing significant reading difficulties. These children were initially identified by staff in the 35 primary schools in the authority, and following individual testing of 118 children, 104 were selected from 32 schools as meeting support criteria. The 14 who were excluded from the sample had a reading age above 9y 6m and these pupils were not viewed as having a reading problem. The final sample comprised 91 in P7, 12 in P6 and 1 in P5. Pre-testing took place in November-December 2002, with post-test in May 2003.

Approximately 120 individual support workers were trained in the use of the programme by the researchers. These were drawn from a wide range of personnel – teachers, classroom assistants and volunteers. Monitoring and support structures were put in place to ensure effective implementation.

The final phase was the extension of the programme to cover every pupil in secondary schools throughout the authority who had not yet achieved functional literacy levels. Following the overall success of the early intervention programme and the second phase of the individual support study, the number remaining was very small. A total of 12 pupils were identified at the beginning of Session 2006-07 from the seven secondary schools in the authority. All of these had been individually assessed on more than one occasion over a period of time on the Neale Analysis of Reading Ability (Neale, 1989). Their reading ages ranged from 6y 9m to 8y 10m.

Results

In relation to the first phase of the study, the 12 experimental pupils in secondary school showed mean reading age gains of 2y 0m (from 8y 2m to 10y 2m) following the three-month Toe By Toe intervention, and with a 12-month interval between tests. The controls gained only four months (from 8y 5m to 8y 9m) during the same period. Analysis of results was carried out by means of independent two-sample *t* tests using the Microsoft Excel Version 8.0 data analysis tools. The test used was the Gapadol Reading Comprehension Test (McLeod & Anderson, 1972). Effect sizes were calculated using the standard deviations in the test manual. These results are shown in Table 4-1.

In relation to the second phase, the 104 children in the primary schools study were all individually tested on the Neale Analysis of Reading Ability, 2nd Revised British Edition (Form 2). Their average pre-test reading age was 8y 0m, this being about three years behind their chronological age. After a period of just under six months their post-test reading age had risen to 9y 2m, giving an average gain score of 1y 2m. The breakdown of gains is shown in Table 4-2.

In relation to the final phase, at the beginning of June 2007 only three pupils remained with Neale Analysis scores below the 9y 6m level of functional literacy. In addition, one pupil was withdrawn from the programme by his parents and was not available for either further intervention or assessment. One other did not attend and was not available for the final assessment. Help continued to be offered to these pupils, and the possible reasons for their difficulties continued to be investigated.

Table 4-1 Changes in reading age: secondary school sample (N = 24)

	Pre	Post	Significance	Effect size
Experimentals	8y 2m	10y 2m	p<0.001, <i>t</i> = 5.65	1.74
Controls	8y 5m	8y 8m		

Table 4-2 Gains in reading age: primary school sample (N = 104)

<i>Gains in reading age</i>	<i>Number of pupils</i>
0-5 months	19
6-11 months	25
12-17 months	30
18-23 months	11
Over 2 years	16
Over 3 years	3

There were no differences of any note in the gains made by children in terms of lower or higher pre-test scores. When the sample was divided in half, those who started with lower reading ages had mean gains of 1y 1m, while those in the higher half gained 1y 3m. The lowest 25 children on pre-test – those who began with a reading age below 8y 0m – gained 1y 0m. Thus there were mean gains of a year or more for the whole sample, irrespective of starting point. At the bottom end, 18% of the sample (19 children) showed gains of less than six months, while at the top end 18% showed gains of two years or more.

Conclusions

The individual support study proved to be highly effective not only in addressing reading difficulties both in the secondary and in the primary school samples but also in achieving the aim of eradicating illiteracy at school leaving age.

The average gain of two years in reading age for the secondary pupils was particularly encouraging. These gain scores were so large that they would have demonstrated the effectiveness of the intervention even without the need for a control group. However, the use of a quasi-experimental design allowed informed comparisons to be made between experimentals and controls. The fact that, even with a good programme of traditional learning support, the gains of the controls were only four months in the course of a full year confirms the routine experience of learning support teachers and educational psychologists – namely, that pupils at this level of reading difficulty tend normally to make annual gains that are less than half of what might be expected. That is, in the course of a year they are making somewhat less than six months of improvement in their reading ages.

There are times when statistically significant changes do not make a real impact either on the perceptions of subject teachers or on the quality of life of the participants. This was not the case in the present study. Teachers of various subjects frequently commented that the pupils on the Toe By Toe programme were reading more competently than had ever been the case before with similar pupils. The position is best summed up in the words of one 14 year old pupil in the study, as cited elsewhere in this work:

When all this started I couldn't read. I was a failure. Now I have a cupboardful of books at home. My favourite authors are Roald Dahl and J.K. Rowling. Now I am a success.

The gains achieved consistently by the secondary school group provided the basis for the extension of the study into the primary schools as an intervention with an established record of effectiveness. Here again the gains were sufficiently large that they showed the programme to be successful without the need for comparison with controls. These pupils would normally have made significantly less than six months of reading gains over a six-month period. The consistent experience of learning support staff, combined both with the results of the control group in the secondary study and with any assessment of normal rate of progress of these pupils over the years, would point to usual gains of between two and three months during any six-month period. The actual gains during the intervention of 1y 2m represent a population shift of 1.75 standard deviations on the Neale Analysis of Reading Ability. In terms of the above observations, the estimated 'expected' shift (that is, the normal mean gains without intervention) would have been about 0.35 standard deviations over the pre-post interval. Therefore, the effect size of the intervention may be estimated at about 1.4. This suggests that the programme was extremely powerful over this short period in increasing levels of reading achievement.

Two further observations may be made regarding the progress of the primary school sample. The first is that in comparison with the secondary sample their gains were 1y 2m as against 2y 0m. However, two main considerations are likely to be of importance here. First, the pre-post test interval was one year for the secondary pupils, while it was less than six months for the primary sample. It was expected that the primary pupils would continue to make reading progress over the succeeding months, thereby narrowing the observed gap in average reading gains. Second, the secondary pupils were all taught by one teacher, who was already an expert in the use of the programme. By comparison, the intervention with the primary pupils was spread across 120 helpers, many of them volunteers, and most of them being inexperienced with the programme other than for a training session of half a day.

Comparisons of the weight and significance of gains in reading made by pupils of different ages and in different intervention conditions can be difficult to interpret. The concept of 'ratio gains' may be applied to the scores reported here, as defined by Topping and Lindsay (1992) – '...the gain in reading age made...on a reading test during a chronological time span, expressed as a ratio of that time span'. The mean ratio gain of 1y 2m in approximately six months for the primary sample is therefore about 2.3.

Overall, the results achieved are clearly very good and point strongly to this programme as being an effective intervention that does not depend on experts or on intensive training. Indeed, follow-up investigations carried out since this phase of the study was completed have indicated that many of the helpers were not even carrying out the programme correctly and needed a higher level of monitoring and support.

Arrangements to undertake this were put in place, and it is expected that future interventions may show still higher gains.

The second observation is that the primary sample included 19 pupils whose progress over the six-month period was less than six months in terms of reading age gains. However, several of these children were in fact making progress at a reasonable rate given their level of difficulties. In many cases they had still not completed the programme at post-test and further progress was therefore expected. In other cases it was clear why less progress had been made. With a sample as large as this, dependent on so many helpers, there will always be some slippage in the intervention, such as helpers being off sick or pupils being absent from school. Every case of low progress was investigated and further help was arranged to ensure success for all pupils.

The fact that for the final phase of the study the number of potential school leavers failing to achieve functional literacy level was reduced to three pupils, with two others unavailable for assessment and intervention, demonstrates the ultimate success of the whole initiative as an effective means of eradicating illiteracy from an entire education authority. Opportunities continued to be available for any who still had not achieved success.

In conclusion, this study provided strong support for strand 7 of the main study, 'identification of and support for children who are failing'. A foundation has been laid for the eradication of illiteracy throughout the primary schools, with a rolling programme in place each year to identify every individual pupil with a reading difficulty. At the end of the brief intervention for primary school children, over one-third of them were no longer described by their teachers as having a 'reading problem'. This leaves very small numbers proceeding to secondary with continuing concerns about not achieving adequate literacy during their school career. This is a crucial area of intervention for these pupils as they prepare to face a future beyond school. As Wells (1998) has observed:

While it's important to get the teaching of literacy and numeracy right in primary schools, early intervention will be too late for some older pupils. So opportunities for catching up in secondary school will need to be given high priority. If they aren't, many pupils won't be able to get much benefit from the wider curriculum. And some will leave school with basic skills that provide hardly any grounding for the world of work and later education and training (p. 1).

While it is clearly of crucial importance to identify children who are failing in literacy at the later stages of their primary schooling or in secondary school, and to provide them with effective interventions, methods must be found at an earlier stage of identifying those children who are likely to experience difficulties at a future date. This may be done by a combination of baseline assessment results and responding to concerns raised by teachers about children who are not making adequate progress in the early stages of literacy. Four-fifths of the children who were identified as reading failures at around Primary 7 level were to be found in the bottom quartile of baseline assessment scores for key literacy skills when they were in Primary 2. Almost half

had scores falling in the bottom 10%. Only one child was identified as having significant difficulty in reading at Primary 7 who scored above the midpoint for reading scores in Primary 2.

Table 4-3 shows a breakdown of the distribution of scores for early reading skills for the 53 children out of the 104 in the primary sample who could be identified in the Primary 2 baseline assessments. Early reading skills (letter sounds, the alphabet, letter names, non-word reading and word reading) proved to be a much more robust predictor of later reading difficulties at this stage than phonological awareness (nursery rhymes, initial sounds, rhyme detection and rhyme production).

These results suggest that a suitable starting point for identification of future reading difficulties at Primary 2 stage would be the bottom 10% of early reading skills scores on the baseline assessments. This one measure would identify approximately half of the children who would be seen to have a reading difficulty in the later primary years.

Table 4-3 Prediction in Primary 2 of later reading difficulties (N = 53)

	Early reading skills		Phonological awareness	
	%	Cumulative %	%	Cumulative %
<i>Bottom 5%</i>	25	25	23	23
<i>Bottom 10%</i>	22	47	3	26
Quartile 1	79	79	51	51
Quartile 2	19	98	26	77
Quartile 3	2	100	17	94
Quartile 4	0	100	6	100

SUMMARY

This chapter describes a subsidiary study designed to develop a key strand in the multiple-component intervention – identification of and support for children who are failing. It examines the rationale for providing individual rather than group support. A quasi-experimental study at secondary school and a gains score study at primary are described. The secondary study involved a comparison of 24 pupils, 12 experimentals who received an intervention programme based on individual support and 12 controls who received the normal learning support package. A commercially-available programme, *Toe By Toe*, that met the specifications for the study, was used. This involved individual, structured tuition in basic literacy skills for 20 minutes each day. The intervention lasted for three months, and pre-post assessments were conducted 12 months apart. Significant gains were made by the experimentals, with an average reading gain of two years compared with 4 months for controls. Following the secondary study, 104 pupils in upper primary school were identified on the basis of low reading scores. They were given the *Toe By Toe* intervention

from volunteers and teachers following a brief training session. In less than six months the average gain scores for these pupils was 1y 2m. Finally, the number of school leavers who had not achieved functional literacy levels at the time of the final assessment early in June 2007 was only three pupils, with two others unavailable for assessment and intervention. These results point to an economical and effective way of addressing reading failure and in contributing to the eradication of illiteracy.

Chapter 5

Conclusions

In terms of the breadth of its vision, the scale of its implementation, the extent of its long-term gathering of follow-up data and the distinctiveness of its novel interventions, it is proposed that this study had made a unique contribution to research in this field, and that it has significant implications for educational policy and practice. This concluding chapter examines these implications, as well as providing a critique of its methodology.

The design, implementation and evaluation of a literacy intervention

The aim of this 10-year research study as outlined at the outset was to design, to implement and to evaluate the effects of, a multiple-component intervention to raise achievement and address illiteracy in areas of socio-economic disadvantage, taking full account of the factors affecting educational change in the context of real world research.

It is concluded that the study has been successful in the achievement of this aim. Not only has a complex, large-scale intervention in literacy been designed from its first conception to the point of comprehensive implementation and evaluation with a whole population of children and young people, but it has also been rewarded with highly encouraging outcomes. These outcomes have led to significant raising of achievement for all groups of pupils from the most to the least able. In tackling the attainment levels of the latter, the intervention has for all practical purposes achieved a goal that seemed impossible – the eradication of illiteracy throughout the school population by the end of the 10th year of the programme.

This study was designed on a grand scale. The main study alone involved assessments of 60,808 children, of whom 30,903 were assessed individually, the remaining 29,905 being group assessments. It also involved 58 educational establishments, of which 23 were nurseries and 35 were primaries. Across these schools over 400 staff had to be trained, monitored and supported to ensure high fidelity of programme delivery.

Implementation involved key personnel working constantly and systematically through the years with every level of education management, from educational directorate, through quality insurance structures and primary and secondary head teachers and heads of pre-five establishments. It also required extensive contact with hundreds of workers at ground level – class teachers, classroom assistants, learning support teachers, SEN auxiliaries and volunteers, together with collaboration with the psychological service and other services contributing to the overall programme. In addition, it involved planning and recommending the key support structures required to carry out the project effectively, as well as the organisational tasks involved in the appointment of early intervention teachers and other key workers.

To evaluate the main study a full baseline assessment scheme was designed and produced. This involved consideration of a complex national and local context regarding baseline assessment in schools (Scottish Office Education and Industry Department, 1998), to take account of which a survey was conducted across all 32 education authorities in Scotland (MacKay, 1999a). To support the baseline assessment, hundreds of teachers and other staff had to be trained in assessment methods, and this had to be repeated every year for new cohorts of workers. The subsidiary studies required setting up and carrying out training programmes for hundreds of additional staff, together with designing and carrying out further individual assessments for hundreds of additional children.

Carrying out all of the above requirements was a considerable logistical exercise. It necessitated the coordination of the work of staff at many different levels, across a number of services and departments and across a wide range of locations, as well as the establishment of training arrangements and support structures. Ensuring that the mechanisms and resources were in place at the right time on every occasion baseline assessment or other assessments were required also involved complex logistics. The resultant data represented an immense mountain of test papers that required to be scored and then processed, and it is estimated that for the main study around three tons of assessment papers were generated in the course of this exercise.

Limitations

Any critique of this study must recognise a whole range of limitations. Some of these are perhaps inherent in a study of this kind and might almost be predicted by Robson's (1993) observation that 'real world research' tends to be done in 'complex, messy, poorly controlled "field" settings' (p. x). His assessment of real world research as mainly solving problems rather than just gaining knowledge, as predicting effects rather than finding causes, as looking for large effects rather than studying relationships between variables, as developing and testing interventions and services rather than theories and as using multiple methods rather than single methods, all resonates with the nature of this study.

As sample size increases from small numbers in a single school to whole populations in many diverse establishments, as co-workers increase from one or two dedicated research assistants to vast numbers of teachers and other staff, and as intervention methods increase from single to multiple-component strategies, so also is there a corresponding increase in the complexity, messiness and poor control that characterise these settings. It is acknowledged that these features were often very prominent in this investigation, and they may provide a context in which at least some of the limitations in the study may be considered. This section notes a range of limitations and weaknesses or possible weaknesses. Others of a more specific nature have already been noted within the studies to which they refer both in this report and in the Phase 1 Report.

First, any study that claims to be driven by a commitment to values in science and the aspiration to 'do good', that is, to seek to apply psychology to human welfare, will almost certainly, among other possible criticisms, be vulnerable to the censure of

those who have a different view of what is 'good'. A key aim of this study was to raise literacy levels among disadvantaged children – indeed, to change their own attitudes and values in the process, and in particular to ensure that all would attain 'functional literacy'. This philosophy might be (and indeed has been) challenged, and is open to Levine's (1986) criticism that this merely serves to increase conforming behaviour in the participants and bring them yet more within bureaucratic communication and authority, with their literacy at just an adequate level to be 'functional' in increasing their usefulness and subservience to society rather than functional to themselves. The alternative view (for example, Freire, 1994, 2000) is that literacy provides empowerment, increases choices and improves quality of life. This is the view that is taken in this study.

A second general weakness is methodological. Throughout the study the researcher's ideal of double-blind randomised control trials was left far behind for designs which, even though they usually reached quasi-experimental level, were far from being blind as far as the people conducting the assessments were concerned, and furthermore, at times the assessor was also the researcher. This is frequently again a feature in real world research in schools, and often is a reflection of resources. For example, in the declaration study described in the Phase 1 Report it was expected that the assessments would be undertaken by others, but when the time came the resources for this were not available. The researcher therefore reluctantly had to carry out the unplanned task of undertaking 120 individual baseline assessments.

Possible vested interests of the assessor in the assessment results were perhaps less of an issue in the vast samples that constituted the main study. Although most of the assessments were carried out by the class teachers and nursery staff themselves there were several factors that mitigated or monitored assessor effects. While undoubtedly the teachers would have had a wish for their own children to do well, the same wish was likely to have been present at pre-test baseline. The differences, however, between pre-test and post-test situations were generally large. Also, as the tests were conducted in November/December each year, the teachers were not assessing situations that reflected only their own work, but also that of the teacher or establishment responsible in the previous session.

Monitoring took place at a high level. Assessment papers were screened for any apparent anomalies that might cast doubt on the validity of the results, using many sources of enquiry such as patterns of test performance, knowledge of children's previous performance levels and the detailed knowledge of the levels of the children in each school held by the early intervention teachers. If such factors gave any cause for concern, and also at many times for other reasons, other staff such as members of the early intervention team undertook half or all of the assessments in a particular class. In addition, parallel forms of key baseline tests were designed and were used in a range of establishments without prior warning, not just where there seemed cause for concern but more generally. Throughout the 10 years of the study only two schools caused concerns that they might not be using the tests appropriately. These situations were investigated thoroughly, and as a result one school was advised of its practice.

Third, despite the vastness of the sample used in the main study, a number of issues may be raised about sample size and characteristics. For example, in the attitudes study, although all but four of the original sample were traced in their secondary schools five and a half years after the original study, the resultant number (after removing one pupil because of significant issues of 'caseness') was only 19, and 11 of these were experimentals. Certainly, the original study, despite its small sample (dictated by the constraints of a single establishment) was a randomised control trial involving three groups, and the sheer size of the experimental v. control differences provided robust results. The differences at follow-up were also large, but the small numbers did constitute an obvious weakness.

Both sample size and sample characteristics were a limitation in the declaration study. The total number of participants was high at 565, but the core data available for the main analysis was 60, reduced to 54 at post-test because of changes of school or being absent at every opportunity for assessment. This meant that the sample was then reduced to 27 experimentals and 27 controls. This somewhat limited the range of analyses available. Of these 27, 18 were primary, leaving only 9 per group at nursery level (reflecting the smaller number of nurseries compared with primaries). The results pointed to real change in favour of the experimentals, supported by a wide range of other more qualitative data. However, the number of possible breakdowns that would have been of interest with this sample – by test, by socio-economic status of establishment, by high v. low achievers, by gender – was clearly very limited.

Sample characteristics proved to be an unexpected and indeed irksome feature in the declaration study. Analysis of pre-test scores in the group who were individually tested pointed to systematic differences in baseline performance in favour of the controls (that is, the controls began with higher scores). Without calibration from other sources, this might have appeared as an accidental or deliberate assessor effect (the assessor being the researcher, as noted above). That is, the assessor might have preferred the scores of the experimentals to be depressed to make later change easier (although it would be obvious that unwelcome pre-test differences in the groups would then be apparent). While randomised allocation to experimental or control conditions would have controlled for this, the practical situation did not support it, as the experimental schools had to be agreed with the education authority from the start. The possible assessor bias was calibrated when assessments provided by the schools showed the same pre-test differences. Despite a good systematic sampling procedure, it is likely that, other than totally random variation, the control schools influenced the ultimate sample by not wanting to have poor scorers as their representatives, while the experimental schools wanted more needy children to have the pre-post intervention assessments. In the event, the analysis took account of these differences and indicated that pupils across both groups who started with higher scores did not differ in their progress from those with lower starting scores.

Fourth, a number of issues may be raised in regard to the range of assessment methods used. The baseline assessment designed for the study had many strengths, and its usefulness was demonstrated by the fact that when education authorities were surveyed a number of them were using it in whole or in part. However, of the seven

principles it was designed to meet – reliability, validity, utility, directness, reactivity, sensitivity and feasibility – an overwhelming one was feasibility. That is, it had to be possible for the education authority to fund its production and administration for thousands of individual assessments each year. For evaluation purposes it was also desirable to design a test that could be used across the full age range of the main study, from the pre-school year to Primary 2. One of the outcomes of this was that some tests were limited in usefulness for the lower age range, having too high a floor (such as word reading), and others for the upper age range because they had too low a ceiling (such as concepts of print). This meant that floor and ceiling effects were often apparent. Nor was it possible to follow a simple expedient of combining scores into a grand total, an exercise with possible advantages but raising further questions about adding together some quite disparate tests in terms of weight and balance, as well as losing data specific to individual tests.

The Norman France Reading Tests presented an unavoidable gap in assessment measures between the stages up to P2 and the stages from P3 onwards. Their scores did not have the sensitivity or detail of the baseline tests, and they clearly had a high loading on attentional and cognitive factors, as well as being difficult to monitor effectively to ensure independent work when undertaken in groups.

Fifth, potential weaknesses in the assessment measures were reflected in the analyses carried out, and these have been discussed both in the Phase 1 Report and in this report in relation to individual studies, and most particularly in relation to data skewed by floor and ceiling effects. The unsophisticated level of analysis reflected Robson's (1993) observation cited above regarding real world research often looking for large effects rather than studying the relationship among variables, and the vast body of data generated has not as yet been exploited in terms of the range and sophistication of the analyses that might illuminate aspects of the study more clearly. Certainly, many of the effects were indeed large, consistent and self-evident, but not all of them were, and these are the ones that might particularly benefit from further appraisal. Also, the advantages and disadvantages of using pre-intervention population cohorts as controls for post-intervention cohorts at the same stage rather than establishing 'true' control groups have been recognised. The issue must be raised as to whether the subsequent year groups might have had better scores without the intervention. This question is of great importance in relation to the main study, and is covered separately below.

Sixth, a further question that may be raised about the project is not so much whether the results rose, but whether the skills measured represented a balanced and meaningful range of literacy abilities. The rationale for selecting the items for the baseline assessment in terms of their usefulness and predictive validity is covered in detail in Chapter 7 of the earlier report. At the same time it may be noted that the assessments were limited to the very mechanical skills of sounding, blending and word recognition that can be easily assessed in the early stages rather than the broader area of reading comprehension. However, apart from the inherent difficulties in assessing comprehension accurately in ways that separate it from broader cognitive abilities in these early years, it is recognised that skill in the mechanical measures assessed is critical to the development of higher order reading abilities (see

Chapter 6 of the Phase 1 Report). When a child has not developed mechanical fluency, but has to spend time deciphering words and their phonic elements, comprehension is impeded. All who are familiar with administering tests covering both accuracy and comprehension, such as the Neale Analysis of Reading Ability, will be very familiar with the issues arising here. Mechanical skills were therefore the focus of baseline assessment for the early stages so that a solid foundation would be laid to develop higher order skills later.

The final limitation noted here again reflects Robson's observations about studies in the 'real world' looking for large effects rather than studying relationships between variables, developing and testing interventions and services rather than theories and using multiple methods rather than single methods. While this study has achieved its key aim of designing, implementing and evaluating this large-scale intervention, and demonstrating that it has raised achievement and tackled illiteracy, it leaves a number of unanswered questions regarding which variables had what effects. The whole sample had the whole intervention. The weight of each strand or component in the intervention may have differed between establishments, and the very objective of maximising 'ownership' of the project in each school tends to highlight differences in approach, but the effects of each component cannot be separately assessed. Might some components have been omitted without weakening the strategy? Might most of the variance have been accounted for by one or two components, such as more time spent on reading? What was the contribution of the context variables, involving such intangible concepts as 'vision' and 'profile'? For many reasons the simplest response to these questions is that they must represent the subject matter of another study. For the present study there was a different ambition that transcended these more detailed questions, namely, to lay the foundation for intergenerational change in the achievement levels of a disadvantaged population, and to seek to maximise this in every way possible. Many strands were interwoven to produce a successful and sustained outcome. The end result was that achievement levels rose, not just statistically, but in a meaningful way that can be crystallised in the experience of a rising generation of successful readers and their families.

Pre-intervention cohorts as post-intervention controls

The issue that affects all year-on-year cohort studies where the baseline results for the first year of a study represent the standard by which future cohorts at the same stage are evaluated is the question as to whether standards might not have been rising anyway from one year to the next. Since this question is of fundamental importance to the main study it is treated here in detail. Four comments may be made. First, the Assessment of Achievement Programme was established by the Scottish Office Education and Industry Department in 1981 to monitor the performance of pupils in Scottish schools in particular areas of the curriculum including English language, on a three-year cycle for each subject. The stages assessed were P4, P7 and S2. Although the results do not allow precise comparisons with the assessments reported here in terms either of age group or of content, the overall drift of findings does not suggest a tendency towards systematic rises in reading attainment from year to year. For example, for pupils in Primary 4, the period from 1984 through to 1995 showed at times a slight fall and at other times no change in reading attainment. Despite

gains reported in relation to a number of early intervention programmes (Fraser, MacDougall, Pirrie & Croxford, 2001), reading standards at all stages showed a marked fall through much of the period reported in this study, from 1998 to 2001 (Scottish Office Education and Industry Department, 1996; Scottish Executive Education Department, 2003). From that perspective there would be no indications that a rise in results might have been expected at that period.

Second, no evidence is available to suggest that standards were rising in core educational attainments in general during these years in the area in which the main study was carried out. While the project was taking place a much less intensive intervention on raising attainment in numeracy was being designed, implemented and evaluated, including the design of a numeracy baseline assessment for the purpose. Only modest gains were reported in numeracy throughout that period, more or less proportional to the size of the intervention, which received much less resourcing than the literacy intervention. Therefore there is no indication that significant improvements in literacy were simply part of a general rise in educational standards over these years.

Third, the changes in the baseline assessment results did not reflect minor variation but showed large effect sizes. As reported for the main study, pupils scoring on the various baseline tests at the 50th percentile in 1997 before the intervention began would by the standards of the 2006 baseline have been described no longer as average but as low scorers. A score at the 50th percentile in 1997 for lower case letter sounds in Primary 1 would have been at the 1st percentile in 2006. In short, the results are consistent with the expectations of a highly successful intervention programme.

Fourth, at a less formal level, the consistent reports of class teachers, specialist teachers and many others associated with the programme overwhelmingly confirmed their view that the intervention had brought about manifest changes in the levels of attainment of the pupils who were participating at all levels of ability.

The process of educational change

The study has been guided throughout by the belief that ‘change is a process, not an event’ (Fullan, 2001, p. 52). Full recognition has been given to Fullan’s (2001) three phases of educational change – initiation, implementation and institutionalisation – together with a commitment to the view that years of systematic collaborative working are required to achieve a successful outcome of these phases.

The preparatory studies not only provided a good foundation for planning a large-scale intervention in literacy: they also contributed a crucial awareness of the inevitable and seemingly incessant factors that can undermine real world research programmes and the changes they seek to effect. An outline has been given earlier (Phase 1 Report, Chapter 2) of the immense and diverse difficulties that arose in a short-term, small-scale study involving only two establishments. It was expected that such factors would be multiplied in a study involving hundreds of teachers and other

workers, many thousands of children and scores of establishments over a period of years.

Indeed, this expectation was fulfilled. The difficulties were manifold, and the following are far from exhaustive. First, there were funding issues. The project required very considerable finances, including funds for the employment of a team of specialist teachers. The bulk of the resourcing was obtained from the Scottish Executive Education Department for the part of the project that related to early intervention – essentially the main study, and overwhelmingly the most costly element. This was available initially for the first three years, and was then obtained for a further two years. At that stage the Council had to make major decisions about mainlining project staff, and also had to work in a different funding context. At times these funding concerns affected the longer-term planning process, and led also to the next difficulty as noted below.

The second major difficulty encountered was in the realm of personnel. Issues over funding and related areas meant that specialist staff were initially seconded or employed on temporary contracts for a fixed term of up to three years. This meant that in the third of these years staff had no guarantee of a future with the project, and understandably many felt they had to look either for permanent posts elsewhere or for a return to the post from which they had been seconded. For a period this resulted in many schools having very limited support. Crucially this affected the position of the person appointed as head teacher for early intervention, who returned to a mainstream head teaching post. She was central to programme implementation and thus a major discontinuity could not be avoided for a period. In addition, other contractual issues arose. For example, the teacher seconded for the implementation of the individual support study was largely recalled from her secondment because of staffing difficulties in her establishment, bringing about significant problems for a period in monitoring and supporting the study.

Third, towards the end of the first phase of the research a protracted period of industrial action affected the nursery nurses in the pre-school sector. This resulted in the implementation of the programme being disrupted to a greater or lesser extent in the 23 pre-five establishments. It affected training, the availability of staff to carry out baseline assessments and the overall readiness of staff to implement any initiative they felt was additional to their established duties.

Finally, an example of the issues that can be much more easily controlled in tidier research projects than in the messy arena of such a large-scale, multiple-component intervention is found in the synthetic phonics study, where the opportunity for continuing with a control sample was overtaken by the proselytising enthusiasm of the experimental schools. Efforts of researchers to stop controls from ‘doing it’ can prove in vain when a group of animated teachers persist in talking about and sharing their practices and materials.

The range of difficulties noted above could be expanded in almost every direction. The intervention had to continue and to succeed through virtually every major change or turmoil taking place in its midst – including a total restructuring of the

educational directorate, together with significant changes in the Council. All of these challenges are fundamental to the process of achieving positive, long-term change in carrying out real world research on a large scale.

Implications for policy and practice

There is a sense in which the key implications of this study for educational policy and practice do not need to be spelt out, as they are axiomatic. A large number of education authorities – in Scotland, in the UK and in other parts of the world – have populations marked by high levels of socio-economic disadvantage. There is a consistent body of evidence on the impact of this factor on health, quality of life, educational achievement in general and literacy levels in particular. Such communities are characterised by educational underachievement and high levels of illiteracy. These problems have not been shown to respond either to generally available educational strategies or to many special initiatives. The multiple-component approach described here is not another ‘package’ to be purchased, but a process for enhancing the effective delivery of the reading curriculum. Its long-term effectiveness has been demonstrated, and while it requires resources and research expertise at significant levels, it is argued here that it is a necessary and cost-effective investment.

Cost-benefit analysis

Visionary aims of raising educational achievement and eradicating illiteracy cannot be pursued without reference to the economic costs involved in implementing effective programmes. Without question, the intervention reported in this study was costly. The target annual budget for the whole initiative at its conception was around £300,000. Translating this into cost-benefit terms is complex in terms both of defining the population that benefited directly from the intervention and of assessing the wider aims and impact of the project.

An attempt was made in the Phase 1 Report to present a cost-benefit analysis and this is summarised again here. In terms of the population that received direct benefits, the main study included every child in the pre-school year, Primary 1 and Primary 2. This was an average of 3,221 pupils per year through the first six years of intervention reported. If the costs of the initiative are related only to this population the cost per pupil would be £93 per year. However, it may be argued that this was an initiative for the whole education system, although with a primary focus on the groups specified in the main study. It involved inputs at many levels throughout nursery, primary and secondary schools. This was reflected not only in the application of the intervention to pupils in the upper primary and secondary stages, but also in more general effects of the initiative, such as enhancing the curriculum from Primary 3 onwards and in affecting other siblings in families of targeted children. Viewed as a whole education authority initiative, the project represented about £13 per pupil per year for pupils attending schools or nurseries in the authority. As a proportion of the education department budget, this represented under 0.5% of education spending at the time this analysis was carried out.

In terms of the wider aims and impact of the project, it is argued that this was an initiative with potential to have significant effects on quality of life and the economy throughout the whole population. If any of its stated goals were going to be achieved – of higher self-esteem, lower disruption in schools, better school ethos, better staff morale, economic savings in remedial support, lower crime, a more skilled workforce and a stronger economy – then in cost-benefit terms the expenses of running the project represented a modest investment indeed.

As a footnote on the cost-benefit analysis, it may be noted that the area in which the main study was conducted – the second poorest Council area in Scotland – despite the financial strictures it was facing, made a decision to fund the entire project after the specific research funding for it had been discontinued. This involved mainlining the posts for the head teacher of the project, the early intervention team and the home-school support teachers.

One of the issues that has been considered in terms of the policy and practice implications of this study is whether the baseline assessments should continue after they were no longer required for the research, as they too have considerable cost implications, not so much in production as in administration. It is a testimony to the usefulness of the scheme that the education authority has opted to continue with the baseline assessments following the completion of the research project. They were found to be of great value in informing teachers of the progress of every individual pupil, in identifying children at risk and setting targets. They supported an essential strand in the multiple-component intervention – raising teacher awareness through focused assessment.

Among the individual components of the intervention, the synthetic phonics study has highlighted the benefits of a strong and structured phonics emphasis. The study indicated the superiority of the synthetic over the analytic or traditional approach, and the clearest policy recommendation would be for schools to adopt this approach. Although this recommendation could be with confidence of good outcomes, caution would still be associated with this area. If there is an aspect in which the synthetic phonics study, and all of the existing evaluations, have left a continuing question it is whether the synthetic approach is ultimately superior because of its distinctive synthetic methods, or whether it has not yet been sufficiently systematically compared with better analytic phonics teaching using a faster pace and more motivating approaches.

In addition to the need for a multiple-component intervention in literacy for the early years, the provision of intensive individual support for failing readers in the later stages of primary and into secondary school is essential. Again, individual tuition is costly – but it is not ultimately so costly as illiteracy. The methods used in this study have proved to be successful, but without the high level of costs associated with other programmes of demonstrated effectiveness. Clay's Reading Recovery programme (Clay, 1979b, 1991, 1993a) requires extensive training of teachers to ensure successful delivery, and Iversen and Tunmer's (1993) modified programme used teachers with a higher degree in reading. The individual support study reported

here was based on a single brief training session for staff, many of whom were volunteers.

Finally, it is essential for education authorities to establish literacy initiatives that are long-term and not short-term, and that take full account of the processes of educational change and the loss of impetus that characterises many projects after the initial enthusiasm had waned. Although early intervention projects for literacy were established throughout Scotland in the late 1990s, many of these were small in scale, limited in scope and not enduring. Informal surveys by the author have highlighted several ambitious projects that have now terminated, some having simply faded out with the passage of time, with personnel changes or with the transfer of central funding to individual schools.

In summary, for areas of socio-economic disadvantage this study supports the establishment of long-term literacy initiatives with multiple components, with intensive individual programmes for failing readers, with detailed assessment and evaluation measures and with high levels of training, monitoring and staff support.

Vision, profile, ownership, commitment and declaration

Throughout this report reference has been made to the recognition given to key context variables: that is, to the idea that major educational change must be supported not only by good programme content but by a much wider context marked by vision, profile, ownership, commitment and declaration. These were built into the project at every opportunity. They were formally articulated and discussed at the meetings of the steering group that met regularly to monitor and plan the progress of the project. They were highlighted as being of fundamental importance at major conferences held to keep the project on track and to celebrate its results from year to year. They were taught to staff and volunteers at every level during virtually every training session. The essential message was: this programme has 10 content variables – the ‘10 strands’ – and it has five context variables.

The result was that the project has been defined and recognised as being ‘visionary’. The language associated with it at all times has therefore been visionary language. The original research proposal on which it was based began with the words, ‘This is a vision for transforming reading standards for all children in all schools throughout the education authority’. It was presented in the highest profile at every opportunity. On over 60 occasions it was seen in newspaper headlines or in magazine articles, and on several occasions it was featured on radio and television. The message to all who participated was that they were involved in something very important. Vision and profile promoted ownership and commitment. The project belonged wholly to everyone – it belonged to the Council, the directorate, to the early intervention team, to the head teachers, to the class teachers, assistants and volunteers, to the parents, to the children and also to the researcher. Commitment began at Councillor level. Few conferences or media reports lacked the visible presence of the Council leader or the chair of the education committee. In turn commitment was expected at every level throughout the authority.

The final context variable was declaration. It was practised informally from the start. There were great expectations – and they were declared boldly. This project was to be a world leader. It was to raise attainment, it was to wipe out illiteracy – and as a result it was to change lives. The aim in tackling low achievement and illiteracy was to tackle everything known to be associated with it. The results anticipated from intergenerational change were higher self-esteem, lower disruption in schools, better school ethos, better staff morale, economic savings in remedial support, lower crime, a more skilled workforce and a stronger economy.

If the success of this research has built the foundation on which these high ambitions will be accomplished, then it will have promoted the framework of values in science that it has endorsed – the values of promoting health, caring and compassion, self-determination and participation, human diversity and social justice.

Epilogue

'For all the money, time, energy and ingenuity we have spent on reading research, we are still at the stage of saying that children learn to read when there is something they want to read and an adult who takes the time and trouble to help them' (Meek, 1983, p. 1).

These were the words with which the epilogue to the Phase 1 Research Report of the West Dunbartonshire Literacy Initiative began, and they remain relevant as the study draws to its close.

In the 20 years or so since Meek published her qualitative, longitudinal studies of adolescents learning to read a great deal of knowledge about the reading process and the basis of effective teaching has been systematically accumulated. Nevertheless, there remains essential truth in the statement that the achievement of competence in reading is based on these two fundamental requirements – the motivation of the learner to learn, and the commitment of the teacher to teach.

It is for the purpose of elucidating this dual foundation of reading achievement that this research has been conducted. It has sought to investigate the circumstances in which the learner will be best encouraged to have the motivation to learn, especially in a socio-cultural context marked by educational underachievement and lack of engagement with formal learning processes. It has also sought to consider the circumstances in which teachers will not only have a high commitment to teach in settings often marked by failure and discouragement, but also how they will be best equipped with the curricular content and methodologies most suited to successful outcomes, together with strategies to address the needs of those whose progress is impaired.

In pursuit of its aims, this study sought explicitly to address the ambitious agenda of applying psychology on a large scale to endemic social and educational problems with a view to laying a foundation for major, intergenerational change in disadvantaged populations. To achieve such a vision it was necessary not only to design effective interventions in terms of basic content and method, but also to manage the processes of large-scale educational change and to sustain and develop these processes over a long period.

There would have been an appealing simplicity in tackling an alternative research agenda – one that avoided the many pitfalls of large samples, of vast individual assessment programmes that at various times looked as if they might be unmanageable and of the messiness of multiple-component interventions with their associated problems of separating the effects of different variables and ensuring fidelity in delivery. The extensive world of literacy research is replete with more straightforward research choices in discrete and manageable areas, and West Dunbartonshire Council could have discharged its responsibilities by conducting a simple early intervention study on a less grand scale, and with a less formal commitment to rigorous research methodology. Instead, the arena of real world

research was embraced in one of its untidiest settings; the agenda of critical psychology was espoused in the selection of a research programme designed to address crucial areas of human need; the thorny issues surrounding the concept of values in science had to be considered; and ways had to be found of incorporating challenging – at times almost mercurial – concepts such as ‘vision’ and ‘declaration’ into the overall equation.

If, in meeting its aim to design, implement and evaluate an intervention for raising achievement and eradicating illiteracy, this study has contributed to the well-being of some of the most vulnerable children and young people in society, then perhaps it will have taken forward to a small degree the vision embraced in the Foreword:

‘Psychology has the potential to help bring about a significantly better world, in keeping with its ethical mandate to promote human welfare. Yet too often we settle for too little.’ (Prilleltensky & Fox, 1997, p. 4).

It is with the commitment to values in science, and to psychology as a force for positive change in society, especially amongst its most needy members, that this epilogue concludes and this work ends.

References

- Barber, T. (1976). *Pitfalls in Human Research*. Oxford: Pergamon.
- Briggs, S., MacKay, T. & Miller, S. (1995). The Edinbarnet playground project: changing aggressive behaviour through structured intervention. *Educational Psychology in Practice*, 11, 37-44.
- Brown, G. (2007). *Britain's Everyday Heroes: The Making of the Good Society*. Edinburgh: Mainstream.
- Burkard, T. (2006). *A World First for West Dunbartonshire: The Elimination of Reading Failure*. London: Centre for Policy Studies.
- Clay, M. (1979b). *The Early Detection of Reading Difficulties* (3rd edition). Portsmouth, NH: Heinemann.
- Clay, M. (1991). *Becoming Literate*. Portsmouth, NH: Heinemann.
- Clay, M. (1993a). *Reading Recovery: A Guidebook for Teachers in Training*. Auckland, NZ: Heinemann.
- Cowling, K. & Cowling, H. (1993). *Toe By Toe: A Highly Structured Multi-Sensory Reading Manual for Teachers and Parents*. Baildon, West Yorks: Author.
- France, N. (1978). *The Primary Reading Test*. Windsor, Berks: NFER.
- France, N. (1981). *The Primary Reading Test (Levels 1 & 2): Teacher's Guide (Revised Edition)*. Windsor, Berks: NFER.
- Fraser, H., MacDougall, A., Pirrie, A. & Croxford, L. (2001). *National Evaluation of the Early Intervention Programme*. Final report to the Scottish Executive Education Department. University of Edinburgh.
- Freire, P. (1994). *The Pedagogy of Hope: Re-living Pedagogy of the Oppressed*. New York: Continuum.
- Freire, P. (2000). *Pedagogy of the Oppressed*. 30th Anniversary Edition. New York: Continuum.
- Fullan, M. (2001). *The New Meaning of Educational Change*. Third edition. London: Routledge Falmer.
- Glass, G., McGaw, B. & Smith, M. (1981). *Meta-Analysis in Social Research*. Beverley Hills, CA: Sage.
- Iverson, S. & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery programme. *Journal of Educational Psychology*, 85, 1, 112-126.
- Levine, K. (1986). *The Social Context of Literacy*. London: Routledge & Kegan Paul.
- Lloyd, S. (1992). *The Phonics Handbook*. Chigwell, Essex: Jolly Learning.
- MacKay, T. (1995a). Reading failure in an area of multiple social disadvantage: response of a psychological service to a school's priorities. In The Scottish Office Education Department, *Matching Service Delivery to Client Needs: Quality Assurance in Psychological Services* (pp. 210-233). Edinburgh: The Scottish Office Education Department.
- MacKay, T. (1999a). Baseline assessment in Scotland. *Journal of Research in Reading*, 22, 1, 81-88.
- MacKay, T. (1999b). Can endemic reading failure in socially disadvantaged children be successfully tackled? *Educational and Child Psychology*, 16, 1, 22-29.

- MacKay, T. (2006). *The West Dunbartonshire Literacy Initiative: The Design, Implementation and Evaluation of an Intervention Strategy to Raise Achievement and Eradicate Illiteracy*. Phase I Research Report. Dumbarton: West Dunbartonshire Council. ISBN: 0-906938-12-0
- MacKay, T. & Briggs, S. (1994). 'Fair Play': the primary school playground. *Educational Psychology in Scotland*, 2, 25-31.
- MacKay, T. & Watson, K. (1996). Literacy, social disadvantage and early intervention: a strategy to enhance reading in primary school. *Final report to Scottish Office Education and Industry Department*.
- MacKay, T. & Watson, K. (1999). Literacy, social disadvantage and early intervention: enhancing reading achievement in primary school. *Educational and Child Psychology*, 16, 1, 30-36.
- McLeod, J. & Anderson, J. (1972). *The Gapadol Reading Comprehension Test*. Port Melbourne, Australia: Heinemann.
- Meek, M. (1983). *Achieving Literacy: Longitudinal Studies of Adolescents Learning to Read*. London: Routledge & Kegan Paul.
- Neale, M. (1989). *The Neale Analysis of Reading Ability: Revised British Edition*. Windsor, Berks: NFER-Nelson.
- Organisation for Economic Cooperation and Development (2000). *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*. Paris: Author.
- Prilleltensky, I. & Fox, D. (1997). Introducing critical psychology: values, assumptions and the status quo. In D. Fox & I. Prilleltensky (Eds), *Critical Psychology: An Introduction* (pp. 3-20). London: Sage.
- Prilleltensky, I. & Nelson, G. (1997). Community psychology: reclaiming social justice. In D. Fox & I. Prilleltensky (Eds), *Critical Psychology: An Introduction* (pp. 166-184). London: Sage.
- Robson, C. (1993). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford: Blackwell.
- Scottish Executive Education Department (2003). *Report of the Sixth Survey of English Language (2001)*. Assessment of Achievement Programme. Edinburgh: SEED.
- Scottish Office Education and Industry Department (1996). *AAP English 1995: A Summary Report of the Fourth Survey*. Assessment of Achievement Programme. Edinburgh: SOEID.
- Scottish Office Education and Industry Department (1998). *Baseline Assessment in Scotland: Pilot Procedures*. Edinburgh: SOEID.
- The Basic Skills Agency (2001). *Adults' Basic Skills*. London: Author.
- Topping, K. & Lindsay, G. (1992). Paired reading: a review of the literature. *Research Papers in Education*, 7, 1-50.
- Wells, A. (1998). *What Works in Secondary Schools? Catching up with Basic Skills*. London: Basic Skills Agency.

Appendix 1

The Structure of Scottish Education

Statutory school age in Scotland is 5-16 years. Most children spend two years in pre-school education at nursery schools or other pre-five establishments from age 3, entering primary school at age 5 and secondary school at age 12. The system may be illustrated as follows:

Nursery school

First nursery school year	–	Age 3
The ‘pre-school year’	–	Age 4

Primary school

Primary 1 (P1)	–	Age 5
Primary 2 (P2)	–	Age 6
Primary 3 (P3)	–	Age 7
Primary 4 (P4)	–	Age 8
Primary 5 (P5)	–	Age 9
Primary 6 (P6)	–	Age 10
Primary 7 (P7)	–	Age 11

Secondary school

Secondary 1 (S1)	–	Age 12
Secondary 2 (S2)	–	Age 13
Secondary 3 (S3)	–	Age 14
Secondary 4 (S4)	–	Age 15
Secondary 5 (S5)	–	Age 16
Secondary 6 (S6)	–	Age 17

Appendix 2

Contents of the Phase 1 Research Report

<i>Abstract</i>	<i>i</i>
<i>Foreword</i>	<i>ii</i>
Summary	1
Chapter 1 Introduction and overview	3
Chapter 2 The application of psychology to human welfare	17
Chapter 3 The research context	27
Chapter 4 The context of socio-economic disadvantage	31
Chapter 5 The preparatory studies	44
Chapter 6 Intervention in literacy: an overview	51
Chapter 7 Baseline assessment	62
Chapter 8 Study 1 – the main study: rationale and design	78
Chapter 9 Study 1 – the main study: results	94
Chapter 10 Developing a phonics intervention: analytic or synthetic phonics?	119
Chapter 11 Study 2 – the synthetic phonics study	126
Chapter 12 Study 3 – the attitudes study	142
Chapter 13 Study 4 – the declaration study: rationale and design	149
Chapter 14 Study 4 – the declaration study: results	165
Chapter 15 Study 5 – individual support: the need for strategies beyond early intervention	176
Chapter 16 The West Dunbartonshire Literacy Initiative in action <i>Kathy Morrison</i>	189
Chapter 17 Conclusions	196
Epilogue	212
References	214
Appendices	231

1 The Structure of Scottish Education; 2 Baseline Assessment; 3 Synthetic Phonics Study – Spelling Test; 4 Jolly Phonics Questionnaire; 5 Toe By Toe – Sample Pages; 6 Impact – The Response of the Media

This book should be cited as:

MacKay, T.A.W.N. (2006). *The West Dunbartonshire Literacy Initiative: The Design, Implementation and Evaluation of an Intervention Strategy to Raise Achievement and Eradicate Illiteracy*. Phase I Research Report. Dumbarton: West Dunbartonshire Council. ISBN: 0-906938-12-0

Please note: This book may not currently be available in hard copy with its full covers and the content of some of its appendices. Any queries regarding this electronic copy should be addressed to the author at: criticalsolutions@btinternet.com

The West Dunbartonshire Literacy Initiative

**The Design, Implementation and Evaluation of an
Intervention Strategy to Raise Achievement
and Eradicate Illiteracy**

Phase 1 Research Report

Tommy MacKay

March 2006

Abstract

Objectives: The aim of this study was to design, implement and evaluate a multiple-component intervention to address underachievement and illiteracy in West Dunbartonshire, taking full account of educational change processes in the context of real world research.

Method: A main study and four supporting studies were conducted. The main study involved the design and implementation through six years of a multiple-component intervention in 58 nurseries and primaries, using a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts of the same age. Children in the early stages (N = 3,000+ annually) were individually assessed on a baseline assessment designed for the study, while older pupils (N = 3,000+ annually) took group tests. The synthetic phonics study used a quasi-experimental design to compare two phonics programmes in 18 schools. The attitudes study was a long-term follow up of 24 children from an earlier randomised control trial. The declaration study designed, implemented and evaluated a novel strategy in 12 nurseries and primaries in another education authority (N = 565), using a quasi-experimental design. It served the purpose of informing aspects of West Dunbartonshire's intervention. The individual support study was a quasi-experimental study in secondary school (N = 24), followed by extension into 35 primaries.

Results: In the main study, comparison of cohorts showed year-on-year gains on all tests and across all age groups, with indications of sustained post-intervention gains in later years. In each of the four supporting studies gains were found for the experimentals, pointing to benefits in the use of synthetic versus traditional phonics, in changing attitudes to reading, in making declarations of future reading achievement and in the use of intensive individual support.

Conclusions: The interventions reported in this study have resulted in raised achievement, have addressed illiteracy in areas of socio-economic disadvantage and have developed a foundation for planning intergenerational change in attainment levels.

Foreword

The West Dunbartonshire Literacy Initiative is unique. Its aim of raising achievement and eradicating illiteracy is based not only on an educational imperative but also on a total commitment to psychological research. At every stage this commitment has informed its design, implementation and evaluation. This Phase 1 Report covers the first six years or so of the project, from its commencement in 1997. Throughout that period many reports and briefings have been provided to inform Councillors, education staff, parents and others of the progress of the intervention. The purpose of the present work is to provide a full research report that will make the results of West Dunbartonshire's initiative available to the wider academic and scientific community.

As a research study the initiative addresses an ambitious agenda. It aims to apply psychology on a large scale to endemic social and educational problems with a view to laying a foundation for major, intergenerational change in disadvantaged populations. Its objectives and methodology are driven by a belief that:

'Psychology has the potential to help bring about a significantly better world, in keeping with its ethical mandate to promote human welfare. Yet too often we settle for too little' (Prilleltensky & Fox, 1997, p. 4).

In pursuit of the above aims, the academic comfort zone of uni-dimensional, randomised control trials in tightly organised experiments is for the most part left behind for an uncertain arena of real world research – an unpredictable world of action research in dynamic settings where researchers have limited control. Applying psychology in this way involves embracing the issue of how to move effectively from short-term, small-scale experiments with a single focus to long-term, multiple-level interventions with whole populations.

Even in the most controlled experiments the results are affected by a number of often unmeasurable, and more often unrecognised, factors. These become even more apparent in a large-scale study of this kind – factors such as the inertia built into the process of educational change on the one hand, and the releasing potential of vision and commitment on the other. There is therefore explicit acknowledgement in this study of the uncertain effects of 'context variables' – vision, profile, ownership, commitment and declaration – and the methodology includes studies relating to the significance of attitudes and expectations.

In academic terms, this study recognises a developing context within psychological enquiry and practice: the rise of qualitative research methods; the increasing focus on real world research and on issues relating to everyday life; the emergence of 'critical psychology'; the place of values in science; and a wider and more centrally important role for psychology and for educational interventions.

While the report focuses on the application of research-based methods to a large population, and the formal analysis of statistics relating to many thousands of children and young people, the ultimate aim of the study is to achieve meaningful and positive change in individual lives. This aim is crystallised in the statement made at one of the dissemination conferences by Kathleen Duncan, a pupil at Braidfield High School, Clydebank:

'When all this started I couldn't read. I was a failure. Now I have a cupboardful of books at home. My favourite authors are Roald Dahl and J.K. Rowlings. Now I am a success.'

It is statements like this, echoed many times by young people on the project, that highlight the place of values in science, and of an applied psychology agenda that addresses the needs of the most disadvantaged and vulnerable in society.

The aim of eradicating illiteracy is not only an ambitious one but one that has far-reaching implications. Each year in the UK over 100,000 young people leave school 'functionally illiterate' (Organisation for Economic Co-operation and Development, 2000; The Basic Skills Agency, 2001). In the main population addressed in this study the problem is endemic. The final achievement of this aim is something which it is the stated ambition of the second phase of this project to attain. The first phase of the study as reported here lays a solid foundation for its accomplishment.

Contents

<i>Abstract</i>	<i>i</i>
<i>Foreword</i>	<i>ii</i>
Summary	1
Chapter1 Introduction and overview	3
Chapter2 The application of psychology to human welfare	17
Chapter3 The research context	27
Chapter4 The context of socio-economic disadvantage	31
Chapter5 The preparatory studies	44
Chapter6 Intervention in literacy: an overview	51
Chapter7 Baseline assessment	62
Chapter8 Study 1 – the main study: rationale and design	78
Chapter9 Study 1 – the main study: results	94
Chapter10 Developing a phonics intervention: analytic or synthetic phonics?	119
Chapter11 Study 2 – the synthetic phonics study	126
Chapter12 Study 3 – the attitudes study	142
Chapter13 Study 4 – the declaration study: rationale and design	149
Chapter14 Study 4 – the declaration study: results	165
Chapter15 Study 5 – individual support: the need for strategies beyond early intervention	176
Chapter16 The West Dunbartonshire Literacy Initiative in action <i>Kathy Morrison</i>	189
Chapter17 Conclusions	196
Epilogue	212
References	214
Appendices	231

Improving Children's Literacy in Areas of Socio-Economic Disadvantage

**The design and evaluation of a strategy
to address underachievement and illiteracy**

Summary

Objectives

Socio-economic disadvantage accounts overwhelmingly for the variance in reading achievement across populations, and illiteracy and underachievement in disadvantaged children are endemic. Short-term interventions with small samples have resulted in increased reading test scores, but comprehensive, sustainable strategies for dealing with large populations are lacking. The aim of this study was to design, to implement and to evaluate the effects of, a multiple-component intervention to address underachievement and illiteracy in areas of socio-economic disadvantage, taking full account of the factors affecting educational change in the context of real world research.

Methods

A main study and four supporting studies were conducted. The main study involved the design and implementation of a multiple-component literacy intervention in all pre-school establishments (N = 23) and primary schools (N = 35) in Scotland's second most disadvantaged education authority. (The structure of pre-school, primary and secondary education in Scotland is shown in Appendix 1.) The study used a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts of the same age. The sample comprised all children in the pre-school year and in Primary 1 and Primary 2 classes (N = between 3,000 and 4,000 per year).

A baseline assessment covering concepts of print, phonological awareness and early reading skills was designed, and was used as a baseline prior to the start of the intervention, and for evaluation in each year of the intervention. All children were tested individually. This allowed comparisons of cohorts of children at the same level each year, as well as the opportunity to trace the progress of individual pupils. Group reading tests were also conducted with all pupils in Primaries 3, 4 and 7 to allow evaluation of the progress of pupils who had moved on from the stages at which the programme was operating. The multiple-component programme consisted of 10 strands, each drawn from existing research on literacy intervention.

The four subsidiary studies were conducted to assess three developing aspects of the 10 strands of intervention, namely: 'a strong and structured phonics emphasis' (one subsidiary study); 'changing attitudes, values and expectations' (two subsidiary studies); and 'identification and support for children who are failing' (one subsidiary study).

The phonics study had for its sample Primary 1 classes in 18 primary schools (9 experimentals, 9 controls). A 'synthetic phonics' programme was introduced for the experimentals, while the controls had a traditional 'analytic phonics' programme. The first study of attitudes and expectations was a long-term follow up of 24 children who had participated in an earlier intervention. The second involved the development of new strategies that were implemented and evaluated in a second education authority, using a sample of pupils (N = 565) in the pre-school year and Primary 1 classes in 12 establishments (6 experimentals, 6 controls). The identification of and support for children who were failing involved a two-fold approach: first, a quasi-experimental study at secondary school involving an individual support programme (12 experimentals, 12 controls); second, the extension of the individual support programme into classes at the top end of primary school in 35 schools.

Results

In the main study, significant year-on-year gains on all baseline assessment tests and across all age groups were found through a six-year intervention period, when cohorts at each age level were compared with the cohorts at the same level in the previous year. Group tests at Primaries 3 and 4 indicated that these gains were being sustained at a reduced level after the intervention was completed. In the synthetic phonics study, significant gains for experimentals on non-word reading and on word reading were found, with indications of sustained improvements into the middle years of primary schooling. In the first study of attitudes and expectations, children whose reading ages had increased more than five years previously as a result of a randomised control trial based on attitude change were reassessed. The experimentals were still reading at significantly higher levels than the controls. The importance of attitudes and expectations was further supported by the second study. The experimentals showed significant gains on baseline assessment tests and also showed significant shifts towards more positive attitudes to reading. Finally, in the individual support study, the secondary school experimentals showed very significant gains, and this pattern was reflected when the programme was introduced across primary schools.

Conclusions

This study has pointed to the effectiveness of a multiple-component literacy intervention in enhancing reading achievement and addressing illiteracy in a large population of children and young people in areas of socio-economic disadvantage. It has also indicated the potential of the intervention as a foundation for planning intergenerational change in achievement levels.

Chapter 1

Introduction and Overview

Organisation and layout

The nature and scale of this study dictate the advantage of organising this research report in a different way from the traditional plan, in which the main elements normally constitute a literature review, methods, results, discussion and conclusions. For a comprehensive literacy intervention on a grand scale, with samples of many thousands, multiple components, a range of quite different but inter-related studies and a wide scope of enquiry covering several disparate fields of research, it is appropriate to unfold the investigation by means of a less orthodox presentation.

For this reason it may be helpful to provide advance notice of how the work has been organised, so that the key elements of the plan may be anticipated and each part may be seen in its relation to the whole. In particular, it may be stated that although all of the key features of a traditional research report are present, they are not to be found gathered in one place. Thus, the literature review for each of the main fields of enquiry is covered in the chapters relating to that field, and similarly accounts of methodology are to be found in the chapters in which the various studies are described. The chapters and their contents are organised as follows.

Chapter 1: Introduction and overview

This chapter follows – and provides an enlargement on – the Summary, which is designed to facilitate anticipation of the investigations that are to follow, together with a brief statement of their key findings. It develops the aims and rationale of the exercise and the scope of the literature that has to be reviewed. A more detailed description of the five studies, with an overview of their methodology and their results concludes the chapter.

Chapter 2: The application of psychology to human welfare

This chapter considers the large-scale application of psychology to human welfare under the four headings of: the scope of psychology and of educational psychology; the context of real-world research; the place of values in science; and the nature of educational change. It contains a brief account of literature relevant to this area.

Chapter 3: The research context

This chapter describes the background from which the whole initiative developed, together with an account of the key demographic and other characteristics of the settings in which the research was carried out.

Chapter 4: The context of socio-economic disadvantage

This chapter explores the concept of socio-disadvantage that is central to this initiative, and reviews literature in this field with reference to social justice, health and quality of life, educational outcomes and implications for literacy.

Chapter 5: The preparatory studies

This chapter outlines three studies previously published by the author and colleagues that informed the development of the current initiative – *Fair Play: the Primary School Playground*, the *Edinbarnet Reading Project* and the *Edinbarnet Early Reading Project*.

Chapter 6: Intervention in literacy: an overview

This is an overview of literature on literacy interventions under a variety of headings. Further coverage of literature in this area is to be found in the chapters that introduce or describe the five main studies – Chapters 8, 10, 12, 13 and 15. Literature specific to baseline assessment in literacy is discussed in Chapter 7.

Chapter 7: Baseline assessment

This chapter describes in detail the baseline assessment scheme designed for the main study, within the context of national developments in baseline assessment. The rationale for the structure and content of the scheme is outlined, with references to the supporting literature in this field.

Chapter 8: Study 1 – the main study: rationale and design

An account of the main study, with a description of the 10 strands of the intervention programme, set in the context of relevant literature. Pointers are provided here to the chapters on the subsidiary studies which develop further the literature for three of these strands. Design, methodology and key features of implementation of the main study are covered here.

Chapter 9: Study 1 – the main study: results

A step-by-step presentation of all the key results, with an account of the statistical analyses used.

Chapter 10: Developing a phonics intervention – analytic or synthetic phonics?

This chapter prepares the ground for the synthetic phonics study (Chapter 11). It describes the difference between analytic (traditional) phonic instruction and synthetic phonics, and reviews key literature, with particular reference to the Scottish context.

Chapter 11: Study 2 – the synthetic phonics study

A presentation of the design and results of this subsidiary study and of its relationship to the main study.

Chapter 12: Study 3 – the attitudes study

This chapter sets the subsidiary study on attitudes within the context of key literature, and describes the method and results of this investigation and its relationship to the main study.

Chapter 13: Study 4 – the declaration study: rationale and design

This chapter prepares the ground for the novel intervention on declaration (Chapter 14). It introduces a rationale for using the concept of declaration to support a literacy intervention, relating this to nine areas of mainstream psychology, with illustrative literature. The place of this study in relation to the main study is outlined, and the design of the study is described.

Chapter 14: Study 4 – the declaration study: results

A presentation of the results of this study and the analyses used.

Chapter 15: Study 5 – individual support: the need for strategies beyond early intervention

A description of the work that prepared the way for the extension of the initiative into the upper years of primary school and beyond that into the secondaries. The rationale for individual support is discussed in the light of relevant literature, and the design and results of this subsidiary study are presented, together with its relationship to the main study.

Chapter 16: The West Dunbartonshire Literacy Initiative in Action

This chapter, written by Kathy Morrison, project leader and head teacher early intervention, provides a practical account of the initiative in operational terms under the headings of support for pupils, assessment and monitoring progress, failing readers in the later primary stages, staff development, ethos and partnership with parents.

Chapter 16: Conclusions

This chapter builds on the discussion issues raised throughout the work as the various studies unfold, and concludes with an overview of the entire study and of its key findings.

The Appendices

Appendix 1 The Structure of Scottish Education

A page showing the way Scottish pre-school, primary and education is structured in terms of ages and stages.

Appendix 2 Baseline Assessment

A full copy of the baseline assessment scheme designed for the study.

Appendix 3 Synthetic Phonics Study – Spelling Test

A copy of the spelling test designed as part of the assessment procedures for this subsidiary study.

Appendix 4 Jolly Phonics Questionnaire

A copy of the questionnaire for Primary 1 and Primary 2 teachers on their experience of the programme used in the synthetic phonics study.

Appendix 5 Toe By Toe – Sample Pages

Examples of the work contained in the programme used in the individual support study.

Appendix 6 Impact – The Response of the Media

Photocopies of press coverage of the project in its various stages.

Aims and rationale of the study

Building on the foundations laid by a number of preparatory studies, outlined below, the aim of this research was to design, implement and evaluate a multiple-component literacy intervention for raising achievement and addressing illiteracy in a whole population in areas of significant socio-economic disadvantage. The principal focus of the study was the early years of education, from the pre-school year in nursery education to the end of the second year of primary schooling. This covered a sample whose ages ranged from under four years at the beginning of their pre-school year to just under seven years at the end of Primary 2. Additional studies were conducted with other age groups in the later primary years and into secondary school.

Many investigations, including the randomised control trial that preceded this research (MacKay, 1999b), have demonstrated that it is possible to raise children's reading scores in small samples, in single establishments, using single intervention strategies over a short-term period. This research represented the first phases of a long-term study addressing a much more complex challenge – tackling the educational impact of socio-economic disadvantage in a large sample, across many establishments, using multiple-component interventions over a long-term period. The aim, in short, was to lay the foundations for intergenerational change in an entire population.

The anticipation of sustainable change on this scale was supported by a clear rationale with two key elements. The first was to design an intervention strategy using the

existing and developing evidence base for enhancing literacy levels. By implementing this strategy primarily as an early intervention programme it was expected not only that the overall literacy levels of children in the early years of schooling would be raised, but also that the number of children experiencing reading failure would be significantly reduced. This planned reduction in the numbers failing as they entered the later primary years would create greater scope and economic feasibility for supporting these pupils with intensive individual help to overcome their difficulties. By following each of these children through with support until they had achieved functional literacy the problem of illiteracy in the school years would be eradicated.

The second element involved recognition of the importance of context as well as content in the design and implementation of large-scale educational interventions. If programmes do not take explicit account of the factors affecting educational change they are unlikely to be sustained effectively in the longer term, irrespective of the quality of their content. This study therefore incorporated a recognition of the process of change. The key context variables were articulated and utilised in the implementation of all aspects of the intervention to ensure that the delivery of the content was sustainable and successful.

Scope of the literature review

The literature most relevant to this study straddles three main evidence bases: the process of educational change, socio-economic disadvantage and literacy.

First, all interventions require understanding of the process of change, and of the factors affecting the progress of intervention other than the research strategies themselves. However, as noted above, this is of particular significance to educational change in the context of long-term studies involving every level of change agency, from council leaders and other policy-makers to individual classroom teachers and helpers. The key factors from the literature on educational change processes are therefore reviewed.

Second, it is necessary to consider the context of socio-economic disadvantage both in general terms and in specific reference to its parameters and significance for the UK and for Scotland. It is a context of central relevance to the population forming the main sample for this investigation. Not only is it of significance in respect of educational achievement and indeed of overall quality of life, but it is also an important factor affecting responsiveness to intervention strategies such as the involvement of parents in supporting children's literacy.

Third, it is necessary to review the vast body of literature on how children learn to read, the obstacles arising for many children in achieving this satisfactorily and the factors most relevant to success. This includes consideration of literacy and illiteracy in a UK context and of the interventions designed to raise reading levels. It also includes an overview of literature in a number of distinct fields relevant to the subsidiary studies – the fields of synthetic phonics, attitudes, expectations and individual support programmes, as well as the novel area of declaration introduced in this study.

The diversity of the areas of key relevance determines that the literature review for this work must be extensive, and the result is a discussion of several hundred sources. This extensiveness and diversity has required careful consideration of the approach to reviewing the literature that would best support the purposes of the study. It may be described in general as a selective, narrative review. Many considerations have guided decisions on the sources to be selected or excluded. For the most part, this has involved judgements regarding the studies that best define or illustrate the known parameters of established areas of enquiry, such as the main factors affecting educational change processes, the impact of socio-economic disadvantage on quality of life and educational attainment, or the essential elements in effective literacy interventions.

Often choices regarding inclusion of studies have been guided by their relevance to the specific context in which the current initiative has been carried out, namely, the context of socio-economic disadvantage within Scottish education authorities. This has dictated the need to include, for example, some 'grey' literature that is unavailable through published sources, such as the internal evaluations by Lothian Regional Council of their Pilton Project in the 1990s. At other times the choice has been to select literature that might best illustrate how an established area, such as self-efficacy, would inform a novel application to an unexplored area, as in the use of declaration to enhance young children's literacy.

Overview of the studies

The research reported here comprises five studies: the first six-year phase of the ongoing main study and four supporting studies, one of these relating to phonics, two to attitudes and expectations and one to individual support of pupils experiencing reading failure. A number of 'preparatory' studies preceded the commencement of this research, and reference to them is of particular relevance as they laid the foundations on which many of its key aspects were developed.

In addition, a baseline assessment scheme (MacKay, 1999a) was designed as part of this research. Its development is outlined in detail, together with an account of other assessment measures used.

The 'preparatory' studies

Three studies or groups of studies were of particular significance in helping to formulate several of the main features of the current research: first, a series of studies of playground behaviour (MacKay & Briggs, 1994; Briggs, MacKay & Miller, 1995); second, a study of attitudes and values among children with reading failure (MacKay, 1995a, 1999b); third, an early intervention literacy study (MacKay & Watson, 1996, 1999).

The playground studies were of importance in highlighting the complexities of the process of change in educational settings and in clarifying a number of key factors for effective interventions in this area. The implications of four such projects were

outlined by MacKay and Briggs (1994). One highly successful playground intervention, using a qualitative research design incorporating several quantitative measures, highlighted the significance of the socio-economic context in relation not only to behaviour but also to difficulties in learning (Briggs, MacKay & Miller, 1995). It also emphasised the importance of factors such as self-esteem in effecting positive change. The groupwork intervention paradigm it developed, with its focus on self-esteem, was adopted in the next study, as outlined below.

The study of attitudes and values moved the focus of a groupwork intervention from playground behaviour to reading failure (MacKay, 1995a, 1999b). This was a randomised control trial in which 24 pupils in Primary 4 and Primary 5 with severe levels of reading failure were matched in triads for age, cognitive ability and reading level and randomly allocated to two experimental groups and a control group. One experimental group was the subject of a 10-week intervention based on changing attitudes and values regarding education and the relevance of reading, while the other, in addition, followed a paired reading programme at home. Both experimental groups achieved significant gains in reading scores compared with controls, together with reported improvements in other areas such as behaviour in school. Measures of attitude change of experimentals versus controls were also significant. This study led to a focus on the importance of attitudes, values and expectations in the main study reported here. It also provides the relevant context for Study 3, 'the attitudes study'.

The above study of reading failure was also important as a preparation for an early intervention literacy study (MacKay & Watson, 1996, 1999). The success of the intervention at mid-primary school level using a single strategy (changing attitudes and values) led to examining the issue of tackling reading failure in the same establishment at school entry age using a wider range of strategies. The sample comprised an experimental group of two Primary 1 (P1) classes (N = 46) in one school and controls (N = 44) matched for age and socio-economic status (SES) in two other schools in the same neighbourhood. All participants were assessed using a pre-test procedure designed specifically for the study and including reading readiness, reading achievement and attitudinal factors. The intervention, which extended over a five month period, was multi-dimensional and utilised variables supported by research in the areas of curriculum and teaching methods, attitudinal factors and home support. Experimentals achieved significantly better post-test results than controls, particularly in phonological skills, and a number of other benefits were reported by the school. The project achieved all of its objectives and led to the introduction of strategies to modify the 'literacy environment' of P1 entrants into all of the pre-school provision in the area.

Each of these studies therefore contributed to the planning and design of the current intervention.

Study 1 ('The main study')

The main study comprised an early intervention programme for all pupils in the pre-school year, Primary 1 and Primary 2 in all of the pre-five establishments (N = 23) and primary schools (N = 35) in West Dunbartonshire. This provided a total sample each year of approximately between 3,000 and 4,000 pupils at these stages. After the first year, one third of this number left the early intervention programme annually as they progressed from Primary 2 to Primary 3, to be replaced by a similar number of new cases entering the programme in the pre-school year.

All of these pupils were individually tested at the end of each calendar year (November-December) using the baseline assessment scheme designed for the programme (MacKay, 1999a). Testing took approximately 20 minutes, and was conducted by classroom teachers and by early intervention teachers employed for the project. All testers were trained in the procedure. About 120 testers were required each year, and training was ongoing to ensure that any new staff joining the programme were trained before carrying out the assessments.

The key elements in the baseline assessment were concepts of print, phonological awareness (for example, rhyme detection and production) and early reading skills (for example, knowledge of letter sounds, blending and word reading). The baseline was carried out prior to the start of the intervention programme and the same assessment measure was used for evaluation of progress each year. This provided a measure of literacy skills for each of the three stages from pre-school to Primary 2 and allowed comparisons of cohorts of children at the same level each year, as well as yielding data charting the progress of individual pupils from year to year.

Further measures of reading ability were also obtained for every child in the higher primary stages, so that there would be a means of evaluating change in the population as the pupils on the programme moved through their primary school years. For this purpose the Norman France Reading Tests (France, 1978, 1981) were conducted in May each year for all pupils in Primaries 3, 4 and 7. These group reading tests were administered by class teachers. Although they lacked the sensitivity of the individually administered baseline assessment they nevertheless provided an overall measure of change in the cohorts at each of these stages.

In terms of content, a multiple-component intervention programme comprising '10 strands' was designed. These strands were drawn from the evidence base for literacy interventions, and recognised areas such as the importance of phonological awareness in the early stages, of a strong and structured emphasis on phonics and of attitudes, values and expectations. An early intervention team comprising a project leader (head teacher) and 10 teachers was appointed. This team constituted the key personnel both in supporting pre-school and primary staff in the delivery of the programme throughout all establishments and in providing additional classroom help.

From the start of the intervention the factors affecting educational change were articulated. A steering group was established and key context variables of 'vision, profile, ownership, commitment and declaration' were promoted as being of the

highest importance. It was considered essential that the project should be marked by these five key factors: by all involved with it having the vision to believe that extraordinary results could be achieved; by being presented at all times in the highest profile as something of great importance; by everyone from the leader of the Council to the parents and the children themselves identifying with it and owning it as their own project; by everyone giving a long-term commitment to making it work effectively; and by constant and bold declarations that this initiative would have outstanding success.

These concepts were constantly and deliberately mediated at all levels of the education authority – Councillors, educational directorate, quality assurance personnel, head teachers, class teachers, other school staff, the early intervention team – indeed all who had an involvement with the project. Fidelity of implementation was maintained through a programme of constant monitoring, assessment and training. Motivation was also maintained both by according the project the highest profile in major dissemination conferences and by the phased introduction of new developments to prevent the initiative from becoming stale or routine.

Study 2 ('The synthetic phonics study')

This study aimed to develop one of the 10 strands of intervention designed for the main study – ‘a strong and structured phonics emphasis’. It adopted a developing area of research and practice in phonics teaching, ‘synthetic’ phonics (starting with letter sounds and learning how to combine these to make words), as an alternative to the traditional or ‘analytic’ phonics approach normally employed in the teaching of reading (beginning at whole word level and breaking words down into letter sounds). The synthetic phonics method had shown good potential, but had not been developed in situations where socio-economic disadvantage was a significant background factor, or where initiatives to enhance the teaching of analytic phonics were also being conducted.

Eighteen primary schools were selected for the study (9 experimentals, 9 controls). This was a quasi-experimental study as random selection of establishments would not have been a practicable possibility. The major curriculum change required by the experimental schools was such that the project could only be expected to work by asking for volunteers. These nine volunteer schools were matched in pairs with the nine schools selected as representing the nearest controls in terms of socio-economic profiles and literacy attainment levels. A synthetic phonics programme was introduced to Primary 1 classes in the nine experimental schools. The initiative was supported by a major training programme for all staff involved in implementation, together with comprehensive support arrangements and regular feedback meetings. A range of quantitative and qualitative measures was used for assessment. This included not only the pre-post baseline assessments for the first year of the study but also follow-up evaluations using the same assessment measure as the sample moved through Primary 2, and group reading tests as they moved through Primaries 3 and 4.

Study 3 ('The attitudes study')

This study served to develop another of the 10 strands of intervention designed for the main study – ‘enhancing attitudes, values and expectations’. Of the 24 children who had taken part in the randomised control trial on reading failure (outlined above in the preparatory studies), 20 were traced to their secondary schools in 1999, five and a half years after the initial study had taken place. At the time of the original study they were around the middle years of their primary schooling, and at follow up they were around the middle years of secondary schooling. All were individually assessed for their level of literacy skills, allowing new experimental v. control comparisons to be made.

Study 4 ('The declaration study')

This study also served to develop the strand of intervention relating to ‘enhancing attitudes, values and expectations’. The study was carried out in 12 primary and nursery schools (six experimental, six control) in East Renfrewshire during Session 1999-2000. The aim of the study was to change children’s expectations regarding their achievement in literacy, and to assess the impact of such change on actual reading scores. A total of 565 pupils participated in the six experimental establishments – 320 at pre-school level and 245 in Primary 1, with 27 teachers plus school management involved in implementation. Schools were matched for SES and included establishments with high and low levels of disadvantage.

Staff were trained in a novel intervention strategy based on changing expectations through declarations by pupils regarding future achievement, and this was implemented daily throughout a period of approximately nine weeks. A systematic sample of 60 children, five from each experimental and control school, was assessed individually before and after the intervention, using the baseline assessment designed for the main study. The sample comprised each n th child from the register, where n equalled one-fifth of the number in the primary class or nursery pre-school year group. Pre-post measures of attitudes to reading were also obtained. These quantitative measures were supported by qualitative indicators obtained from children who were assessed individually and also from staff in relation to all of the participating pupils in the experimental establishments.

Study 5 ('The individual support study')

The individual support study (MacKay & McDonald, in preparation) recognised that two factors relating to the later years of schooling required to be addressed. First, there were many pupils in the upper primary years and in secondary school who were already experiencing reading failure and who were not going to benefit from a literacy intervention focused on the early years. Second, a strategy was needed for the identification and support of children who were still failing even after they had been through the early intervention programme.

This study was carried out in two phases. First, a quasi-experimental study was conducted in one West Dunbartonshire secondary school, with 24 pupils referred for

learning support because of low reading levels. Of these pupils, 12 controls were assigned to the normal learning support programme while 12 experimentals were enrolled in an intensive individual support programme. The programme selected as meeting the criteria determined by the researchers was 'Toe By Toe' (Cowling & Cowling, 1993). The allocation of cases to the two conditions was not random, as it had to be subject to the normal timetabling and other constraints encountered in a large secondary school. However, the two samples were matched as closely as possible. All experimentals received individual tuition for 20 minutes a day, using a programme that lasted approximately three months. Pre-post assessments were conducted at the start and finish of a 12-month period.

Following further piloting of the programme in secondary and primary schools in the authority, the second phase of the study involved the identification of pupils at upper primary level (mainly Primary 7) who were experiencing significant reading difficulties. These children were initially identified by staff in the 35 primary schools in the authority, and following individual testing 104 were selected from 32 schools as meeting support criteria.

Approximately 120 individual support workers were trained in the use of the programme by the author and the learning support teacher who delivered the secondary school intervention. These were drawn from a wide range of personnel – teachers, classroom assistants and volunteers. Monitoring and support structures were put in place to ensure effective implementation.

Overview of methodology

Table 1-1 provides an overview of the methodology of the studies in terms of design, sample, assessment measures and analysis.

Table 1-1 Overview of methodology

	Design	Sample	Assessment measures	Intervention	Analysis
Study 1: Main study	Long-term; a cross-lagged design in which pre-intervention population cohorts were controls for subsequent same age intervention cohorts	Total N=43,370 of whom: Individual assessments: 22,986 (pre-test 3,659; post-test: 6,997 pre-school; 7,897 P1; 8,092 P2) Group assessments: 20,384 (6,683 P3; 6,752 P4; 6,949 P7) in 58 establishments (23 pre-school, 35 primary)	Individual: Baseline assessment scheme (MacKay, 1999a, designed for study – Appendix 2) Group: Norman France Reading Tests (France, 1978, 1981). All assessments administered annually	A multiple-component literacy intervention with 10 strands; designed for study but embedded within the curriculum as determined by national and local authority guidance	Independent two-sample <i>t</i> tests with effect sizes calculated on the standard deviations of the pre-test cohorts
Study 2: Synthetic phonics	Quasi-experimental, triangulated by a range of qualitative measures	18 primary schools (9 experimental, 9 control). Total N=590 (315 experimental, 275 control) Targeted N=180 (90 experimental, 90 control)	Baseline assessment scheme (MacKay, 1999a); spelling test (designed for study – Appendix 3); qualitative measures (Appendix 4)	A synthetic phonics programme (based on Lloyd, 1992)	Total sample: independent two-sample <i>t</i> tests with effect sizes calculated on the standard deviations of the pre-test cohorts Targeted sample: related two-sample <i>t</i> tests
Study 3: Attitudes	A comparison study: long-term follow-up of randomised control trial, comparing original experimental and control groups after 5½ years	N=19 (11 experimental, 8 control) from sample of 24 in original RCT	Neale Analysis of Reading Ability, Revised British Edition (Neale, 1989)	Not applicable. Original intervention described in MacKay (1995a, 1999b)	Independent two-sample <i>t</i> tests, with effect sizes calculated on the standard deviations of the standard assessment measure
Study 4: Declaration	Quasi-experimental, triangulated by a range of qualitative measures	12 schools (nursery and primary, 6 experimental, 6 control) Total sample: N=565 (320 nursery, 245 primary) Targeted sample: N=60 (30 experimental, 30 control)	Baseline assessment scheme (MacKay, 1999a); attitudes test, designed for study; qualitative measures	All children in experimental schools and nurseries made declarations regarding future literacy achievements (as described in Chapter 13)	Baseline assessment: independent two-sample <i>t</i> tests, with effect sizes calculated as for main study Attitude change: chi-square test
Study 5: Individual support	A quasi-experimental study in secondary school, supported by a gains-score study in primary school	Secondary: one school, N=24 (12 experimental, 12 control) Primary: 35 schools, N=104	Secondary: Gapadol Reading Comprehension Test (McLeod & Anderson, 1972) Primary: Neale Analysis of Reading Ability, Revised British Edition (Neale, 1989)	An individual support package for developing basic reading skills (Cowling & Cowling, 1993)	Secondary: Independent two-sample <i>t</i> tests with effect sizes shown; effect sizes were calculated on the standard deviations of the standardised assessment measure Primary: inspection of individual gain scores

Overview of results

The preparatory studies had already provided a basis from which to plan the design of a long-term, population-wide intervention in literacy using multiple-component strategies. They had highlighted not only the extent of underachievement in literacy and difficulties in related areas in disadvantaged populations, but also the importance of variables other than the content of the curriculum, most particularly attitudes, values and expectations. They had also pointed to the value of a broad-based approach to early intervention.

In the main study, a consistent pattern of higher achievement levels was found for all baseline assessment tests across each of the pre-school, Primary 1 and Primary 2 cohorts. These enhanced results were obtained year-on-year throughout a six-year intervention period when each cohort was compared with the cohort at the same age level in the previous year. Children at all levels of achievement benefited from the intervention. The proportion of pupils obtaining high scores rose significantly, while at the other end of the scale those with very low scores reduced dramatically in numbers. The group reading tests at Primary 3 and Primary 4, although very much less sensitive than the baseline assessments in their scoring, indicated that the gains made as a result of early intervention were being reflected in reading scores in the years following the intervention. The practical effect of these improvements was very apparent. Indeed, the class teachers in Primary 3 noted that the new reading levels of children entering their classes since the programme began were challenging the delivery of the normal P3 curriculum, which was requiring to be re-appraised.

In the synthetic phonics study it was hypothesised that the baseline assessment tests that should show the effects of the programme were those requiring word attack skills, namely, the non-word reading test and the word reading test. It was on these two tests that significant gains for pupils in the nine experimental schools were found, together with overall improvements in reading performance on group tests at Primary 4. Extensive qualitative data were also obtained from the staff in these schools. This provided very strong support for the effectiveness of the programme, since teachers were virtually unanimous in asserting that their pupils were working at higher levels of skill than had ever been known before. This view was expressed so universally that the nine volunteer schools were joined within a year or so by virtually every other primary school, so that the area became, in effect, a 'synthetic phonics' authority.

The attitudes study and the declaration study strengthened the base for ensuring that addressing attitudes, values and expectations should be built into any literacy intervention programme as a variable which could affect outcomes but which was essentially separate from the literacy content of the programme itself. Both studies addressed this variable in its own right, while keeping the content of the reading curriculum constant. The attitudes study indicated that over five years after a brief intervention to raise literacy scores by changing attitudes and values, the experimentals were still reading at a significantly higher level than the controls, even though they had received no differential treatment during these intervening years.

Again, this study was of importance in planning the main long-term intervention with a view to intergenerational change.

The declaration study resulted not only in significantly higher scores on early literacy skills for the experimentals but also in significant shifts towards more positive attitudes and expectations regarding reading. As a study of children in their pre-school year and Primary 1 it was particularly relevant to informing a large-scale early intervention.

In the individual support study dramatic mean gains were made by the experimentals in their reading scores, while the controls progressed only at the expected rate – that is, they made considerably less than one year's gain over the 12-month period between pre-test and post-test. This result supported the pilot work carried out in preparation for the study, during which very high gain scores were reported for pupils receiving the intervention. This provided a firm basis for setting up the training and implementation arrangements by which the study was extended throughout the upper classes in the primary sector. Following the established success of the programme at secondary using quasi-experimental methods, the primary project was carried out without selecting further controls. A high level of gain scores was achieved across the sample during a five-month intervention period.

These results have consistently supported the two main objectives of the whole initiative – significantly raising the achievement levels of this entire population and providing a basis for addressing the problem of illiteracy.

Chapter 2

The Application of Psychology to Human Welfare

The large-scale application of psychology to human welfare, specifically within the educational arena, is considered here under four heads: the scope of psychology and of educational psychology; the context of real world research; the place of values in science; and the nature of educational change.

The scope of psychology and of educational psychology

The foundations of applied psychology are to be found mainly in a response to human problems, that is, in the application of psychology to individuals or situations where things have become wrong or disordered. This may be illustrated by reference to the origins of applied psychology in Britain in the main settings of health, education, work and the law (MacKay, 2001b). Psychologists were employed in these fields to deal respectively with disorders and disabilities, special educational needs, meeting the demands of war and the aftermath of crime – that is, to respond to the problem areas of life.

The reasons for this focus were both obvious and commendable. Funding and support for applied psychology were most likely to be found in dealing with these key areas of need, and they highlighted the demonstrable significance of psychology to the ‘human condition’ in terms of equity and relevance. However, at the same time this led to a narrow emphasis for much of applied psychology in relation both to its target population and to its methods of operation. The population of most interest tended to be one with psychological problems, and the method of approaching it was generally based on individual assessment and intervention. Thus, clinical psychology dealt with individuals with mental health problems; educational psychology dealt with children and young people with special needs; occupational psychology developed a focus on individual personnel selection, often specialising in identifying those who failed selection tests for the armed forces, and then for the wider employment context; and forensic psychologists assessed and treated individuals with problems of delinquent or criminal behaviour.

All of these interests continue to be necessary fields of psychological enquiry and practice. Nevertheless, the result became a narrowing of vision that may be summarised in three key concepts: prevention, breadth and positive focus. In terms of prevention, the response of applied psychology was reactive rather than proactive. The context of the problems was not being addressed and challenged (MacKay, 2000a; Prilleltensky & Nelson, 2000). The case for prevention is overwhelming (Albee & Gullotta, 1997). In social and economic terms, the cost of failure to foster human welfare is enormous, and falls heavily on the families concerned as well as on health, education and social services (see, for example, Knapp, Scott & Davies, 1999). Yet the focus of intervention has continued to be ‘ounces of prevention and pounds of cure’ (Ross, 1998).

In terms of breadth, dealing with the individual as the main focus of psychology was unsustainable and uneconomical, and had little overall impact. It was the sheer scale of the problem that rendered individual interventions ineffective as a basis for dealing with psychological and social problems. It was not so much a question of whether clinical or educational psychologists could answer the specific challenges raised about the effectiveness of their individual techniques (for example, Dawes, 1994). It was more a requirement to recognise that psychology would never have a major impact on society if preoccupied with individual interventions. The problems were too widespread, whether the incidence of mental health difficulties (for example, Kramer, 1992), of educational underachievement (for example, Cox, 2000) or of the recognised level of criminal behaviour in society. It is this context that determines that the priority for interventions must shift from a central focus on individual needs towards addressing institutions and structures.

In terms of a positive focus, it was necessary for applied psychology – whether clinical, educational, occupational, forensic or otherwise – to look beyond the traditional territory of disorders and difficulties to a much wider and more crucial role in society's affairs: that of promoting health and quality of life, fostering learning and raising achievement, enhancing work satisfaction and motivation and fostering harmonious communities (MacKay, 2001b).

It is the question of fostering learning and raising achievement that is of the most central interest to educational psychology. It has the scope to meet all of the aspirations indicated above. It is essentially preventative rather than reactive; it calls for systemic rather than individual interventions; and it promotes the development of positive outcomes for all, rather than responding solely to a problem context.

A turning point in establishing a wider and more systemic role for educational psychology was the publication of *Reconstructing Educational Psychology* (Gillham, 1978). The agenda for change was set out by a wide range of authors, and included a robust critique of reactive, individual models of working. Nevertheless, despite a constant focus on this change agenda in the educational psychology literature (Acklaw, 1990; Jensen, Malcolm, Phelps & Stoker, 2002; Stobie, 2002a, 2002b), the following 20 years or so were characterised not only by a continued emphasis on traditional roles (Boyle & MacKay, 1990; Lokke, Irvine, M'gadzah & Frederickson, 1997; MacKay, 2000b; MacKay & Boyle, 1994; MacKay & Vassie, 1998; Stobie, 1996) but by a clear signal from teachers and other service users that these were the roles they valued. For example, a survey of head teachers' opinions at the beginning of this period of change (Topping, 1978) indicated that they valued 'individual casework' rather than general school work, policy development or research. This preference was echoed in unambiguous terms several years later in a Scottish study of the perceptions of teachers in primary and special schools (O'Hagen & Swanson, 1983). New and wider consultancy roles were treated with suspicion while the traditional functions in relation to assessment of children with special needs were emphasised.

However, there have been indications that new roles for psychologists have been appreciated when they have been introduced in the context of a good 'conventional'

service. Evans and Wright (1987) examined teachers' perceptions of service delivery in the Surrey psychological service during a period of change between 1981 and 1985 when the service moved from a traditional, psychometric, referral-based approach to a model offering regular visits to schools, curriculum-based assessment, increased involvement in in-service training and educational policy-making and a system of writing reports to schools in terms of agreed aims and actions. Teacher ratings indicated that the service was viewed in a positive light both before and after these changes. At the same time, in terms of priorities the teachers' views had not changed. They still placed assessment of individual pupils with learning difficulties as their top priority, with considerably less interest in the new roles on offer.

Within a Scottish context in recent years there has been considerable evidence that a wider role for educational psychologists is increasingly welcomed by service users. A series of studies examining primary and secondary teachers' perceptions of the role of the psychologist in relation to learning difficulties has shown a clear development here (Boyle & MacKay, 1990, submitted; MacKay & Boyle, 1994, submitted). In the earlier studies, while roles such as consultancy, research and in-service training were valued, teachers continued to view the traditional activities of individual assessment and counselling as being of prime importance. Vassie and Watson (1990) reported similar findings in a study of more general service delivery to primary schools. However, in the most recent study, conducted in 2003 with 112 primary and 24 secondary schools, not only had there been a significant increase in the involvement of psychologists in new roles such as research and policy development, but the extent of involvement in these roles accounted for most of the variance in how highly schools valued the psychologist's contribution.

The current context of Scottish educational psychology may be described as representing a tension between continuity and change (Boyle & MacKay, submitted; MacKay, 2003b; Stobie, 2002a, 2002b). Nevertheless, the presumption in favour of change is a compelling one, and has considerable official backing. A research role for the educational psychologist has not only been promoted vigorously in the literature (Greig, 2001; MacKay, 1987, 1989, 1997, 1999d) but has also been endorsed by the Scottish Executive in its review of services as a core function (Scottish Executive, 2002). The review has highlighted a central remit for the profession in relation to the national priorities for education in Scotland, arising from the Standards in Scotland's Schools etc. Act 2000 – raising attainment, supporting teachers, enhancing school environments, promoting equality, work with parents to teach pupils self-respect and enhancing pupils' skill, attitudes and expectations. All of these priority areas are of central relevance to the current study.

The context of real world research

A survey of the views of over 800 UK psychologists regarding future trends using the Delphi technique (Haste, Hogan & Zachariou, 2001) pointed to two developments for psychology as a science. First, there would be an increased research emphasis on everyday life, quality of life and the whole person. Second, research would move increasingly from the laboratory to real world settings. These predictions represent a continuing trend in psychology, which has seen an expansion

of interest in qualitative research paradigms applied to the relatively uncontrolled and unpredictable arena of real world research.

Conducting real world research with practitioners in schools requires a recognition that work must be done in 'complex, messy, poorly controlled "field" settings' (Robson, 1993, p. x). As sample size increases from small numbers in a single school to whole populations in many diverse establishments, as co-workers increase from one or two dedicated research assistants to vast numbers of teachers and other staff, and as intervention methods increase from single to multiple-component strategies, so also is there a corresponding increase in the complexity, messiness and poor control that characterise these settings.

These features of real world research are prominent even where small samples form the basis for carefully controlled, short-term interventions. The preparatory investigations that laid the foundations for this study, in particular a randomised control trial (MacKay, 1995a, 1999b) and a quasi-experimental study (MacKay & Watson, 1999), exemplified the messy nature of real world research. In the latter case it was clearly articulated:

'In many ways the study reported here serves as an excellent example of such messiness. The planned groupwork intervention for the experimental subgroup could not proceed because the groupworker took ill; the home link teacher met with discouragement in her home support plan; the peer tutoring by older pupils in the Easter holidays did not get off the ground because of lack of response; one of the two researchers made a career change in the middle of the study; no sooner had the intervention begun than the school had to prepare for a merger with another school in the area; and in the middle of it all local government reorganisation took place, resulting among other things in the termination of funding and the necessity for a new grant application' (MacKay & Watson, 1999).

These difficult features do not constitute a reason for avoiding research in real-world settings. Rather, they represent a reason for researchers to acknowledge the special challenges in taking research from controlled, laboratory settings to complex naturalistic settings for which effective interventions must be designed, implemented and evaluated. Despite the difficulties encountered in the above study, the authors went on to note:

'Nevertheless, the intervention was successful, and the experimentals made greater gains in their early reading skills than the controls. It is important to highlight the difficulties which arose, because in the view of the researchers these "messy" features are characteristic of all real research which takes place in schools, yet these aspects are scarcely ever mentioned' (MacKay & Watson, 1999).

It is this context of real world research that informs all aspects of the present study. The comfort zone of uni-dimensional, randomised control trials in tightly organised experiments is left behind for the unpredictable world of action research in dynamic settings where researchers have limited control. It is an arena where researchers design studies, but where politicians, funders, managers and workers follow many

other often conflicting agendas; where mundane issues such as council policies, financial crises, breakdowns in communication, staff turnover and vacant posts conspire to undermine the most carefully laid plans for experimental rigour. It is a world in which high fidelity of implementation is often vitiated by staff illness, maternity leave or burnout, or by the competing demands of other initiatives; a world also in which the most stringent plans for experimental and control groups are subject to the political ethics of those whose remit is to do good, and to ensure equality of opportunity and equity of service provision.

It is also a world in which the researcher becomes increasingly aware of factors other than the substance of the intervention itself that seem to be important in determining outcomes – ethereal factors such as vision, profile, ownership, commitment and declaration. These concepts are difficult to identify, far less to measure. Yet they may define whether an intervention is dramatically effective or merely mediocre. Each contains inherent dangers for the researcher. Vision can only ever be first-hand, and never second-hand. It can be inspired and encouraged, but ultimately it must be the vision of those who see, and not of those others who advise and guide them. For this reason those with vision will think independently, and will aspire beyond the narrow boundaries dictated by researchers, as they adapt and improve the designs prescribed for them. Profile likewise is hazardous. The projects that are given the highest profile are characterised by excitement and vibrancy. But is it the intervention strategy or the profile that contributes most to the positive outcomes? So also with commitment. Staff with low commitment may complete records indicating the level of fidelity with which they have adhered to the research procedures, yet their records will reveal little of the dynamic process that leads to results. Those with high levels of commitment will complete the same records the same way, but leave the researcher often pondering whether they would achieve their results whatever the intervention strategy. As to ownership, it is only achieved at the cost of ensuring that control is with the practitioners and not the researcher. It ensures that the project will be pursued in the face of difficulties and obstacles; but those who have taken ownership of their work will adapt it and enhance it, and will develop aims and aspirations that do not always coincide with those of the researcher.

These mercurial factors – the *context variables* as opposed to the *content variables* – have been explicitly articulated throughout this study. Indeed, they have been fostered at every level of the intervention as part of the interactive context needed for ensuring that a large-scale, long-term programme will remain vibrant and viable throughout its lifespan.

Robson (1993) characterises real world research as: solving problems rather than just gaining knowledge; predicting effects rather than finding causes; looking for large effects rather than studying relationships between variables; developing and testing interventions and services rather than theories; and using multiple methods rather than single methods. All of these are features of the present enquiry.

The place of values in science

The focus of real world research on solving problems, testing interventions and looking for the kinds of large effects that make meaningful differences to people's lives goes hand in hand with an increasing debate on the place of values in science. It is a discussion that challenges the entire psychological research agenda, by asking questions about the priorities, aims and methods of research. The debate has many able exponents who are promoting a 'critical psychology' agenda and who are challenging the assumptions on which traditional psychological paradigms for research and intervention are based (Fox & Prilleltensky, 1997; Prilleltensky, 1994; Prilleltensky & Nelson, 1997, 2000; Rappaport, 2000). This is an arena marked by 'irony, tensions and contradictions' (Rappaport & Stewart, 1997), but essentially the discussion about values is an extension of the established field of psychology and ethics. It moves beyond traditional ethical issues regarding the conduct of psychologists and how they act in relation to their clients and their research participants, and raises wider questions about psychology, social justice and human welfare. The central argument is that psychology (or indeed science in general) is not and cannot be apolitical and value free. Prilleltensky and Fox (1997) crystallise the position in citing Pettifor (1996) in her presidential address to the Canadian Psychological Association:

'The discipline of psychology is moving beyond the myth of detached neutrality to discover virtue and to recognise politics as forces which determine ethical behaviour... The concept of personal individual ethics needs to be extended to the social and cultural environment, because the environment influences the ethical behaviour of psychologists and the quality of life of all citizens. A morally responsible perspective includes a political role for psychologists which encompasses strategies to shape that environment' (in Prilleltensky & Fox, 1997, p. 8).

The attempt to provide an agreed values framework to underpin the priorities of psychological interventions is not without significant challenges. In promoting the belief that 'fundamental human needs, values and rights must be met and upheld for a better and more just society to emerge', Prilleltensky and Nelson (1997) propose five core values that may be found generally agreeable among psychologists. These are: health, caring and compassion, self-determination and participation, human diversity and social justice. Perhaps the common foundation governing the ethics of social and psychological interventions has never been better expressed than by Brill (1962):

'We believe that every single human being is of equal and infinite value... This remains true however strange, unpleasant, or socially unacceptable that human being may be. It is true of helpless babyhood, lunacy and senility... it is true of the grossly subnormal as of the highly intelligent; it is as true of the... tramp as of the managing director'.

The recognition of social justice issues in governing the priorities of this research is not put forward as a political alternative to objective scientific enquiry. Nor is it proposed without reference to the implicit dangers in a value-driven psychological

agenda. MacKay (1999c, 2000a, 2000b, 2001b, 2002c) has argued that explicit acknowledgement of the place of values in science must recognise key principles that are often absent from critical and political constructions of psychological enquiry. First, research is not obliged to be 'real world'. If the place of pure psychological science were undermined, and if ethical values meant that human behaviour could not be studied simply for its own sake, psychology would be placed in a position of unique disadvantage among all academic disciplines. Second, if a campaign for social justice is not primarily to be that of the social and political activist but rather that of the psychologist, then it must be firmly grounded in the principles and methods of psychology and in its evidential base.

It is here that the critical psychology paradigm and the positivist-empiricist paradigm of mainstream psychology must surely coincide. If we can show, for example, that the distribution of physical and mental disabilities is most concentrated among the poor, but that it is not matched in the distribution of resources designed to meet those needs (MacKay, 1999c), we are using the normal tools of empirical enquiry and statistical analysis. Thus, our values lead us to enquire into areas where social justice issues arise, and our concepts and methods provide the distinctive tools for addressing these issues effectively. This approach to science and values has been illustrated in a similar way in relation to the structure of social justice in areas such as government spending, poverty and the council tax, the language of education, discrimination in public services and the distribution of public resources (MacKay, 1982, 1983, 1999c, 2000c, 2001a, 2003a, 2003c). Psychologists are almost uniquely placed to deal with the study of social justice, not as political campaigners but as scientists.

The aspiration to design strategies to contribute to shaping the social and educational environment of disadvantaged and vulnerable populations was the driving force underpinning this study. The place of values in determining the aims and objectives of the research, and in guiding its methodology and priorities for implementation, has been to the fore throughout. The issue of values and social justice has been made explicit and has been brought from the periphery to the centre of enquiry and intervention.

The nature of educational change

The whole concept of shaping the educational environment to the benefit of disadvantaged populations requires recognition of a wider question: the nature of educational change. This is not only a complex area but a subject of study in its own right, with a substantial evidence base (for example, Crawford, Kydd & Riches, 1997; Fullan, 1991, 1993, 1999, 2001; Hargreaves, 1994; Sarason, 1982). A failure to understand and plan effectively for the process of change in organisations is a primary cause of the widespread failure of school projects, and indeed of the general failure of educational reform (Fullan, 2001; Sarason, 1990).

Fullan (1993, 1999) has demonstrated the ease with which 'restructuring' occurs time and again in educational organisations, but the much greater difficulty of 'reculturing'. The former is done by fiat. It is the outcome of the mandates of policy-

makers. The latter refers to the process by which school staffs come to question and change their beliefs and habits. This latter task, a fundamental requirement for fruitful change, has proved to be a far more difficult undertaking than was previously believed. For example, Stigler and Hiebert's (1999) international study based on videotapes of mathematics teaching showed that despite the presumption of significant reform in this area in US schools, little evidence of any real reform was to be found. Some surface features had been changed, but the element most clearly requiring reform – the basic approach to teaching – had remained the same.

The pace of educational reform has increased dramatically in recent years, not only in the North American context studied most extensively by Fullan (2001), but also within the UK and in Scotland specifically (Humes, 2003). Large scale change has been envisaged not only for education in general, but also, and most especially, for schools in areas of socio-economic disadvantage. The Scottish Executive promised to focus support on those vulnerable young people who were in greatest danger of becoming permanently excluded (Scottish Executive, 1999). This included a target of reducing by one third the days lost every year through exclusion from school and truancy. Yet much of the research on the effects of educational reform is dispiriting. As Elmore (1995) has noted, it is very straightforward to produce countless examples of how educational practice might be reformed, but very difficult to find examples of large numbers of teachers actually engaging in these practices.

Indeed, the problem of reform and of educational change is even more challenging than simply recognising the gap between policy and plans on the one hand and classroom practice on the other. Frequently, the results of reforms are not merely neutral but negative. Staff suffer from the effects of overload and confusion. Fullan (2001) cites a survey by Hatch (2000) of schools in California and Texas. Some establishments were working with nine or more different improvement programmes simultaneously. The results were that frustration and anger had never been higher. Hatch's conclusion was that

'...rather than contributing to substantial improvements, adopting improvement programmes may also add to the endless cycle of initiatives that seem to sap the strength and spirit of schools and their communities' (in Fullan, 2001, p. 22).

In his extensive investigation of 'the meaning of educational change', Fullan (2001) summarises the change process in terms of three broad phases now generally recognised by researchers. These are described in a variety of ways but essentially comprise: Phase I – initiation (the process that leads up to and includes a decision to adopt or proceed with a change); Phase II – implementation (the first experiences of attempting to put an idea or reform into practice); Phase III – institutionalisation (the process by which the change, if effective, is built in as an ongoing part of the system). Since the implementation phase will usually occupy the first two or three years of use, it is clear that an effective and lasting change process is unlikely to be accomplished quickly. Fullan's studies lead him to the view that:

'The total time frame from initiation to institutionalisation is lengthy; even moderately complex changes take from three to five years, while larger scale efforts

can take five to 10 years with sustaining improvements still being problematic. The single most important idea... is that *change is a process, not an event* – a lesson learned the hard way by those who put all their energies into thinking through what would have to happen beyond that point.’ (2001, p. 52).

It is clear from a study of change processes in an educational context that the whole area is beset with major difficulties – innovation overload on schools and teachers because of the pace of reform; the imposition of policy and legislation leading to restructuring rather than reculturing; and lack of recognition of the length of time during which an effective change process must be planned and managed. These difficulties are often a function of governmental and administrative structures. Governments must be seen to be reforming and legislating, but their lifespan before the next election will seldom exceed the initiation and very early implementation phases of any reform. Educational administrations in turn will introduce new policies, procedures and initiatives, but will often lack the mechanisms and resources to embed these in the culture of schools in a meaningful way. There is also a danger for researchers, who will initiate successful models of reform for which they themselves have had time to develop ownership and commitment, but without allowing for the processes leading to ownership and commitment in those responsible for implementation. All real change involves loss, anxiety and struggle (Marris, 1975), and requires a decision to leave behind the comfort of established and familiar ways.

Despite these difficulties, and the general inertia attendant on educational change, there is evidence that school-wide and authority-wide educational innovation can be successful. The American Institutes of Research (1999) reviewed the research supporting a large number of educational initiatives to find evidence of impact. The overall results were in many ways as disappointing as the above review might predict, in that two-thirds of the programmes showed marginal or no evidence to support them. Nevertheless, three of the models were viewed as demonstrating strong evidence of positive effects.

One of these, ‘Success for All’ (Slavin & Madden, 1998, 2000) has been used by schools extensively in the United States and has also been adopted by some UK schools. The programme has wide implications for school reform, but its specific literacy focus and explicit recognition of the nature of organisational change have particular relevance to this study. Its main components as described by the American Institutes of Research (1999) include: a reading curriculum designed to give at least 90 minutes of daily instruction in classes; frequent assessment of student progress; one-to-one reading tutors; an early learning programme at nursery level emphasising language development and reading; a family support team to encourage parental involvement; facilitators to assist teachers in the implementation process; and training in key aspects of the programme.

Within an educational psychology context, MacKay (1995a) proposed that the aims of project work in schools should be three-fold: first, direct effects – changes that occur as a direct result of the intervention; second, training effects – changes that become incorporated into the school’s methods and policies; third, generalised

effects – approaches that can be extended to other situations within the school or to other schools. Each of these aims required a plan for educational change, including follow up and review. Otherwise projects would quickly become irrelevant and their impact would be lost.

This study sought from the outset to take account of these lessons regarding the nature of educational change. The content of interventions, and more particularly the implementation strategies, were designed to maximise the potential for effective and enduring change. The key issue was how to operationalise the existing and developing corpus of psychological knowledge about literacy and illiteracy into a long-term intervention that would work effectively across an entire education authority. There was therefore a crucial emphasis not only on initiation but also on implementation and institutionalisation, with built-in structures for monitoring, review, support and training.

SUMMARY

This chapter has considered issues relating to the large-scale application of psychology to human welfare under the headings of the scope of psychology and of educational psychology, the context of real world research, the place of values in science and the nature of educational change. It argues that there is a central remit for educational psychology in raising attainment, but that the real world research context in which this will be achieved is one of ‘complex, messy, poorly controlled “field” settings’. It also argues for a central place for values in science, and sets a research agenda for the current study based on the aspiration to enhance the social and educational environment of disadvantaged and vulnerable populations. It also recognises from the outset that the large-scale application of psychology to real world settings must take account of the complex processes of educational and organisational change.

Chapter 3

The Research Context

The studies described here represent the initiation and a major part of the implementation of a wide-ranging and long-term literacy intervention. The overall aims of the intervention were to raise achievement levels and to eradicate illiteracy in an entire population of children and young people in West Dunbartonshire, an area characterised by high levels of socio-economic disadvantage. This chapter provides a background to the studies and a description of the research settings in which they were carried out.

The background to the studies

In 1996 the Director of Education in West Dunbartonshire considered a paper submitted by the author of this report, and research consultant to the project, entitled *Transforming Reading Achievement for All Children – A Vision for an Education Authority*. It was a time of significant change for local government in Scotland following the reorganisation brought about by the Local Government etc. (Scotland) Act 1994. It was also a time of significant opportunity. The 32 new Councils were, for the most part, smaller than the 12 regional authorities that had preceded them. They had extensive powers in governing their own affairs at local level, and with many new personnel in newly-created posts, together with a fresh start in a new context, the time was ripe for new thinking and bold initiatives.

It was in this context that the visionary proposal raised in the paper was put forward. It embraced all of the concepts and ideals endorsed so far in the earlier chapters of this work: the potential of applied psychology; the opportunity for effecting large-scale educational change; the challenge of real world research; the place of values in science, and the contribution it could make to the lives of people who are vulnerable and disadvantaged; the power of transforming leadership; and the need for vision, profile, ownership, commitment and declaration. The project was endorsed by the educational directorate, subject to obtaining research funding from external sources, and the vision was embraced by the Council. Shortly afterwards, the Scottish Executive made funding available to education authorities for carrying out early intervention programmes in literacy, and a successful bid was prepared under that scheme. Thus the *West Dunbartonshire Literacy Initiative* was born.

The commitment to a visionary strategy as forming the driving force for large-scale educational change was beyond question. Vision was the keynote of the paper in its title and in its content. Its stated purpose was to outline ‘a proposal for achieving something that has not been done before... It is a vision for transforming reading standards for all children in all schools throughout the education authority’.

West Dunbartonshire was ideally suited to a visionary programme of this kind for two reasons. First, it was the second poorest Council area in Scotland (following the

City of Glasgow), and indeed some of its economic statistics were in fact the poorest in the country. Second, the author and colleagues had already conducted a number of relevant research studies in schools forming part of the new authority ('the preparatory studies').

A very high level of commitment was expected from the Council, from the educational directorate and from the researcher. It was anticipated that the Council and directorate would: adopt the proposal as a vision which they intended to achieve; give this vision high public profile as Council policy; and collaborate in securing the commitment of every member of the education department. In summary, the Council was making 'a bold declaration that it is committed to extraordinary achievement'. In return the commitment from the researcher would be: to collaborate in setting up an appropriate research design; to ensure that the strategy was made operational in establishments throughout the authority; to provide technical advice whenever required; to participate in a training strategy for key personnel; to keep the project on target; to design evaluation procedures; and to remain with the project until completion.

Literacy was selected as the focus of the initiative for clear reasons. It was the single best indicator and predictor of educational achievement or failure in primary school children, and was the basis on which adequate access to other subjects in the curriculum at secondary school would be facilitated. It was also the most prominent aspect of educational underachievement in areas of socio-economic disadvantage, and was strongly linked with many other problem areas such as behavioural difficulties during school years and crime and delinquency levels in later life.

The underlying rationale and method outlined in the original paper have continued to guide the project throughout its development to date. In terms of rationale, it was proposed that there is an existing research base that defines most of what needs to be known about achieving or failing to achieve literacy; that the real challenge was to 'operationalise' this research base and apply it effectively throughout an entire education authority; and that the only context in which this could be attained was one that embraced vision and a total commitment to achieving extraordinary results. In terms of method, all aspects of the project would be research based both in content, drawing from the evidence base on enhancing reading skills, and in process, drawing from a wider evidence base including the studies of organisational change. Thus the paper delineated the requirements for monitoring, support and training, and the key 'context variables' referred to in this work – vision, profile, ownership, commitment and declaration.

It was in this context of the West Dunbartonshire Literacy Initiative that the main study, the synthetic phonics study, the attitudes study and the individual support study were carried out. The declaration study was developed in 1999 following an approach from East Renfrewshire Council for research to support their work on early intervention in literacy. Their chosen area was teacher and pupil expectations, and they wanted a study that would develop its own unique methodology. In response, 'declaration' was designed as a new tool for raising achievement in literacy. In turn,

the work on declaration informed the development of the main initiative, and its rationale, methodology and results are therefore provided in full in this report.

The research setting

West Dunbartonshire is a mixed urban/rural area situated to the west of Glasgow on the north or the River Clyde. The greater part of its population clusters in and around its two principal towns, Clydebank in the east and Dumbarton in the west. As the second poorest Council area in Scotland, it is marked by significant levels of socio-economic disadvantage. Its territory marks the heartland of the old Dunbartonshire County Council, which disappeared with the previous reorganisation in 1974 to become almost co-terminous with the Dumbarton District of Strathclyde Region. Already one of the poorer areas in the country at that time, it lost its only two significant pockets of economic prosperity in 1996 – the area around Helensburgh, which went to Argyll and Bute Council, and the area around Bearsden and Milngavie, which went to East Dunbartonshire.

Following the 1996 reorganisation, and around the time of the commencement of this study, the population was 96,000. Of these, 24,000 were in the 0-18 age range, with the breakdown into age sectors approximately as follows: pre-school (0-4), 6,000; primary school (5-11), 9,000; secondary school (12-16), 7,000; post-school (17-18) 2,000. There were 23 pre-school establishments, 35 primary schools and seven secondaries.

As a poor authority, all of its key economic statistics were well below national averages (General Registrar's Office, 1993; Office of Population Censuses and Surveys, 1991; West Dunbartonshire Childcare Partnership, 1999). The proportion of 'economically active' people (those of working age who were either in work or seeking a job) was 44%, compared with 75% in the West of Scotland as a whole, and over 80% nationally. The proportion of lone parents with dependent children was 14%, compared with a national average of 9%. Over 50% of children lived in households defined as being 'in poverty', and in some areas this rose to 86%.

These statistics were reflected in the indicators most commonly used to measure socio-economic disadvantage in schools, such as entitlement to free school meals and footwear and clothing grants. Compared with a Scottish average of 19%, the proportion of school pupils with free school meal entitlement was 27%, the figure being second highest nationally both at primary and at secondary stages. All of the secondaries and 32 of the 35 primaries met criteria for being described as 'disadvantaged'.

This general picture of poverty was considerably magnified in two parts of the authority viewed as areas of multiple disadvantage, the East Clydebank corridor and the Vale of Leven corridor. A snapshot of the extent of poverty in one area, the district of Faifley adjoining East Clydebank, was provided in one of the preparatory studies conducted in the mid-1990s prior to the commencement of the current project (MacKay, 1995a). It illustrated not only the conditions in the poorest parts of the territory that later formed the new Council area, but also the effects of social policy

in the distribution of funds to areas for priority treatment (APTs). Because Faifley was paired with a nearby district with lower poverty levels it was not accorded APT status, yet its economic profile was the worst in the authority, as shown in Table 3-1.

Table 3-1 Demographic statistics for project school and other areas (percent)

	Project School	Nearest APT	Strathclyde	Scotland
Free school meals (primary)	60.2	43.9	27.2	23.0
Footwear and clothing grants	68.1	46.3	32.9	N/A
Owner occupied housing	3.4	27.8	47.2	52.1
Lone parent families	30.4	29.9	18.1	6.0
Male unemployment	21.8	18.1	15.7	12.6

Source: The 1991 Census Report for Scotland (General Registrar's Office, 1993), supplemented by school records and data provided by the Scottish Office Central Statistics Unit and by Strathclyde Regional Council Chief Executive's Department.

In sharp contrast, East Renfrewshire, situated to the south of Glasgow is one of the wealthiest Council areas in Scotland, with low unemployment and a high proportion of owner-occupied housing.. With a population of around 90,000 it is of similar size to West Dunbartonshire, but at the time of the declaration study it had the fifth lowest rate overall of free school meal entitlement in the country (9.2%), and third lowest in its primary sector (8.7%). Nevertheless, it too had pockets of significant disadvantage in the Barrhead and Thornliebank areas, allowing a mixed study of high and low socio-economic areas to be carried out.

Both in West Dunbartonshire and in East Renfrewshire the studies were firmly embedded in the Council's policies and objectives for raising educational achievement and for addressing the needs of the most vulnerable sections of the community.

SUMMARY

This chapter describes the research context in which the West Dunbartonshire Literacy Initiative originated and developed – one that combined a research vision of transforming the reading achievement of all children in one of Scotland's poorest Council areas with the support of a Council that was willing to commit itself to a visionary strategy for change. The study proceeded on an agreed basis of an intervention in which the content and methods would be governed by the evidence base and research principles of psychology. It also articulated and promoted the importance of 'context variables' of vision, profile, ownership, commitment and declaration. The demographics of West Dunbartonshire, an area of marked socio-economic disadvantage, are described and contrasted with those of the setting for one of the subsidiary studies.

Chapter 4

The Context of Socio-Economic Disadvantage

The context in which the studies reported in this investigation were carried out was one of significant socio-economic disadvantage. Of the five studies, four were conducted with a population which fell almost totally in the disadvantaged category, and the remaining study (the declaration study), was based on a sample drawn equally from high and low socio-economic groups.

Initiatives aimed at raising achievement in areas of socio-economic disadvantage, and especially those that focus on children and young people, have particular relevance at the present time to the government's agenda in Scotland and more generally in the UK. The impetus to tackle this area is fundamental to the programme of social inclusion, promoting mental health and fostering effective working among agencies. The government considers this aspect to be of such importance that it has described tackling social and economic inequalities as an 'overarching aim' (Scottish Office, 1999).

Social inclusion is at the heart of the government's programme of social and economic reform. Socio-economic disadvantage has been seen as having a profound influence on health, mental health and education, and ready access to appropriate services and support have been emphasised for those living in disadvantaged circumstances. The *Research Strategy for the National Health Service in Scotland* (Scottish Office Department of Health, 1998) identified addressing inequalities as the key strategic aim for public health research, while the White Paper *Towards a Healthier Scotland* promised 'a sustained attack on social exclusion' (Scottish Office, 1999). The aim, as outlined in the report *Social Justice: A Scotland Where Everyone Matters* was 'to focus support on the most vulnerable young people who are in greatest danger of becoming permanently excluded' (Scottish Executive, 1999). As previously noted, this included a target of reducing by one third the days lost every year through exclusion from school and truancy.

All attempts to define the construct of disadvantage and to establish an agreed terminology involve a number of difficulties, and no single term or interpretation is fully satisfactory. Indeed, it must be recognised that disadvantage is a dimensional and not a bi-polar concept. It is not therefore a concept that may be of uniform application to a whole area or even to a whole school population. It is something that occurs in every degree from the most advantaged to the most disadvantaged sectors of any society. In addition to 'socio-economic disadvantage', the term 'deprivation', sometimes prefaced by adjectives such as 'social' and 'cultural', is frequently used, together with other terms such as 'low social class' and 'poverty'. Reference must be made to each of these as they occur frequently not only in general parlance but also in official documents.

The term 'deprivation' is in very wide use, and indeed has probably been the most common term found in the official descriptions used by Scottish Councils regarding

their 'areas of deprivation'. To speak of a population as 'deprived', however, involves two difficulties. The first is that it begs the question, 'Who deprived them?' This is essentially a philosophical and political question, and one regarding which there is far from consensus of viewpoint. The second is that it is a term that is not always found acceptable by those whom it describes. It carries pejorative overtones and, in particular, terminology such as 'culturally deprived' may be interpreted as the judgement of one culture on the norms of another.

A similar criticism may be made of a term such as 'low social class'. Its wide usage stems in part from the Registrar General's classification of occupations into Social Classes I-V. Again, the concept of class is seen to carry many potential value judgements reflecting not only on people's occupational or economic position but on their behaviour, attitudes and values.

'Poverty' has been viewed widely as a useful description of the underlying constructs covered here. It is simple and direct, and appears frequently in government and other publications. It has been particularly used in the UK by the Labour administration first elected in 1997, which made a Manifesto commitment to putting an end to child poverty as part of a long-term strategy. This term too, however, is not without difficulties. As Brown, Scott, Mooney and Duncan (2002) have noted:

'Ideas about poverty are complex, often contradictory and influenced by factors such as personal experiences, value judgements and belief systems. Inevitably definitions of poverty are contested. There is no single, universally accepted definition' (p. 11).

A particular difficulty lies in the fact that when used in affluent and highly developed countries such as the UK, the term means something very different from its use in other nations. It is not referring to 'absolute poverty', a term indicating lack of the resources needed to sustain physical survival at the level of being fed, clothed and housed, but rather to 'relative poverty'. This relative concept is an important one, because it locates the idea of poverty within the standards of living in a particular society at a given time. People are viewed as living in poverty in this sense when they do not have the resources sufficient for them to take part in the activities normally available to others in the same society. This is of central significance to the government's agenda for 'social exclusion' and 'social justice'.

The concept of relative poverty is important for another reason. There is considerable evidence that quality of life outcomes are affected not only by absolute socio-economic circumstances but also by comparative ones, and this is reflected in much of the research cited throughout this chapter.

The term 'socio-economic disadvantage' has been preferred here, partly because it may be seen to involve fewer political or cultural value judgements, and partly because it already has wide currency in the academic literature. However, where the other terms are used in direct quotations or for other purposes, and in particular the term 'poverty', they should be taken as synonymous with 'socio-economic disadvantage'.

Socio-economic disadvantage and social justice

MacKay (1999c) has argued that socio-economic disadvantage is not only a major dimension of inequality of opportunity in its own right, but is also the most significant dimension through which all other forms of inequality of opportunity are mediated. The most vulnerable members of society in terms of all of the traditional areas of gender, race and disability are over-represented in the lowest socio-economic groups. While the relationship is a complex one, the general facts are well documented. Those who belong to a minority ethnic group are three times as likely to be long-term unemployed and twice as likely to be permanently sick or disabled. Those who are employed and belong to these groups earn less, and if they are women they earn less again (Equal Opportunities Commission, 1996; Whitmarsh & Summerfield, 1996). In summary, the groups over-represented in the poorest sectors of society include minority ethnic families, lone parent families (in nine out of 10 of which the parent is a woman), large families, families with very young children and families with disabled people (Equal Opportunities Commission, 1996; Kumar, 1993; Office of Population Censuses and Surveys, 1991).

Although the government has made a commitment to addressing poverty, it has also been argued that the structure of public expenditure and services often exacerbates rather than ameliorates disadvantage and inequality of opportunity. As indicated in the discussion on values in science (Chapter 2), this has been illustrated in analyses of the structure of social justice in relation to government spending, poverty and the council tax, the language of education, discrimination in public services and the distribution of public resources (MacKay, 1982, 1983, 1999c, 2000c, 2001a, 2003a, 2003c). Bramley (1997) has noted that the poor are particularly disadvantaged in respect of 'local public goods', for example, the cleanliness of the local area, the quality of open space, access to school resources and the extent of home disrepair.

Attempts by the government to address some of these issues by giving priority to disadvantaged areas often fail to address the real problems. In the context of the current research, this issue is discussed in relation to the preparatory studies described in Chapter 5. The area with the greatest level of need had no special status for extra funding because of being grouped together with a more wealthy nearby area. A similar difficulty was raised by Barnes and Lucas (1975), who found that in the Inner London Education Authority the majority of disadvantaged primary school children were not located in 'educational priority area' schools and, even in such schools, the majority of children were not seriously disadvantaged. They argued for educational policies designed to locate and help disadvantaged children wherever they are instead of global policies targeted on whole schools or areas. Cox and Jones (1983) underlined this need to review policy and emphasised that public authorities must 'adopt more sensitive measures for locating individual children and their families who are disadvantaged in a clearly defined way or otherwise educationally at risk' (p. 125).

Indeed, inequality of opportunity for disadvantaged groups can often be seen in terms of access to the very services that have been established for their benefit. Many years ago it was shown that while the problems with which the caring professions are

concerned – mental illness, physical disability, educational underachievement – are over-represented in the lower socio-economic groups, these groups are often under-represented in the distribution of resources and in the extent to which they can gain access to them (Hollingshead & Redlich, 1958). Tudor Hart (1971) described this as the ‘inverse care law’, whereby resources are distributed in inverse proportion to need.

In a Scottish educational context this has been demonstrated in relation to the distribution of the Record of Needs. In an analysis of recording patterns in mainstream schools, using data gathered from all 2,367 primary schools in the country, MacKay (2000c) concluded that ‘the distribution of the Record of Needs in Scottish schools does not reflect real levels of need in terms of the clear implications of socio-economic disadvantage’. It was also noted that ‘the threshold of recording differs between less disadvantaged and more disadvantaged areas, and that this threshold may be influenced both by schools and by parents’.

Further evidence of bias in the distribution of resources was found by Sacker, Schoon and Bartley (2001) using data from two British cohorts (the National Child Development Study and the 1970 British Cohort Study). While there were more children from manual homes receiving special needs support in school than children from professional homes, when level of need was taken into account the gradient reversed, so that children from professional homes were more likely to be getting help. They concluded that ‘schools with the greatest requirement for special help provision are those least able to supply it’.

While low socio-economic status and its inevitable association with a poorer quality of life in terms of housing, diet and other factors can lead to direct effects of disadvantage on general well-being, there is considerable evidence of the effects of relative, as opposed to absolute, poverty. This has highlighted the importance of psychological factors when sections within a society have significantly diminished access to the range of benefits, privileges and opportunities available to the rest of the population (Wilkinson, 1992). This issue is of considerable significance in the UK context in recent years, where two trends have occurred in relation to relative poverty. First, there is evidence of a marked decline in the real income of the lowest 10% of the population during the 1980s and 1990s, amounting to a drop of 13% (Department of Social Security, 1996). However, the same government source indicates a growth in the average real income of the general population of 39% in the same period. These comparisons may be placed within an international context. Britain has the highest relative child poverty rate in the European Union (European Union, 2001), with the rate increasing more than in any other country between the mid-1980s and the mid-1990s (Bradshaw, 2002). Thus, although many indicators of poverty have improved over a long period of time, there are statistics which point to a recent increase in the level of disadvantage of the poorest sector of the population both in relative and in absolute terms.

Socio-economic disadvantage, health and quality of life

The impact of social disadvantage is experienced literally from the cradle to the grave. It is associated with significantly higher infant mortality rates (Kumar, 1993) and significantly lower longevity (Davey Smith & Egger, 1992). The entire span of life in between is marked by poorer health on virtually every measurable indicator and by a higher incidence of physical and mental disabilities (for example, Blaxter, 1990). It is this pervasive nature of disadvantage, in affecting these three key areas of human well-being, mortality, morbidity and quality of life, that marks it out as a comprehensive vehicle of inequality of opportunity.

The second half of the 20th century brought a marked decline in infant and child mortality rates, which are already very much lower in economically prosperous nations than in developing countries. Although data recorded on SES are incomplete, all studies have shown a clear social class gradient, with significantly higher figures for Social Classes IV and V (Office for National Statistics, 2000, 2001a, 2001b; Whitehead & Drever, 1999). This pattern is sustained throughout the entire age span of childhood. By far the highest level of child deaths from all causes is for Social Class V, where for every age group it is more than twice the level for Social Classes I and II (Botting, 1997). Most of the causes of infant and child deaths, except cancer, show this class gradient, with accidental deaths showing the strongest gradients of all (Office of Population Censuses and Surveys, 1988). For example, children in Social Class V are more than four times as likely to die as pedestrians as children in Social Class I (Towner, Jarvis, Walsh & Aynsley-Green, 1994). In Glasgow, a city with a high index of poverty, children are twice as likely to be knocked down as elsewhere in the UK, and in Drumchapel, one of its poorest areas, six times as likely (Thomson & Whelan, 1997).

The social class gradient in mortality is sustained not only throughout childhood but for all ages, and this is supported by a vast number of studies (for example, Walker & Walker, 1997). A Scottish study by McLoone and Boddy (1994) reported mortality ratios in disadvantaged areas. They showed that between the early 1980s and the early 1990s the ratios for the most disadvantaged areas of Glasgow relative to the least disadvantaged had risen from 96% to 124% higher. This trend has not been reversed since then but has increased. The differences are not only statistically significant: they are important and meaningful. Probably the most direct way to translate the current official figures into real life is to note that male life expectancy in the poorest areas of Glasgow is a full 12 years less than in neighbouring wealthier areas (NHS Scotland, 2004).

It is almost axiomatic, given the statistics for mortality in terms of infant and child deaths and longevity of life, to state that the health of people who live in disadvantaged areas is significantly worse than for those who have greater prosperity. The official statistics and research data covering this area are so extensive that to speak of the relationship between health and social class becomes a truism. For example, between 1985 and 1993, over 400 empirical studies were published in Britain documenting the nature and extent of the health divide (Mackenbach, 1994, cited in Walker & Walker, 1997). It is also probable that morbidity in the lower

socio-economic groups is under-reported, since individuals with less education tend to report less chronic illness than those with more education, when compared with medical diagnoses (Mackenbach, Caspar & Joost, 1996). When self-reports of overall health among children are invited, significantly fewer in Social Class V than in Social Class I report being in 'good' or 'very good' health (Prescott-Clarke & Primatesta, 1998).

A similar pattern is reflected in relation to mental health. While many facets of this area are complex, the overall data indicate that the prevalence of mental disorders in Social Classes IV and V is more than twice that in Social Classes I and II (Meltzer, Gatward, Goodman & Ford, 2000). Apart from mental disorders, when the more positive issues of 'mental health' as such are considered, the lower socio-economic groups score considerably lower on measures of self-esteem (Rosenberg, 1965) and happiness (Harding, 1985).

This pattern of disadvantage in relation to health, mental health and other factors is reflected also in the distribution of special educational needs at every level. Maxwell (1994), in a study of special needs and socio-economic disadvantage in Aberdeen school catchment zones, concluded that disadvantage in comparison with all other factors is 'a much more powerful predictor of the number of pupils from a zone who will be found to have major special educational needs of an intellectual/cognitive nature, physical or sensory nature'. The socio-economic disadvantage variable alone accounted for 87% of the variance.

In summary, the picture is a very clear one. People who are poor are more likely: to die in infancy or childhood, to have lower life expectancy, to have congenital abnormalities, to be disabled, to have worse sequelae of serious illnesses, to have serious head injuries in childhood, to be burned to death, to have a mental disorder, to have low levels of self-esteem and happiness or to be knocked down and killed.

Socio-economic disadvantage and educational outcomes

The relationship between social disadvantage and educational underachievement has long been established. Writing in 1918 in relation to arithmetical skills, Ballard pointed to a difference of more than two years between the addition and subtraction performance of high and low socio-economic groups, in favour of the former (Ballard, 1918). In a discussion of the same area over 30 years ago, MacKay (1971) noted a similar finding and its significance for attempting to establish meaningful norms of performance across schools:

'When the results were being calculated, it became evident that there were marked differences between one school and another. A comparison was made between two schools, School A and School B, the former having the advantage over the latter in terms of socio-economic status. Analyses of variance were carried out on the differences, the *t* values for all of them being significant beyond the 0.01 level' (p. 15).

Other early studies have demonstrated the clear relationship between educational attainments and socio-economic disadvantage. Douglas (1964) in his home and school study found that at age eight years children of parents of non-manual workers performed better than those of manual workers, with the differences increasing in later years. The evidence gathered by the Plowden Committee (1967) on this subject fuelled the impetus towards comprehensive education in the 1970s.

The social class gradient in performance is again robust across every facet of educational attainments, and children from culturally and materially disadvantaged homes score significantly lower than controls on all measures of linguistic and scholastic attainments, as well as being less well adjusted socially and emotionally to their school environment (Cox & Jones, 1983). This has been demonstrated in a very wide range of studies (for example, McCallum, 1993; Patterson, 1992; Sacker, Schoon & Bartley, 2002; Sammons, 1995; Shuttleworth, 1995; Strand, 1999). While school factors are also important, a strong correlation between disadvantage and underachievement that is independent of school attended can be demonstrated (Mortimore & Blackstone, 1982). The effects of disadvantage are long-term, persisting throughout an individual's school career, and the gap in attainment over time tends to become still wider (Smith & Noble, 1995).

Kumar (1993) reviews some of the evidence to support a direct causal relationship between disadvantage and educational underachievement. Low income means that parents have less money to spend on books, toys and extra-curricular activities. Parents in low income jobs tend to work long, unsocial hours, frequently doing shift duty. There is also a relative lack of knowledge of the education system and of the skills to influence it. Higher levels of unemployment lead to stressful effects on physical and mental health and increased marital discord. There is also evidence of a higher incidence of undiagnosed hearing and visual impairment, together with unsuitable facilities for homework, inadequate diet and poorer sleeping patterns.

The effects of disadvantage, however, are complex and do not relate crudely to socio-economic status. Essen and Wedge (1982) defined social disadvantage using the three measures of atypical family, poor housing and low income and noted a 'clustering of adversities'. Fifteen times as many 16 year olds were multiply disadvantaged as there would be if the problems were distributed randomly. Garner (1989) studied disadvantage in Lothian at four levels – individual, family, school and neighbourhood – and demonstrated significant neighbourhood effects. Two young people with identical home and family characteristics and attending the same school could expect quite different outcomes in terms of educational achievements.

Socio-economic disadvantage and literacy

The association of social disadvantage with poor attainment in reading and with serious reading failure is particularly marked. Davie, Butler and Goldstein (1972) in the National Child Development Study reported that 50% of children aged seven in Social Class V had poor reading scores compared with 7% in Social Class I, with an unskilled manual worker's child having a six times greater chance of being a poor reader. Essen and Wedge (1982) found that the disadvantaged children in their study

included ‘an unusually high proportion of particularly poor readers’ whose reading scores were more depressed relative to controls than were their maths scores. Cox and Jones (1983) noted that their control and disadvantaged groups ‘were most sharply differentiated in terms of their reading test scores’, and that a significantly higher proportion of the disadvantaged children were seriously retarded in reading at the end of their primary school careers. Results from their original infant school study showed that the differences were apparent during the first year in school but also that at age 11 these differences were even greater than would have been predicted on the basis of performance at age seven.

The Scottish study by McMillan, Fox and Wood (1994) again confirmed that ‘there are massive differences in attainment in reading between children from relatively affluent areas and children from relatively impoverished areas’, and that ‘simply by identifying schools with high proportions of pupils who receive free school meals, one can predict schools with high proportions of children with reading difficulties’ (p. 29). The evaluation of early intervention initiatives for literacy in Scotland continues to demonstrate that this gap has not been diminished (Fraser, MacDougall, Pirrie & Croxford, 2001).

As well as highlighting the gap in literacy attainment between advantaged and disadvantaged groups, the McMillan et al. (1994) study was of critical importance in defining a key problem in service provision within the Scottish context. For the first time it provided data relating to literacy in Scotland that demonstrated the operation of the ‘inverse care law’ of Tudor Hart (1971) described above, whereby resources are distributed in inverse proportion to need. This study had a major political impact, and was so central both to the future of Scottish Executive initiatives to support literacy in disadvantaged areas and to the key values and aims of the present study that its methodology is considered here in detail.

The McMillan et al. (1994) study was conducted with very modest resources in a small number of schools in one pocket of a Scottish education authority, Lothian Region. It was carried out under the professional development initiatives of the Scottish Executive (at that time the Scottish Office Education Department), whereby a small number of research projects in educational psychology was sponsored annually. While its aim was limited to examining attainments in an area of socio-economic disadvantage, to observing classroom activities and to discussing possible interventions with school staff, it also used a simple but effective methodology to demonstrate inequality in service provision.

Four of the five schools that comprised the sample were in the Pilton area of Edinburgh, an area described by Garner (1989) as one of the most socio-economically disadvantaged in Lothian Region. A fifth school in a relatively advantaged area was selected for comparison. Percentage of pupils receiving free school meals was used as an indicator of socio-economic disadvantage. This was on average above 50% in the four target schools, but was only 3.6% in the comparison school. First, pupils in one Primary 4 class in each school (total N = 112; average age = 8y 2m) were assessed on a range of literacy skills. Stark differences were found in attainment levels, and these closely reflected the socio-economic differences between

the target schools and the comparison school. While three-quarters of the pupils in the P4 classes in the disadvantaged schools had reading accuracy ages below 7 years, fewer than one in 10 was at this level in the comparison school. The comparison for scores below 5y 9m was 42% against 0%, and for spelling age below 7 years was 62% against 4%.

Second, a possible relationship between literacy attainments at this stage and phonological skills was investigated. The three pupils with the best and poorest attainments in reading in each of the P4 classes in the four target schools were assessed on measures of receptive vocabulary and phonological knowledge. This provided evidence of lower scores on these measures for the poorest readers.

Third, an obvious recommendation arising from the high occurrence of literacy difficulties in P4 in disadvantaged areas, together with an association with poor phonological skills, would be to ensure that higher levels of effective learning support were targeted on those pupils with the highest level of need. However, when the distribution of learning support to pupils across these schools was investigated a quite different pattern emerged. There was virtually no overlap in scores between those on learning support in the more advantaged school and those in the target schools. The authors noted that if the same entry criteria were applied to the two poorest schools as to the most advantaged, then about 90% of the pupils in the poorest schools would have been receiving learning support.

This study made a powerful statement regarding the needs of disadvantaged areas and the distribution of learning support resources. It did, however, have a number of methodological weaknesses. Only one class in one school was used for the comparison school, and it may not have been representative in terms of pupil characteristics, teaching methods or attainment levels. Also, while impressive percentage differences in numbers reaching specified attainment targets were quoted, no statistical analysis was available to point to their size or significance.

In addition, both the rationale for selecting phonological skills for investigation, and the conclusions drawn from it, must be questioned. It was the only factor investigated as having an association with the literacy levels cited, but there may have been others equally or more important. The authors concluded that 'there is extremely strong evidence that good readers are more proficient in their phonological skills' (McMillan et al., 1994, p. 21), yet some of the evidence to support this conclusion was not significant for some of the analyses used, and some was at only a modest level of significance. Furthermore, the sample did not look at all the scores but only those for a few children at the two ends of the available distribution of scores. Given that most of the low-scoring children were so low as to be below the norms on the reading test used, it is possible that the combination of low literacy and low SES simply pointed to a sample that was likely to have lower scores on any measure, and that phonological skills may have had little real relevance. Indeed, the sample was a very small one (three high and three low attaining pupils from each of four target schools).

Similar criticisms may be made of the conclusions about learning support levels. The comparison was with only one school class, where six children were receiving learning support. This may not have been representative of overall distribution of learning support resources throughout the school or in other schools. Indeed, figures referred to in the study suggest that the advantaged school may have had larger class sizes, and the actual allocation of learning support was somewhat lower than for any of the target schools. The authors also raised the additional issue of too little time spent on reading in school on the basis of observations of pupils in Primary 2 classes, and this led to a key recommendation of having more time devoted to this area of the curriculum. However, the observation was limited to the target schools, so it did not provide any information for comparison regarding the time spent in the advantaged school.

While these methodological criticisms raise issues for this study, they do not remove the force of its central arguments. It is acknowledged that the differences in attainment between the target schools and the comparison school were stark, that the threshold at which learning support was able to be provided was clearly much higher in the comparison school and also that these results were fully consistent with the trend of findings reported throughout this chapter in relation to socio-economic disadvantage. The study was undoubtedly an important one in the context of Scotland. Most significantly, it recommended the introduction of an early intervention scheme to support literacy in schools in disadvantaged areas. When the report was published it created a considerable political stir, and within three years of its publication the Scottish Executive announced a comprehensive early intervention scheme for literacy in socio-economically disadvantaged areas.

The complex relationship between disadvantage and literacy

Although the relationship between socio-economic disadvantage and poor achievement in literacy is robust it is nevertheless complex and, in common with all other outcomes associated with socio-economic disadvantage, it is not a simple correlate of poor economic circumstances. The influence of parental behaviour is clearly an important factor in literacy outcomes irrespective of socio-economic status (Adams, 1990; Gregory, 2000; Kumar, 1993). The same may be said for the influence of siblings, grandparents and other family members (Gregory, 2000; McQuillan & Tse, 1995; Rogoff, 1990).

The importance of home and parental factors is also demonstrated in international comparisons. The Progress in International Reading Literacy Study (PIRLS) assessed performance in reading in children aged 9-10 years in 35 countries. Across all countries, students with the highest level of reading achievement had parents who themselves spent time reading and had favourable attitudes towards reading (Mullis, Martin, Gonzalez & Kennedy, 2003). Home-school connections were also a major focus in the Programme for International Student Assessment (PISA), a three-yearly survey of the knowledge and skills of 15-year-olds in the principal industrialised countries (Organisation for Economic Co-operation and Development, 2004).

International comparisons also provide an interesting commentary on the relevance of relative (as opposed to absolute) poverty to literacy. MacKay (2002a, in preparation a) examined the context of illiteracy in South Africa, where the problem is extensive and has vast economic, social and political implications. The analysis provides a clear indication of the effects not only of socio-economic disadvantage but more particularly of relative poverty, and it is described here in some detail.

South African illiteracy is complex and, on initial examination of available data, represents a paradox. First, it is an objective problem. The UN Human Development Report estimated illiteracy in Europe, USA, Japan, Australia and New Zealand at 1%, but in South Africa at 15% (United Nations, 2001). However, it is also a comparative problem. There is a dramatic difference between where one might assume the country should be in terms of literacy and where it actually is. Table 4-1 shows a comparison for 162 countries, derived from data published in the UN Human Development Report. These data indicate a considerable gulf between South Africa's 'inputs' and its 'outcomes'. It is a country with a comparatively high gross domestic product (GDP), high educational enrolment and a high proportion of its resources devoted to education. Yet it has much lower rankings for adult literacy and for the broader Human Development Index (HDI) of which literacy forms a part. The HDI is a composite index based on 'a long and healthy life' (life expectancy at birth), 'knowledge' (adult literacy and educational enrolment levels), and 'a decent standard of living' (GDP per capita). Since GDP and educational enrolment are therefore represented in both columns of Table 4-1, a final outcome figure has been shown in which the HDI is adjusted for these two variables. This highlights the difference between inputs and outcomes as being still more prominent. For a country that ranks in the top 10 in the world in terms of its educational spending it ranks bottom in terms of the key quality of life indicators of life expectancy and education.

Table 4-1 Illiteracy in South Africa: a comparison across 162 countries

Inputs	Outcomes
GDP 72%ile (rank 45/162)	Adult literacy 48%ile (rank 85/162)
Educational enrolment 91%ile (rank 15/162)	Human Development Index 42%ile (rank 94/162)
Educational spending 93%ile (rank 10/141)	HDI adjusted for GDP and enrolment <1%ile (rank 162/162)

However, there is a key variable that becomes apparent when the published international statistics are examined in further detail. It is the context of comparative social and economic inequality. The total share of national income/consumption available to the poorest 10% of the South African population is 1.1%. Social inequality at this level places South Africa at the 6th percentile, 105th out of 111 countries for which comparative data are available. The Gini Index provides a

measurement over the entire distribution of income/consumption, where a score of zero represents perfect equality and a score of 100 perfect inequality. South Africa's Gini Index of 59.3 places it at the 2nd percentile, 109th out of 111 countries.

It is then necessary to ask whether a correlation can be established between social and economic inequality and illiteracy. This can be demonstrated quite simply by looking at South Africa's nearest GDP neighbours. Table 4-2 shows the results of a comparison with the 16 countries closest to South Africa in GDP per capita, eight of these having a higher GDP and eight a lower. When these are ranked by their Gini Index and then placed in two groups there is a marked divergence between their illiteracy figures, despite almost identical average GDP.

Table 4-2 Illiteracy and inequality: South Africa compared with her GDP neighbours

	Lowest inequality	Highest inequality
Gini Index	27.8	50.0
Illiteracy	1.9%	10.7%
GDP per capita (US\$)	8171	8215

These figures represent a context of relative poverty. In comparison with other nations South Africa has greater inputs in terms of national wealth and educational spending and enrolment, but worse outcomes for literacy and other quality of life factors.

All of the above issues regarding the effects of socio-economic disadvantage are crucial to the current study. It was conducted principally in the second poorest Council area in Scotland – an area affected by all the problems that have been outlined in terms of higher infant mortality and lower life expectancy, poor physical and mental health and endemic educational underachievement. These outcomes interacted with the compounding issues that have been raised in terms of the structure of government spending, the distribution of resources and access to public services. It is an area where, in its most disadvantaged quarters there are high levels of crime and vandalism and of alcohol and drug abuse, together with an impoverished physical environment.

In terms of literacy outcomes the figures were stark. In 2002, over a quarter of the entire school population left primary school with reading scores described as 'functionally illiterate'. Many who had preceded them in previous years with similar scores had gone on to become demotivated and alienated in the secondary school environment, and as a result to contribute to poor school ethos, low staff morale and high levels of disruptive behaviour.

Many attempts have been made to address underachievement in disadvantaged groups. However, a problem has always remained: the large number of young people who continue to leave school without having developed the basic skills of competent literacy, and who are unable either to make a full contribution to a prosperous economy or to enjoy a higher quality of life. It was the challenge of this study to address this issue, and to lay the foundations not only for raising educational achievement but also for eradicating illiteracy.

SUMMARY

This chapter portrays socio-economic disadvantage as being not only a major dimension of inequality of opportunity in its own right but also the most significant dimension through which all other forms of inequality of opportunity are mediated. It summarises the evidence for the impact of socio-economic disadvantage on health and quality of life, on educational achievement and on literacy in particular. It also explores the concept of relative poverty in relation to international comparisons and their relevance to literacy outcomes. While recognising that the concepts of advantage and disadvantage are dimensional and not bi-polar and that the relationship between socio-economic disadvantage and illiteracy is a complex one, it is nevertheless asserted that this factor is of the most fundamental significance in terms of quality of life and educational outcomes. It is presented as being of central importance to this study, and to its aims of raising achievement and addressing endemic problems of illiteracy.

Chapter 5

The Preparatory Studies

The studies reported here were preceded by three studies, or groups of studies, which were completed and published prior to the commencement of the current project. They were of particular importance to this project in that they provided, albeit on a small scale, part of the essential context from which it could be developed. They are therefore reported in this chapter in some detail. The first was a series of studies of playground behaviour. In common with the other preparatory studies, these were conducted in areas of multiple socio-economic disadvantage within the territory that became West Dunbartonshire at local government reorganisation. They were significant not only in highlighting the relationship between difficult behaviour and educational underachievement, but also in the development of an intervention paradigm that was then applied to enhancing literacy.

The second study was a project on addressing reading failure. Its significance was twofold: first, it illustrated just how profound the problems of educational underachievement are in disadvantaged areas, and therefore the need to explore methods for moving effectively from small-scale, short-term to large-scale, long-term interventions; second, it illuminated key factors other than the content of the reading curriculum, particularly attitudes and values, that were crucial to enhancing literacy skills. The third study involved setting up an early literacy intervention programme. It represented an important step towards addressing reading skills in the earlier years of schooling, and in establishing a broad base for a multiple-component intervention.

‘Fair Play’: the primary school playground

In the mid-1990s Briggs and her colleagues conducted a series of studies of enhancing playground behaviour in schools with high levels of socio-economic disadvantage under the title of ‘Fair Play’: the primary school playground’ (MacKay & Briggs, 1994; Briggs, MacKay & Miller, 1995). These studies represented an attempt by educational psychologists to introduce systemic interventions into schools with high levels of individual referrals to psychological services for behaviour problems and learning difficulties.

The playground was selected because it was considered to be as important a determinant of school ethos as the classroom, and often more so. Traditionally it has been viewed as peripheral to the ‘real’ curriculum of the school. Yet in a longitudinal study of children’s progress in London primary schools, Blatchford, Creeser and Mooney (1990) deduced from observations of 7-year old children that playtime and lunch breaks took up around a quarter of the school day – as much time as was spent on the combined areas of reading, writing and mathematics. Many serious problems of disruptive behaviour and bullying that were referred to the psychological service were not occurring within the context of the classroom or the school building but in the playground. The result was an impairment in the quality of relationships within

the school because of violence and unresolved conflicts between children at playtime (Axford, Blundell, Brown & Macphee, 1994). It was appropriate to address the behaviour in its context, and this pointed to the value of a more general approach to the structure of playtime and the nature of the playground environment in addition to individual intervention.

The centre-piece of the above initiatives was the 'Edinbarnet Playground Project', a study with a focus on changing aggressive behaviour through a structured intervention (Briggs, MacKay & Miller, 1995). There is a strong argument for intervening with aggressive children as the long-term implications can be serious. Aggressive behaviour has been identified as a major predictor of later alcohol and drug problems (Achenbach, 1982), delinquency (Roff & Wirt, 1984) and mental health problems (Eron, Huesmann, Brice, Fisher & Mermstein 1983; Robins, 1978). Problems of aggressive behaviour in the playground need to be dealt with on a number of levels. A whole school approach in conjunction with parents, pupils and all staff must be developed and the role of the physical playground environment must be examined. Children who are referred to educational psychologists frequently have problems with aggressive behaviour but the level of satisfaction with the outcome of such referrals is variable (Sloane, Endo, Hawkes & Jenson, 1990).

The aims of the study were: to boost self-esteem and promote an increased awareness in participating children of the negative effects of their behaviour on themselves and others; and to offer the school an alternative model of intervention for aggressive behaviour. The primary school that was the setting for this study had the highest rating on all indices of socio-economic disadvantage among over 120 primaries in the authority. Although not situated in one of the official 'areas for priority treatment' (APTs), it had a significantly higher percentage of footwear and clothing grants, free school meals, lone parent families and male unemployment, and a lower level of owner-occupied housing than in the nearest APT. The project was the response of the school psychologists to a request for help with a very high incidence of disruptive behaviour that had proved resilient to all measures used by the school.

At the start of the study the head teacher had described the playground as a 'battleground'. A pupil survey conducted by the authors with all children in Primaries 4 and 5 (N = 90) pointed to very high levels of bullying, with the playground being identified as by far the most frequent problem area. To address this, a structured groupwork intervention was designed, focusing both on skills and on relationships. The school selected 12 pupils who constituted their most difficult and challenging cases. These were allocated to two groups of six which were the subject of weekly groupwork sessions lasting one hour for approximately one school term. The value of groupwork in changing behaviour is well documented and there is evidence to suggest that a focus on cooperation in small groups in school can enhance the level of children's cooperative play in general (Daniels, 1990; Smith, Boulton & Cowie, 1990). A class teacher who knew the children well was released to work with an experienced groupworker provided by the social work department. A strong sense of group identity was fostered, and a key feature was the inclusion of games and enjoyment as a context for teaching social skills and personal qualities such as honesty and trust.

This intervention was highly successful in achieving its aims. In terms of quantitative results, significant positive changes were found with pre-post pupil questionnaires and playground incident records. All qualitative measures also showed large shifts in a positive direction. Teachers' ratings of the behaviour of the children in the groups showed a consistent change from 'bad' to 'good' and pupil interviews were highly positive. Staff interviews indicated that the school and its ethos had been totally transformed. The improvements in the project children had generalised through the school as a whole, and classroom behaviour was reported to have improved. Overall there was a strong perception that this study had not only achieved significant results but that it had also made meaningful changes to quality of life for the project children and for the school as a whole. A systematic follow up two years later demonstrated that enduring gains had been achieved.

This project was of central importance to the studies reported throughout this work. It highlighted the contribution of applied educational psychology at the level of systemic interventions in schools; it pointed the researchers towards the crucial role of context variables such as school ownership and the importance of profile and commitment in achieving successful outcomes; it confirmed the benefit of highly structured and focused methodologies in effecting change; and it indicated the value of recognising factors that are central to the process of enduring educational change, such as monitoring, support and follow up. It also paved the way for the following study on reading failure, by creating a context in which attention could be turned from dealing constantly with behaviour problems to addressing educational underachievement.

The Edinbarnet Reading Project

This randomised control trial (MacKay, 1995a, 1999b) was a piece of naturalistic action research that recognised the parameters dictated by the setting of a single school and necessarily small samples. It followed on as a coherent development from the playground project. The success of the playground study in transforming the ethos of the school and reducing behavioural disruption had allowed school management to develop a new focus on curricular issues. This led the school to identify reading failure as a priority area and to ask for project work in setting up an intervention to address it.

The need for an intervention in this area was beyond question. At the start of the study the school, in co-operation with the psychological service, carried out an analysis of the special educational needs of its 350 pupils in terms of difficulties in learning, development and behaviour. The number meeting the criteria that would normally justify consideration for formal individual referral was 140, and the number who could have been considered for opening a Record of Needs was 91. This made a poignant commentary on the extent of needs found in areas of multiple socio-economic disadvantage. It also showed that the occurrence of difficulties was so extensive that individual referral to psychologists would be totally ineffective, and that a different methodology was required.

The aim of the study was to design an intervention that would address reading failure, and its main objective was to enhance performance through changing attitudes and values regarding reading. There was a clear rationale underlying the idea of addressing attitudes and values. Most literacy interventions were concerned with curricular and methodological issues (reading schemes, programmes, strategies, approaches) or with a number of within-child variables (cognitive style, information processing, visual memory, auditory sequencing). However, knowledge of the school and of its pupils did not support a hypothesis of reading failure in which these factors represented the significant variables. On the other hand, the school was situated in an area of low incomes, high levels of crime and vandalism, single parent families and a mismatch between the culture and values of the school and its staff and those of significant sectors of the neighbourhood which it served.

The sample was drawn from Primaries 4 and 5, where the largest group with severe reading difficulties clustered. All pupils at these two primary stages were screened using the Norman France Primary Reading Test (France, 1978, 1981). Of these, 36 (47%) had a score of less than seven years. These were further screened for age, cognitive ability and reading performance. All pupils were tested individually on the revised British edition of the Neale Analysis of Reading Ability (Neale, 1989), and those with accuracy scores above 6 years 11 months were excluded. The final sample of 24 was placed in matched triads. One child from each triad was then selected blind by an independent observer and allocated to the groups, which were designated Control, Experimental 1 and Experimental 2. Mean age was 9y 0m, and mean reading age was 6y 0m.

The groupwork paradigm developed successfully for the earlier playground project was adopted as the main intervention strategy, again using a class teacher and an experienced groupworker. The aim of the groupwork was not to provide a context of reading tuition but to change attitudes and values regarding reading. The two experimental groups were each subject to a one-hour groupwork session each week during school hours for a period of 10 weeks. Experimental Group 1 received groupwork only. Experimental Group 2 received groupwork plus home support. The main part of the support took the form of paired reading, planned on a home visit by the author with each child and parent. Written details outlining the paired reading procedure were explained fully during the visit and were left with the parents for reference. Details and comments were passed between home and school each week on a card prepared by the school.

The results of this study were extremely encouraging, and positive changes for the experimentals were demonstrated both in measures of attitude and on reading scores. For measurement of attitude, prior to the experimental intervention a questionnaire developed by the two school psychologists, 'What You Think About School', was administered by class teachers to all children in P4 and P5. The questionnaire was again administered to these classes following the intervention period. Significant differences were found for experimentals versus class controls for pre-post attitude scores, supporting the hypothesis that children in the experimental groups would show positive attitude changes towards school compared with their peers.

Significant differences in reading scores in favour of each of the experimental groups were also found. These showed increases of approximately one year in reading accuracy scores (range 4-21 months) and one and a half years in reading comprehension (range 3-30 months). For the experimentals, no significant differences were found between the group receiving attitude change only and those who also had home support. However, a number of factors are of relevance here. First, samples were small and the duration of the intervention was short. Second, some of the parents in the home support group indicated that they themselves experienced difficulties with reading, and they may not have been able to provide the type of support required. Third, in some of the homes there were multiple difficulties and stresses at levels hardly compatible with providing effective support for a child in relation to reading skills.

In addition to the quantitative results, all qualitative measures supported the view that positive change had taken place for the pupils in the intervention. Teachers, parents and children emphatically endorsed the view that the project had been enjoyable and successful, that reading skills had improved and that there had been enhancement of factors associated with attitudes and self-confidence. Teachers indicated that in terms of improved attitudes and raised confidence levels, the comparison between the first group session and the last one was 'dramatic'. Despite the problems faced by some of the parents, all but one managed to complete a questionnaire, the results of which indicated improved attitudes and more reading at home. Semi-structured interviews with children showed a similar pattern, with greater liking of reading and increased confidence across the curriculum.

In addition to these encouraging results, this project again highlighted the importance of context variables. In terms of factors such as profile, ownership and commitment, school management were centrally involved in every stage of the planning process. Ownership rested firmly with the school and not with the psychological service or any outside agency. It was clear to both staff and pupils that the school viewed this initiative as being very important, and a high level of commitment to it was secured.

The Edinbarnet Early Reading Project

A very important outcome of the reading project was that as well as having immediate benefits for the participating children it led to a literacy study with a broader base. So far the focus had been on pupils in mid-primary who had already experienced reading failure, using mainly a single intervention strategy, the modification of attitudes and values. The 'Edinbarnet Early Reading Project' was set up as a broad-based early intervention study using a quasi-experimental design (MacKay & Watson, 1996, 1999). It focused on children at school entry who had not yet reached the stage of experiencing significant reading difficulties. Also, it adopted a multiple-component intervention, using a number of different strands designed to enhance literacy.

The experimental sample comprised two Primary 1 classes (N = 46) in the base school. These were matched in terms of stage and socio-economic status with pupils (N = 51) in two other primary schools in the same geographical area. The sample

was further subdivided by selecting 15 of the experimentals on the basis of their pre-test performance and targeting them for additional support (the experimental subgroup). These were matched with the 15 controls whose pre-test scores corresponded most closely (the control subgroup). All children in the experimental and control groups were assessed individually prior to intervention using assessment materials devised by the authors and covering readiness for reading, phonological and phonic awareness and knowledge of nursery rhymes. The attitude questionnaire referred to above, 'What You Think About School', was adapted for use by younger children and was administered to the entire sample.

Four main strands of research findings were identified as a basis for a multiple-component approach: first, enhancing classroom resources to allow additional direct instruction in reading; second, an increased focus on phonological training; third, a recognition of attitude and ethos factors; fourth, home support. Two ancillary staff were employed for five months to provide support to each of the experimental classes. Increased time was spent on activities relating to rhyme, rhythm, phonics, listening and matching, using curricular materials devised for the project. Every day reading groups from both classes were given priority for having extra teaching support provided by staff from within the school's normal staffing complement. A structured daily programme was also devised for raising confidence and self-esteem.

Pre-post comparisons indicated that the overall gains in literacy skills were significantly greater for the experimentals. No meaningful analysis of the attitudinal data was possible for this sample. The attempt to modify an attitude test down to 5-year level essentially served to highlight the difficulties in addressing this area with young children. However, it provided a valuable lesson that proved to be beneficial when a new method of attitude assessment for this age group was devised later for 'the declaration study'.

The positive quantitative results were again supported by a number of qualitative indicators. Staff described benefits that were absolutely central to the main aims of the research. The focus of the study was not mainly a short-term objective of enhancing P1 reading scores but a longer-term strategy of raising achievement through highlighting the centrality of reading and encouraging fresh approaches to the curriculum. The view of the head teacher was that the research had 'put reading high on the agenda'. It became a talking point not only in the infant department but through the whole school. Everybody in the school saw it as important. The study had stimulated a fresh look at the teaching of reading and all aspects of it. Class teachers had developed new ways of working and pupils were benefiting as a result. Reading and language had also become integrated into the whole curriculum in a pervasive way, resulting in a more effective and cohesive delivery of education in general. Other reported benefits included enhancement of staff morale, an interest in applying research findings to the curriculum, a fresh look at home links and working with parents, and the establishment of an initiative to foster a literacy environment at pre-school stage.

Together these three initiatives – the playground studies, the project on reading and attitudes and the early intervention study – provided the essential preparation for

designing the multiple-component intervention at the centre of this work. They highlighted key features for developing baseline assessment, for building several different intervention strands into one broad programme and for recognising the importance of context variables and change processes in successful initiatives.

SUMMARY

This chapter provides a description of three ‘preparatory studies’ that laid the foundations for this study. These studies, previously published by the author and colleagues, were *Fair Play – The Primary School Playground*, with particular reference to the *Edinbarnet Playground Project*, the *Edinbarnet Reading Project* and the *Edinbarnet Early Reading Project*. All were conducted in the area that was the main setting for this study, and their significance to the study is described. They highlighted the extent of behavioural problems and educational underachievement in the population being studied, but also pointed to ways in which these could be addressed. In particular, the *Edinbarnet Reading Project* pointed to the significance of attitudes and values in relation to literacy, and the *Edinbarnet Early Reading Project* prepared the ground for the multiple-component intervention at the heart of this study and the development of its baseline assessment measures. All studies also highlighted the significance of the context variables articulated throughout this study and the importance of taking account of the processes of educational and organisational change.

Chapter 6

Intervention in Literacy: An Overview

Effective multiple-component interventions to raise reading achievement and address illiteracy in areas of socio-economic disadvantage require considerable resources. The rationale for providing enhanced resources to schools in such areas has long been recognised at policy level (Thomson, 1996). However, if such resources are to be used effectively it is essential to plan them in a way that utilises available research findings. The overview provided here represents a general review of the literature on intervention in literacy. It is supplemented later by a further discussion of studies that support the key strands of the multiple-component intervention. This will be found in the chapters covering the rationale for the main study (Chapter 8), the synthetic phonics study (Chapter 10), the attitudes study (Chapter 12), the declaration study (Chapter 13) and the individual support study (Chapter 15).

The socio-cultural context of literacy

Traditionally the focus of reading research has been on psycholinguistic processes and reading itself has been viewed as a discrete skill that can be studied largely at the level of the interaction between the individual and the text. During recent years this view of the 'autonomous' nature of the text has been seriously challenged from within a number of disciplines and there has been a clear shift in interest to the importance of social and cultural variables. It is important to consider this socio-cultural context in a study focusing on socio-economic disadvantage.

In 1993 the United Kingdom Reading Association devoted an entire issue of its *Journal of Research in Reading* to the 'new literacy' studies. Street, a leading exponent in this field, provided a guest editorial introducing a series of articles exploring the social, cultural and indeed political context of literacy (Street, 1993a). The attitudes towards, and uses of, text vary considerably from one society or community to another. The emphasis of the new studies is on the part text plays in social lives and in determining status within a community. Street proposed an ideological model of literacy in which literacy practices were seen as inextricably linked to cultural and power structures in society. The apparent neutrality of literacy disguised its 'significance for the distribution of power in society and for authority relations' (Street, 1993b, p. 2).

Acknowledgement of this socio-cultural context almost dictates the need for a justification of the present study and its methodology. It would be inappropriate to introduce interventions that include attempts to influence cultural values and attitudes, assessment procedures based to a large extent on reading test scores and strategies designed to change literacy levels, without at least acknowledging the challenges implicit in the new literacy studies and their socio-cultural foundations.

This study does not ignore the critique that may be made of the traditional pedagogical, linguistic and cognitive approaches to literacy and the assumptions on

which these are based. Attempts to remedy reading failure are themselves embedded in power relationships within society, and these are reflected in the nature of educational interventions and the ways in which disadvantaged populations are described. This may be illustrated by reference to the concept of 'functional literacy', a critique of which is offered by Levine (1986). He proposes that there is little or no research to demonstrate that illiterate and disadvantaged adults are assisted significantly by achieving functional levels of competence. Literacy levels adequate to deal with instructions, signs and application forms, which tend to elicit conforming behaviour or institutionalise existing social arrangements, are just enough to bring people within the reach of bureaucratic modes of communication and authority. In so doing they may serve to 'domesticate and subordinate the previously illiterate person further rather than to increase his or her autonomy and social standing' (p. 41).

The present study, in acknowledging the centrality of cultural values to the development of literacy within the educational system, recognises the place of these socio-cultural considerations. However, despite the difficulties for non-readers of even achieving 'functional literacy', the development of reading skills is viewed as potentially empowering. This is the position adopted by Freire in his 'pedagogy of the oppressed'. His articulation of the concept of 'empowerment' was within the context of providing literacy programmes to thousands of Brazilian peasants. He viewed education in literacy as empowering people to engage in the political process through a critical awareness of the world (Freire, 1994, 2000).

As to the assessment instruments adopted, it is recognised that they measure only one dimension of achievement. Nevertheless, the basic skills measured by reading tests are of central importance. 'There is no serious challenge to the idea that they constitute the core of literacy' (Levine, 1986, p. 24).

Developing literacy: the importance of early experience

In her study of young fluent readers Clark sought to move the emphasis of reading research away from a primary concern with levels of literacy and difficulties in reading to an emphasis on the factors associated with the development of successful reading (Clark, 1976). While factors such as general cognitive ability, auditory discrimination and other psycholinguistic variables were not discounted, the main focus was on early experience and home background. All parents of the fluent readers in the sample seemed to value education and to wish for their children what they themselves had experienced or had missed. In general the children had available to them an interested adult with time to devote to them and to encourage an interest in reading. Their early experience therefore was one of valuing literacy, of having adults to talk to them, read to them and answer their questions and of having a wide range of reading materials available.

A primary emphasis of studies of children who are successful in achieving high standards of literacy has continued to be the importance of the early environment as one in which literacy skills are valued. The typical, well-prepared child in terms of pre-school reading skills at school entry age will have received thousands of hours of

literacy-related activities. Adams (1990) illustrates this in relation to the home experience of her eldest son:

‘Since he was six weeks old, we have spent 30 to 45 minutes reading to him each day. By the time he reaches first grade at age six and a quarter, that will amount to 1,000 to 1,700 hours of storybook reading – one to one, with his face in the books. He will also have spent more than 1,000 hours watching “Sesame Street”. And he will have spent at least as many hours fooling around with magnetic letters on the refrigerator, writing, participating in reading/writing/language activities in pre-school, playing word and “spelling” games in the car, on the computer, with us, with his sister, with his friends, and by himself’ (p. 85).

These observations are of central significance to interventions in areas of socio-economic disadvantage. In contrast to the home experience outlined above, Teale (1986) counted and timed the literacy events that occurred in the presence of each of 24 pre-school children from low-income homes in the United States. In the bulk of these cases storybook reading took place only about five times per year. For all but one of the children it averaged less than four hours a year. Adams (1990) estimated from Teale’s study that by the time these children were six and a half years old and entering first grade in the American school system their home experiences would have amounted to about 25 hours of storybook time. Other aspects of exposure to literacy events were equally impoverished in comparison with the type of middle-class environment Adams described.

While the causes of reading failure are multi-factorial and include low intelligence, sensory deficits and neurological impairments, the association with home environment and the encouragement of early literacy skills is of paramount importance. For this reason the impact of socio-economic disadvantage must be of central concern in the study of reading failure. What children bring with them to their first year of primary schooling, or before that to their nursery school experience, shows a crucial social class gradient of such proportions that it is almost inconceivable how the most disadvantaged children can catch up in the course of classroom reading instruction. It leads to what Stanovich (1986) has described as the ‘Matthew effect’ in which those who already have receive more, while the have-nots lose even what they already had. This process whereby individual differences in early literacy achievement in the nursery tend to be maintained or magnified over the school years has been confirmed in several studies (Tizard, Blatchford, Burke, Farquhar & Plewis, 1988; Wells, 1987; Weinberger, 1996).

Effective schools and effective teaching

‘How children learn to read well is still almost a secret. What kind of teaching helps them is a scarcely penetrated mystery.’ In this way Meek (1983, p. 1)) introduced her series of qualitative, longitudinal studies of adolescents learning to read. She continued by highlighting the importance of teaching, and the fact that the context of successful teaching is one in which the pupil is motivated to learn. Nevertheless, while there is a great deal of truth in the assertion that the process by which reading

is acquired involves much that is little understood, the vast research base on literacy and on effective schooling has clarified many aspects of effective teaching.

The account given above has emphasised the significance of early home experience. Rutter, Maughan, Mortimore and Ouston (1979) in their 'Fifteen Thousand Hours' study of the impact of schooling concluded that pupil background factors such as intelligence, home circumstances and pre-school learning accounted for 85% of literacy outcomes. The remaining 15% was attributed to schooling, of which 5% was school factors, 7% teacher factors and 3% method factors. Nevertheless, the fundamental importance of early experience and the home environment as predictors of success in reading does not diminish the role of schools and teachers but rather highlights its potentially crucial role.

The National Commission on Education (1996) undertook a project designed to explain how some schools in disadvantaged areas had succeeded against the odds. They identified a number of key variables. These included: the importance of leadership, with effective communication and team participation in decision-making; a vision of success and of how the school could achieve improvements; the careful use of targets; the improvement of the physical environment; high expectations about pupils' behavioural standards and academic success; and an investment in good relations with parents and the community.

There are fewer studies of 'good reading schools' than simply of 'good schools', but there is evidence that schools which are effective in one area of the curriculum and for one group of pupils are usually effective in other areas and for other pupils (Inner London Education Authority, 1986, cited by Wragg, Wragg, Haynes and Chamberlin, 1998). In relation to literacy as such, information gathered by Her Majesty's Inspectorate has led to the conclusion that, in schools successful in teaching reading, firm leadership is given at head teacher and other management levels, reading is given high priority, there are well-documented school policies that translate into actual classroom practice, the needs of individual pupils are addressed and there is a wide variety of appropriate reading materials (Department of Education and Science, 1991).

The impact of disadvantage, however, is pervasive and enduring. The basic question, 'Can school improvement overcome the effects of disadvantage?', has been addressed by Mortimore and Whitty (2000). There is no doubt that the question is a complex one, and many years of experience of school improvement projects such as those referred to above has led the authors to recognise the force of Bernstein's (1970) assertion that 'education cannot compensate for society'. At the same time, they do not regard educational disadvantage as an unchangeable fact of life, and they propose a variety of actions required both at government level and at school level. These include increased funding for early intervention programmes in disadvantaged areas, reconsideration of the approaches to learning and teaching used with disadvantaged pupils and extra support for pupils with disadvantaged backgrounds in school improvement programmes.

While school leadership, ethos and organisational factors play an important role, the quality of teaching at classroom level is crucial. In the study by Rutter et al. (1979) cited above, teacher factors were the largest single variable in the percentage of later educational outcomes attributable to schooling. In his review of school effectiveness research, Reynolds (1985) has indicated that teachers and the learner level are three to four times more powerful than the school level. The teacher is central not only in providing direct instruction in basic reading skills but also as the one who is responsible for creating much of the literacy environment in the school-age years. The resurgence of interest in Vygotsky has emphasised the importance of adult-assisted learning in promoting progress at the 'zone of proximal development'. This concept defines those functions that have not yet matured but are in the process of maturation – 'functions that will mature tomorrow but are currently in an embryonic state' (Vygotsky, 1978, p. 86). The quality of teaching in promoting learning and development in this way, and in nurturing the motivational context in which it is likely to flourish, is clearly of great significance.

Hopkins (1996) has outlined the attributes of effective teachers. They are clear about their instructional goals; they are knowledgeable about their content and the strategies for teaching it; they accept responsibility for pupil outcomes; they integrate their instruction with that in other areas; they communicate to their pupils what is expected of them and why, and offer regular appropriate feedback; they teach pupils meta-cognitive strategies and give them opportunities to master them; and they are knowledgeable about their pupils, adapting instruction to their needs.

The content of the curriculum

In relation to literacy, teacher variables are important, however, not only in terms of teaching style and quality but also in terms of the actual content of what is taught. At the heart of the early reading curriculum there must be structured and systematic phonics teaching, built on a sound foundation of competent phonological awareness. While there are still many aspects of these basic principles that require further research, their overall importance has been demonstrated for a very long period in a vast body of literature.

Phonological awareness plays a key role in the development of reading competence. The term refers to knowledge of the internal sound structure of spoken words. It is assessed by such tasks as deciding if two words rhyme or if they start with the same sound, and is reflected also in young children's knowledge of nursery rhymes. The strong relationship between these skills and learning to read has been shown in numerous studies (for example, Blachman, 1984, 1989; Bradley & Bryant, 1978, 1983; Bryant, Bradley, MacLean & Crossland, 1989; MacDonald & Cornwall, 1995; MacLean, Bryant & Bradley, 1987; Torgesen, Wagner & Rashotte, 1994; Wagner & Torgesen, 1987; Williams, 1986). These findings have been shown to be robust for phonological awareness independently of its association with general intelligence. Leaving aside the known effects of significant general learning difficulties, studies of children who learn to read early indicate that the relationship between IQ and early reading, while discernible, is not a strong one (Briggs & Elkind, 1973). Reading achievement overall in the early years of school is only weakly related to IQ

(Stanovich, Cunningham & Feeman, 1984), and children who have difficulty in learning to read often have above average IQ (Rawson, 1995).

The need for structured phonics teaching was for a long period the subject of a ‘great debate’ in reading (Chall, 1967). The opponents of phonics teaching supported a ‘whole language’ approach that viewed reading as a ‘natural’ task. Perhaps the most extreme, if by no means unique, representation of this viewpoint is to be found in the writings of Goodman, who was of the opinion that teachers make reading hard for children by breaking it up into small bits to make it easy – that is, by teaching phonics.

‘By isolating print from its functional use, by teaching skills out of context and focusing on written language as an end in itself, we make the task harder, impossible for some children’ (1986, p. 24).

Research, however, has consistently supported the view that reading is not acquired naturally, in the same way as speech, but rather that it is best viewed as an ‘unnatural act’ (Gough & Hillinger, 1980). Therefore, it is necessary systematically to provide instruction in those alphabetical building blocks that are essential to the reading process. A definitive body of literature has now supported this viewpoint (for a comprehensive review of the evidence see Adams, 1990; Stanovich & Stanovich, 1995). Indeed, the research base supporting the need for phonics teaching has been available for over 30 years (Bond & Dykstra, 1967; Chall, 1967). Despite the criticisms of the whole-language movement the phonics approach was never pursued in mechanical isolation from all the other strategies necessary for developing successful reading, such as listening to stories and exposure to a wide range of books. Nevertheless, the impact of the opponents of phonics teaching was widespread, and in many schools the key alphabetical principles were not taught in a structured way (Turner & Burkard, 1996). This is of particular relevance in the context of socio-economic disadvantage, where often children arrive at school with a very scant knowledge of print and the alphabet, compared with more advantaged pupils whose home environment has already exposed them extensively to the foundations of reading.

Paired learning

It is not only the teacher who may have a central role in supporting the learning of children during their years of schooling. Since the 1980s a substantial body of research has been accumulated on the subject of ‘interactive learning’ in which the focus of learning is a collaborative or interactive setting. This may be in the context of groupwork, peer-assisted or parent-assisted learning, or collaboration of students in interaction with computers (Foot, Howe, Anderson, Tolmie & Warden, 1994).

It is particularly the literature on paired learning that has made a significant contribution to literacy interventions. The paired learning approach has been developed both for children and adults (Topping, 1996). It has been applied to a wide variety of curricular areas including reading, spelling, writing, mathematics and thinking, not only for those experiencing difficulty in their learning but also for

students at all levels of ability (Topping, 1988, 2001). In this approach the individual is paired with a tutor, who may be a parent or other adult, or with a peer, in the form of same-age or cross-age peer tutoring.

Paired reading is a particularly attractive method for inclusion in literacy interventions, both because of its appealing and simple nature and because of its established effectiveness. In terms of its appeal, it has the advantage of using reading material chosen by the tutee in terms of its interest level. It takes place in a social and positive context in which praise for success is an essential ingredient. There are no negative approaches, such as saying 'No', and the tutee is not made to struggle or sound words out. It is therefore likely to be not only a rewarding experience but also an empowering one for the tutee. In terms of its simplicity, it does not depend on access to particular materials or reading schemes, and it can therefore be carried out anywhere, so long as interesting reading material is available. The method is straightforward and easy for both tutee and tutor to grasp. Reading material somewhat above the tutee's current level of competence is chosen, and the pair read together, with the tutor correcting any word read incorrectly and the tutee repeating it. When easier text is encountered the tutee signals to read without support, returning to a cycle of reading together again as required. Pairs commit themselves to an initial trial period of at least 15 minutes a day at least three times a week for eight weeks (Topping, 2001).

In terms of its effectiveness, the extensive reviews of the literature on paired reading by Topping and Lindsay (1992) and Topping (1995) have clearly demonstrated its value and have clarified the circumstances in which it is most likely to be successful. The consistent tendency of a large number of research studies and school projects has been to show gains both in reading accuracy and in comprehension. Attitudinal and affective gains have also been reported. Toomey (1993) reviewed nine controlled studies that used pre-post Neale Analysis scores. During an average intervention period of 2.7 months the mean gains in months for reading accuracy were 2.9 for controls and 6.5 for experimentals. Topping (2001) has reviewed more recent research including the 'Read On' project in Scotland, reporting paired reading gains for both tutees and tutors in comparison with controls. He reports a general picture in published studies indicating that paired readers progress at over four times normal rates in reading accuracy during the initial period of commitment, with good results sustained in follow-up studies.

Topping and Lindsay (1991) also considered aspects of paired reading in relation to socio-economic status. While their data were complex, they concluded that their results 'clearly give the lie to the claim that socio-economically disadvantaged families are difficult or impossible to recruit and deploy as effective tutors in parental involvement in reading projects' (p. 315).

Other factors associated with reading and reading failure

The traditional remedy for supporting children experiencing reading difficulties has been direct instruction, usually taking the form of individual or small group teaching, using multi-modal approaches and materials that are carefully planned for pace, level

and interest. While gains achieved through this type of instruction have often been recorded, the early literature on the effects of remedial education consistently indicated that gains were not sustained after the special help was withdrawn (Carroll, 1972; Topping, 1977). The individual approach returned to prominence with the claims of Clay's Reading Recovery programme, which has been evaluated in a UK setting by Wright (1992). Of pupils entering the programme with very poor reading and writing skills, over 96% reached average levels of attainment after a mean of 16.8 weeks of teaching. Reading Recovery, together with the general issue of intensive individual support, is considered further in Chapter 15.

A number of studies have focused on psychological factors such as self-esteem, motivation, attitudes and expectations, and these are particularly relevant in the context of socio-economic disadvantage. Thomas (1980) summarised the research on self-concept and educational achievement. The relationship was complex and difficult to interpret in causal terms. Nevertheless, he concluded that self-esteem is integral to academic performance, and an increase in the one may be expected to lead to an increase in the other. A review by Gurney (1987) examined controlled experimental studies that have sought to enhance children's self-esteem in classroom settings. He reviewed curriculum packages directly aimed at self-esteem enhancement, interventions designed to improve academic performance, counselling interventions and attempts to change teacher or pupil behaviour. Among factors considered to be important were the value of making positive statements about oneself and the advantage of working on both self-esteem and educational skills at the same time. In a randomised control trial MacKay (1995a, 1999b) focused on socio-cultural attitudes and values relating to reading. The reading gains reported following a short-term intervention were sustained several years after the completion of the study, despite the fact that no further special help had been provided (see Chapters 5 and 12, where these studies are covered in greater detail, together with other factors relating to this general area).

Literacy interventions in disadvantaged areas

Many interventions have been carried out to raise educational achievement in areas of socio-economic disadvantage. The readings edited by Cox (2000) provide both an overview of the range of interventions conducted and an analysis of the extent of the enduring obstacles encountered in effecting sustainable change in disadvantaged populations. Further comment is made here on two interventions of particular relevance to the present study – the High/Scope (originally called the Perry Pre-School) Project in the United States and the Pilton project in Scotland.

Several evaluations of the longer-term effects of early intervention programmes in the United States have shown good results (see Barnett & Escobar, 1990, for a review), but the High/Scope was the most carefully controlled early intervention research. It has been subject to evaluation for about 30 years and significant social and economic benefits have consistently been shown. The project adopted an experimental design with randomised assignment of children to early childhood education or control groups. Although initial IQ gains for experimentals were not sustained in later years, wider outcome measures at age 27 showed striking benefits

in social adjustment, educational achievement, earnings level, home ownership, lower dependence on social services and lower crime rates (Schweinhart, Barnes & Weikart, 1993).

It was not just the fact that the experimentals had received a pre-school programme that made the difference. Schweinhart, Weikart and Larner (1986) compared the effects of three different curricula: the 'active learning' curriculum used in the High/Scope Project, a traditional nursery group whose curriculum emphasised free play and a formal skills group whose programme was based on direct instruction. At age 15 the two play-based programmes (the active learning and traditional nursery groups) showed superior outcomes in terms of pro-social behaviour and school adjustment. The same sample was followed up at age 23 (Schweinhart & Weikart, 1997) with outcomes that demonstrated advantages for the play-based programmes in terms of emotional adjustment, educational achievement and anti-social or criminal behaviour. Few real differences were to be found between the High/Scope sample and the traditional nursery sample in these comparisons, although the trend of any differences did on the whole tend to favour the High/Scope children. It is also to be noted that the ultimate sample was relatively small (N = 65).

Sylva (2000) reported a wider application of the same comparisons within a European context. She described a study by Nabuco (1996) in Portugal replicating the Schweinhart and Weikart (1997) study, but with a larger sample (15 pre-schools in Portugal, N = 219 children, compared with one pre-school in the US, N = 65 children). Nabuco's work indicated that the High/Scope sample generally did best in various areas of adjustment such as self-esteem, with the formal skills sample again having the poorest outcomes. Significantly, in terms of the present study, they also did best in reading. Although the sample size was larger it should be noted that this was a relatively short-term comparison (start and end of first grade in Lisbon). Nevertheless, in combination with the smaller but long-term sample from the US, there was a consistent trend of outcomes in these studies. Sylva (2000) summarised this area by referring to three distinct outcomes from early education: cognitive skills underpinning academic learning, personal attributes, giving a child confidence, curiosity and perseverance, and social commitment, bonding the child to the community.

The Pilton Project (McMillan, Fox & Wood, 1994; Lothian Regional Council, 1995) deserves particular mention for several reasons. First, it was a Scottish study carried out in an area of multiple socio-economic disadvantage, and it is therefore of direct relevance to this study. Second, it was the first attempt in Scotland to carry out a multiple-component intervention to address reading achievement and reading failure in a population of this kind. Third, it had crucial political significance that has gone undocumented and therefore largely unrecognised. The 1994 study highlighted the disadvantages suffered by poorer socio-economic groups even in terms of access to the services needed to support their difficulties. This received very prominent national press coverage in Scotland which undoubtedly had a political impact. It is the view of this author that it was a crucial factor in the discussions that led the (then) Scottish Office Education and Industry Department to formulate its early intervention funding scheme, aimed primarily at addressing literacy in disadvantaged

populations. This scheme made an impact throughout Scotland in that it provided considerable funding to take forward projects for raising reading achievement throughout the country.

Although the Pilton project covered just four Edinburgh primary schools, its contribution lies in the fact that it identified a range of robust, research-based interventions for raising literacy levels and it had successful outcomes. It focused on supporting family literacy through the provision of home-link teachers. It provided a basis for improving nursery pupils' knowledge about books, letters, rhyme and phonology. Its most fundamental aim in the early years of schooling was to increase the amount of time children spent reading, and to support this by providing extra help in the classroom. In addition, it identified children who needed extra support and provided a remedial programme based on Clay's Reading Recovery and Bradley's phonological approach. All of these concepts and principles, together with the research base from which they had been developed, informed the interventions adopted in the current research.

It was the evaluation of the Pilton Project that made a deserved public impact. Careful comparisons of knowledge of rhyme, knowledge of letters, concepts of print, the alphabet, reading concepts, letter naming, word recognition and reading fluency were made from nursery through to Primary 4, comparing cohorts at these levels with their equivalent cohorts before intervention (as has been done in the main study for this research). All changes favoured the intervention cohort.

Implications for this intervention

The implications of the literature covering interventions in literacy for this study were clear. Any intervention must take account of and respect the socio-cultural context of the population for whom the programme was designed, aiming to empower rather than to control. It must foster family literacy and the more general literacy environment in school and community. It must recognise the process of educational change, and include a focus on developing effective schools and effective teaching, fostering leadership, vision and high expectations. It must address adequately the content of the curriculum, emphasising the need for good phonological skills and systematic phonics instruction. It must develop a context to encourage interactive learning, including the benefits of paired reading utilising family members and peers. It must identify and support children experiencing difficulties in their reading development, and build self-esteem and positive attitudes. While encouraging reading development in nursery it must promote a literacy-friendly environment in a context that respects play and active learning as crucial components for developmental progress in early childhood. Finally, it must ensure adequate time and additional support for reading in the early primary classroom.

SUMMARY

This chapter provides an overview of key literature on interventions in literacy. It highlights the importance of the socio-cultural context and the need for interventions

in areas of socio-economic disadvantage by comparing typical early development of literacy in different social settings. A wide range of factors affecting literacy outcomes are considered, including the role played by effective schools and effective teachers. The curricular content of interventions is presented as being crucial, and in particular the debate regarding the place of phonics instruction is examined. It is asserted that such instruction must be central to early teaching in literacy, together with promoting phonological awareness. Other factors such as the benefits of paired learning, the place of individual support for failing readers and the potential contribution of enhancing attitudes and self-esteem are outlined. Implications for the development of the intervention strategy in this study are considered.

Chapter 7

Baseline Assessment

This chapter provides an account of the development of a baseline assessment scheme for the main study. This is set within the perspective of the relevant literature on baseline assessment and also of the developing Scottish educational context.

Three principal requirements were established for the design of the baseline assessment scheme. First, as well as relating to general developments within the UK for baseline assessment, it had to be workable within the current Scottish political and educational setting; second, its content had to reflect assessment items of central relevance to the development of early literacy skills; third, it had to meet effectively the general requirements for the design of an assessment instrument.

The development of the baseline assessment in terms of these three requirements is discussed below, together with a description of the assessment measures adopted for the later years. A copy of the baseline assessment scheme is shown in Appendix 2.

The context of baseline assessment in Scotland and in the UK

Education in Scotland shares a common foundation with the rest of the UK. In terms of its legislation and practice it is based on the same philosophy of making adequate and efficient provision of education, including special education, and its progress and development similarly reflect the initiatives of central government in areas such as raising achievement and promoting excellence. Beyond this common foundation, however, the position in Scotland is different from the rest of the UK in many important respects (MacKay, 1995b). It is governed by separate Scottish education Acts which do not always correspond in timing or in purpose with Acts applying to England and Wales.

This context is of considerable importance in relation to baseline assessment. There is no Scottish equivalent of the Education Act, 1997, and baseline assessment therefore has no place within Scottish educational legislation. Nevertheless, the then Scottish Office Education and Industry Department (SOEID) took forward proposals for a national framework for baseline assessment and a national scheme was being piloted in several education authorities during the early stages of the research reported here (Scottish Office Education and Industry Department, 1998).

This pattern of developments is a characteristic one in Scotland and mirrors the arrangements made in relation to the curriculum. While there is no 'National Curriculum', there is a curricular framework for children in their pre-school year (The Scottish Office HMI, 1997) as well as 5-14 curriculum guidelines (for example, Scottish Office Education Department, 1992, 1993). It is within this curricular framework that the proposals for baseline assessment in Scotland were firmly embedded.

The plans for Scotland, however, were very different in one significant respect from the national framework initially endorsed for baseline assessment in England and Wales. The latter served to regulate local schemes and to allow flexibility. The Scottish framework was designed for use as a single scheme in all primary schools. This was in line with the recommendations of a review of literature and relevant baseline assessment schemes commissioned by the SOEID Research and Intelligence Unit:

‘A single national scheme is preferable to an accreditation process. National accreditation procedures promote an inefficient use of resources by encouraging each local authority to devise local schemes’ (Wilkinson & Napuk, 1997, p.28).

Subsequent developments almost saw the reversal of these differences between Scotland and the rest of the UK. In England and Wales by 2002, the Department for Education was moving towards the development of a more standardised national framework. Meanwhile in Scotland, the Executive announced their abandonment of the national pilot scheme described below in favour of more flexible local approaches. The scheme contained a number of weaknesses, some of which had been raised in the academic literature (MacKay, 1999a).

Early intervention in literacy

Prior to these national developments in baseline assessment, the SOEID in April 1997 invited education authorities in Scotland to submit projects for early intervention in literacy and numeracy in schools where educational underachievement was a prevalent and recurrent problem. Project proposals were received from all 32 authorities, and a high level of additional funding was made available to finance these through a three year period until March 2000.

The criteria for projects were that they should focus on children in Primaries 1 and 2 (ages 5 to 6 years), that they should build on pre-school early reading activity and that they should have systems for improving the flow of information from nursery to primary. The last-mentioned criterion implied a focus on the type of information that would be available about children at the time of school entry, and indeed it was a requirement of the projects that they should include measures for baseline assessment at Primary 1 and for subsequent re-assessment. This would provide a basis for the authority’s own evaluation of its early intervention programme. In response to these requirements, education authorities began to develop approaches to on-entry assessment of Primary 1 pupils. It was within this framework of early intervention projects that the concept of baseline assessment developed in most of the education authorities in Scotland.

Survey of Scottish education authorities

The design of a baseline assessment scheme for this study took account of these incipient developments elsewhere in Scotland. Since the national context for baseline assessment was changing rapidly it was also necessary that developments across the country should continue to be monitored. This would serve to ensure that the scheme

designed for the main study would continue to be workable and politically acceptable to the education authority, and to allow any necessary changes to be considered.

With a view to such monitoring, and following the launch of the draft national framework for baseline assessment, key personnel in all of the 32 Scottish education authorities were interviewed by the author in March 1998 (MacKay, 1999a). This survey allowed a picture to be built up of the current position in Scotland and also to gauge the views of those involved regarding the proposed national scheme. Respondents were asked questions covering four areas: first, whether they were carrying out any form of baseline assessment, and if so whether it was for literacy only or for the wider curriculum; second, whether the schemes they were using had been designed in-house, and if not, which commercially available schemes they had chosen; third, at what stage in the school session they were carrying out the assessments and the purposes for which they were being used; fourth, their knowledge and views regarding the national scheme.

All 32 education authorities were operating some form of baseline assessment scheme for their early intervention projects in literacy. Schemes showed a clear division in their nature between those designed mainly to monitor children's progress and inform curriculum planning and those designed mainly for value-added assessment. In 23 cases, these were in-house schemes that had either been devised by professionals working within the education service or had involved support from the universities and colleges. In 22 of the authorities, including those that were piloting the national scheme, there were also general baseline assessment schemes covering the wider curriculum, and in 10 cases these had been devised in-house. No fewer than 17 commercially available schemes were also being used, with some authorities adopting a mixture of these. The only commercial scheme reported frequently enough to merit individual comment was the *Early Years Easy Screen (EYES)* (Clerehugh, Hart, Pither, Rider & Turner, 1994), used in seven authorities.

The survey indicated that most authorities had consciously opted for one of two basic approaches to baseline assessment. The first was pupil profiles based on teacher observations, designed mainly for the purpose of monitoring pupils' progress and curriculum planning. The second was a more formal approach to individual assessment, designed mainly for value-added purposes. While these approaches were clearly not mutually exclusive, they did tend to dictate the general structure and content of the schemes. In a number of cases the former was adopted for general baseline assessment across the curriculum, while the latter was chosen to provide quantitative evaluation of the early interventions in literacy.

The Scottish national scheme pilot study

It was in parallel with these developments in education authorities, that the SOEID launched its pilot procedures for a national baseline assessment scheme. The main purposes of the scheme were similar to those outlined by the School Curriculum and Assessment Authority (1997): first, to identify the principal features of individual children's progress and achievement, in order to plan appropriate future provision for them; second, to identify levels of attainment in 'basic' aspects of learning for groups

of pupils, in order to be able to evaluate effectiveness of provision and subsequent 'rates of gain', and set targets for their improvement. In other words, the scheme was designed to meet the needs both of curriculum planning and of value-added assessment.

In relation to the first of these purposes the national scheme, as well as seeking to ensure that more able pupils were suitably challenged, attached particular importance to the use of baseline assessment in identifying those who needed special help to develop skills essential to effective learning. In relation to the second, the aim was to provide evaluation based on aspects of performance agreed as being important for learning, using methods of assessment and forms of reporting that were reliable, valid, manageable and clearly understood by users.

Assessment was based on structured observation by teachers of identified 'key aspects' of learning, drawn from the curriculum framework for the pre-school year and the 5-14 curriculum guidelines. As well as literacy and numeracy, these key aspects covered personal, emotional and social development, physical co-ordination, expressive communication and understanding the environment. Suggestions were provided on a variety of appropriate contexts in which each of the key aspects could be assessed. An entry was then made to indicate which of four 'attainment statements' most closely matched the child's level of performance, ranging from displaying 'very few' to displaying 'almost all' of the features listed.

Three of the key aspects used in the scheme were most closely related to literacy. These were 'listening and talking', 'reading' and 'writing'. The features for reading and writing are listed in Boxes 7-1 and 7-2.

Box 7-1 Baseline assessment in Scotland pilot procedures: Reading (*Baseline Assessment in Scotland: Pilot Procedures*, SOEID 1998)

Key aspect: Reading

- Knows how to hold a book and turns the pages appropriately
- Knows the conventions of the layout of English language texts
- Knows the difference between letters and words
- Understands the difference and relationship between print and illustration
- Selects text(s) appropriate to purpose
- Uses a range of strategies to 'read' familiar text
- Recognises own name, signs, labelling in the environment
- Identifies letters of the alphabet in upper and lower case by shape, name and sound
- Knows by sight the most common words in use
- Sounds out phonically regular words
- Identifies words which start/end with same or different sound(s)
- Reads continuous text aloud with reasonable fluency
- Shows understanding of the content and structure of text by retelling/predicting content

Box 7-2 Baseline assessment in Scotland pilot procedures: Writing (*Baseline Assessment in Scotland: Pilot Procedures*, SOEID 1998)

Key aspect: Writing

- Makes marks on paper to 'write' ideas
- Composes own stories through drawing, scribing or other means
- Copies and forms letters and words legibly and correctly
- Writes independently to label, recall, describe and record
- Uses punctuation to mark sentences
- Spells correctly most commonly used and phonically regular words

The same 'attainment statement' record was used for all of the key aspects. This is shown in Box 7-3.

Box 7-3 Baseline assessment in Scotland pilot procedures: 'Attainment statements' (*Baseline Assessment in Scotland: Pilot Procedures*, SOEID 1998)

Attainment statement	
<input type="checkbox"/>	The child displays very few of these features. Immediate investigation and structured intervention are essential.
<input type="checkbox"/>	The child displays some of these features. Most others require attention and planned support.
<input type="checkbox"/>	The child displays the majority of these features successfully and is making good progress with most of the others.
<input type="checkbox"/>	The child displays almost all of these features consistently and with confidence

Starting in 1998 the scheme was piloted in six education authorities with children in their pre-school year during the month of May, and in another seven authorities with approximately 1,000 Primary 1 children during March.

Issues arising

The Scottish national baseline assessment scheme had a number of strengths. The most significant of these was its central relevance to the educational process, since it not only related to the curriculum but was altogether derived from it. It was also broad in scope. It covered the whole curriculum and in no way restricted the scope of education in order to meet narrow assessment indicators. In terms of demands on teacher time it was manageable, and it encouraged assessment while avoiding some of the difficulties associated with formal testing. The method of assessment built on teachers' knowledge of their children, and development of the scheme involved ongoing consultation with researchers in this field. As a scheme that was thoroughly

embedded in the curriculum it was likely to enhance the teaching of literacy skills by raising teachers' awareness and by providing a basis for assessing the next step in each pupil's learning and planning appropriate activities.

The scheme provided a record of progress in a large number of skills relevant to achievement in reading. The three key aspects of listening and talking, reading and writing included a total of 28 items, each of which could be assessed in a variety of different contexts. Although in the form of a checklist, it was not a simple 'yes-no' format and, if used consistently, the attainment statements gave a range of options for indicating level of performance. Within the limitations of a schedule of this kind, and subject to ongoing evaluation, the scheme therefore provided a useful basis for assessing on-entry attainments in literacy skills both for those at the pre-reading stage and for those who have acquired the rudiments of reading.

Nevertheless, the national scheme posed significant issues both for education authorities in general and most particularly for the current study. First, while providing an overall picture of school entry skills in literacy and other aspects of the curriculum, the scheme was not adequate in giving the measures that would allow progress to be assessed precisely or the impact of interventions to be evaluated. Although one of the purposes of the scheme was to allow value-added assessment, the view generally expressed by the SOEID on this matter was that the 'jury is out', and that the scheme might or might not serve this purpose well. However, it was not possible to see how the scheme as it stood could meet a specification for reliability and sensitivity that would allow it to be a suitable tool for this purpose.

The problems in using baseline measures for value-added assessment are widely recognised (for example, Lindsay, 1998). It may be that there is a simple choice between using broadly-based observational profiles that are useful for informing the educational process but are insensitive as value-added measures, or using narrower assessment instruments that are likely to have high value-added reliability. The former are more easily carried out by teachers as part of the day-to-day work of the classroom, while the latter, such as the scheme used in this study, require more structured individual assessment and are therefore likely to have greater resource implications.

A second and related difficulty of the national scheme was in using it to identify the difficulties of pupils who were not progressing as expected. Although the scheme provided a record that would assist teachers in the identification of such pupils and in planning appropriate assessment, it was not sufficiently detailed to yield more precise information, and to show the nature and extent of change in performance over time. Most education authorities were carrying out interventions designed to address the needs of vulnerable pupils, and a different type of baseline assessment was therefore required. For the purposes of the research reported here it was particularly essential that pupils with difficulties could be identified using measures sensitive to change. As one of the key aims of the project was the eradication of illiteracy, measures had to be available that would clearly demonstrate the performance in key areas of children who were failing.

A third difficulty with the national scheme was raised very frequently in the survey of Scottish authorities described here. While the efficiency of having a single national scheme was recognised, there was considerable concern about this being imposed on all education authorities. Having been encouraged by the government to develop baseline schemes for their early intervention projects, many authorities had already endorsed schemes to which they had a high level of commitment. While it would have been possible to operate the national scheme for general baseline assessment, and use local schemes in addition for evaluating early intervention, this would have been quite impracticable in those areas where the early intervention programme covered most schools in the authority.

This was the main difficulty the introduction of a national baseline assessment scheme raised for the current study in its early stages. It posed for the education authority the dilemma as to whether it might be expected to use the national scheme, either instead of or in addition to the scheme designed for the project. Using it as an alternative was not an option in terms of robust evaluation of the intervention. The pre-intervention baseline had already provided a measure of achievement for every child in the study, and the national scheme had neither the comparability, the specificity nor the sensitivity to replace it. Using it as an addition was also a problem. It would require staff to be assessing the same children at different times in the session on two quite disparate frameworks. In the event, the education authority opted to continue with the scheme designed for the study, and not to introduce the national baseline assessment scheme.

These national issues are of particular relevance to this study. The survey of Scottish authorities ultimately served two key purposes in taking forward the project. First, it demonstrated the continuing suitability of the baseline assessment scheme at a time when there might have been a political imperative to replace it with the national scheme. The survey provided key information on how baseline assessment was being addressed throughout the country, and established a sound basis for the view that the scheme used for this study should continue. It was firmly embedded in the curriculum and was appropriate to the Scottish educational context. Second, it helped to inform the debate on the future of the national scheme, and the more flexible approach based on local arrangements that replaced it.

The baseline assessment scheme: establishing relevant content

Many baseline assessment and other screening instruments have been available in education over the last 30 years (Blatchford & Cline, 1992; Lindsay & Deforges, 1998). Among the most widely used is the *Infant Index* (Deforges & Lindsay, 1995; Lindsay & Deforges, 1999). This was developed from the earlier *Infant Rating Scale* (Lindsay, 1980) and comprises 15 items covering literacy (reading, writing, spelling), mathematical skills, social behaviour and independent learning. Each item is scored against explicit criteria on a scale of 0-3. The item for reading has three elements: shows an enjoyment of books, and knows how books work (front/back, left/right, top/bottom), can recognise individual words and letters in familiar contexts and can read from a simple story book. During recent years computerised schemes have also come to be widely used, such as the *CoPS Baseline Assessment System*

(Singleton, Thomas & Horne, 1998). This scheme comprises 56 items grouped into eight different skill/concept domains including literacy.

The baseline assessment scheme used for this study was designed to be an effective instrument for assessing pre-reading and early reading skills in children in their pre-school year and in Primaries 1 and 2. It therefore focused on items that had been shown in the overall research on literacy to reflect reading ability in this age group, to identify literacy difficulties and to have high predictive validity. In doing so it contained items that reflected areas almost universally assessed in other baseline schemes, such as concepts of print and letter recognition. Both of these items featured, for example, in the widely-used schemes described above. Box 7-4 provides a summary of the structure of the scheme.

Of the four sections comprising the assessment, only the first three – concepts of print, phonological awareness and early reading skills – were used for the purposes of the present study. The fourth section, developmental tasks, was designed to provide measures of a wider range of related skills relevant to further research arising within the overall project. Altogether, the scheme provided measures in 13 different areas.

Box 7-4 The baseline assessment scheme

Section 1 Concepts of Print

Section 2 Phonological Awareness

- Test 1: Nursery Rhymes
- Test 2: Initial Letter Sounds
- Test 3: Rhyme Detection
- Test 4: Rhyme Production

Section 3 Early Reading Skills

- Test 1: Lower Case Letter Sounds
- Test 2: The Alphabet
- Test 3: Letter Names
- Test 4: Non-word Reading Test
- Test 5: Word Reading Test

Section 4 Developmental Tasks

- Test 1: Copying Forms
- Test 2: Writing Name
- Test 3: Draw-a-Person

The scheme had its origin in work piloted for the purpose of evaluating earlier literacy research with Primary 1 children, the Edinbarnet Early Reading Project (MacKay & Watson, 1996, 1999), as outlined in the account given of the preparatory studies. Of the 13 tests in the full baseline assessment, five were either used in, or were modified from, the earlier project. Of these, two were sufficiently universal in their use and application that they would be expected in any scheme of this kind –

Section 1, Concepts of Print, and Section 3, Test 1: Lower Case Letter Sounds. The other three were: Section 2, Test 1: Nursery Rhymes, Test 2: Initial Letter Sounds and Test 4: Rhyme Production. One other test in the previous assessment, 'Sounding and Blending', provided a basis for developing Section 3, Test 4: Non-word Reading Test.

Section 1, Concepts of Print, was taken as an essential starting point for the assessment of emerging literacy skills in young children. Progress in learning to read may be viewed as a series of stages of which the earliest can be described as learning associations between visual features of graphic forms (not complete orthographic word forms) and spoken words (Chall, 1983; Ehri, 1991; Gough and Hillinger, 1980). This idea of 'emergent literacy' (Clay, 1991) characterises young children as developing concepts about the components of literacy. The idea of a 'Concepts about Print' test, designed to test children's knowledge of the nature and function of written text, was developed by Clay (1979a). Assessing these emergent concepts not only provides information on current early literacy abilities but also taps into abilities that have predictive validity. In a longitudinal study, Tunmer, Herriman and Nesdale (1988) assessed children at the beginning and end of their first year of schooling, and again at the end of their second year, and found a strong relationship between concepts of print and reading achievement.

The relevance of the tests in Section 2, Phonological Awareness, to the development of reading has been outlined in Chapter 6. A significant body of evidence has shown that phonological awareness is a good predictor of later reading abilities (for example, Blachman, 1984, 1989; Bryant, Bradley, MacLean & Crossland, 1989; MacDonald & Cornwall, 1995; Torgesen, Wagner & Rashotte, 1994; Wagner & Torgesen, 1987).

The Early Reading Skills tests in Section 3 are central to later reading development. Knowledge of the alphabet and of the relationship between letters and their sounds, followed by skills in blending and in recognising words are all part of the essential elements in becoming a skilled reader (Adams, 1990; Mason, 1981, 1985). In particular, knowledge of letter names has been identified as the best single predictor of early reading achievement (Chall, 1967; Bond & Dykstra, 1967; Riley, 1996; Vellutino & Scanlon, 1987).

The baseline assessment scheme: properties and design characteristics

The third requirement of the baseline assessment, that it should meet effectively the general requirements for the design of an assessment instrument, was addressed with reference to seven basic principles of measurement outlined by Bloom and Fisher (1982) – reliability, validity, utility, directness, reactivity, sensitivity and feasibility.

Reliability

In terms of reliability, the baseline assessment scheme already drew from such a large body of existing knowledge that there was from the start a high level of confidence in its stability and consistency. As already detailed, it utilised several of the tests previously used in the Edinbarnet Early Reading Project. These tests had already proved themselves to have high levels of reliability. In addition, it brought together a number of well-established measures, such as concepts of print, letter sounds and word reading. All of these in a variety of formats had been used with high reliability in literacy research in general, and indeed one form of the word reading test included earlier standardisation data with high reliability figures derived from a Scottish population in the 1970s (The Burt Word Reading Test). In addition, feedback from the field from a wide range of users working with very large samples had supported the view that this scheme offered a stable and consistent measure. This included allocation of different testers to share assessment across the same classes. Informal analysis of the pattern of results in these circumstances again showed completely consistent patterns of performance.

Two steps were taken to provide confirmatory data regarding reliability. First, the test forms of a small sample of just over 100 children were analysed for internal consistency, splitting the test by alternate items. This was done at three levels: Sections 1 and 2: Concepts of Print and Phonological Awareness; Section 3: Early Reading Skills; Sections 1, 2 and 3 combined. This covered the 10 core tests in the scheme, but with the omission of Section 3, Test 2: The Alphabet, which did not lend itself to splitting. High levels of split-half reliability were obtained, as shown in Table 7-1.

Table 7-1 Baseline assessment scheme: internal consistency

N = 108	Reliability co-efficient
Concepts of Print/ Phonological Awareness	0.83
Early Reading Skills	0.94
Whole Test	0.89

Second, as a later check on external consistency, the total test scores across the 10 baseline tests of 100 children who took the test in Primary 1 in 1998 were correlated with the scores of the same 100 children in Primary 2 in 1999. A correlation coefficient of 0.82 was obtained.

Validity

The scheme clearly had high face and content validity. In terms of face validity, it measured the precise skills that the intervention was designed to enhance in relation

to pre-reading and early reading abilities. In terms of content validity, it adequately provided a representative sample of these skills, and in the case of Letter Sounds, Letter Names and The Alphabet it contained the entire population of items. It was also expected that the test would have good predictive validity. As outlined above, it drew extensively from those items that had been shown to be good predictors of later reading achievement.

For the purposes of the work reported here it was not intended to conduct any extensive analysis of the properties of the assessment in relation to its predictive validity. However, the duration of the project to date has resulted in data relevant to this purpose, and the following checks using small samples were carried out to allow some comment to be made. First, some pupils who took the Norman France test in P4 in May 2002 were traced back and identified in their original baseline assessment scores from November 1997 when they were at nursery. Second, the same was done for some pupils who took the Norman France test in P7 in May 2003 in comparison with their P2 baseline assessment scores from 1997.

The results of these first two comparisons are shown in Tables 7-2 and 7-3.

Table 7-2 Baseline assessment scheme predictive validity: Correlation with Norman France at P4 and P7

Year groups	Test items	Results
Pre-school (1997) v. P4 (2002) Norman France N = 21	Concepts of Print/ Phonological Awareness	Pearson's $r = 0.55$
P2 (1997) v. P7 (2003) Norman France N = 116	Early Reading Skills	Pearson's $r = 0.43$

The pre-school baseline scores for Concepts of Print plus Phonological Awareness showed a correlation of 0.55 with the P4 group test scores almost five years later, using just a small sample of 21 children from one nursery.

Table 7-3 Baseline assessment scheme predictive validity: Mean Norman France scores at P7 (N = 116)

Test items	1997 P2 baseline quartiles	2003 P7 Norman France mean scores
Early Reading Skills	1 st quartile	8y 4m
	2 nd quartile	9y 3m
	3 rd quartile	9y 4m
	4 th quartile	9y 8m

With a sample of 116 for the comparison over almost six years, the correlation between P2 baselines for Early Reading Skills and P7 group tests was 0.43. When

the group tests scores were associated with baseline quartiles, a clear gradient was shown, with a range of 1 year 4 months in mean scores at P7 between pupils scoring in the lowest quartile in P2 and those scoring in the highest quartile.

A third comparison was made using the results of the P7 pupils identified by schools as having significant reading difficulties at the end of 2002. These pupils were traced back to see how well their original baseline assessment scores identified their later difficulties. All of these pupils were tested individually on the Neale Analysis of Reading Ability. Of those who fell below a cut-off score of 9y 6m, 53 were able to be traced in the 1997 P2 baseline assessments. A summary of the results, which are discussed in greater detail in the individual support study (Chapter 15) is shown in Table 7-4. Four-fifths of these pupils fell in the 1st quartile of baseline assessment scores for Early Reading Skills, with approximately half of them being in the lowest 10% of scores, and a quarter in the lowest 5%. Only one of these pupils scored above the second quartile, reaching the 52nd percentile. Except for the lowest 5% of scores, phonological awareness was not such a good predictor of later difficulties as early reading skills. Phonological skills are in general well established by Primary 2, and only the poorest children were scoring much below ceiling level on these tests.

In summary regarding validity, the baseline assessment scheme had clear face and content validity, and in predictive terms there was also a clear relationship between early baseline scores and performance on other measures of literacy up to six years later. However, it is noted with caution that the lowest of the correlations cited, that of 0.43 between baseline scores at P2 in 1997 and Norman France group test scores at P7 in 2003, accounts for only one-fifth of the variance.

Table 7-4 Baseline assessment scheme predictive validity: Identification of children with later reading difficulties (N = 53)

1997 P2 baseline tests	Distribution of baseline scores of children identified with reading difficulty at P7 (Cumulative scores)		
	Lowest 5%	Lowest 10%	1 st quartile
Phonological Awareness	23%	26%	51%
Early Reading Skills	25%	47%	79%

Utility

The question of utility was addressed by considering: does the scheme give the information that is required for assessment and intervention planning? Is it relatively easy to produce and to use? Is it suitable for the resources available in terms of the financial and organisational implications of staffing, time and production costs? Will

there be significant training implications to ensure proper use, and will those using it accept it as something they are happy to work with?

All of these issues were considered carefully in designing the assessment. The information provided was seen to be precisely what had to be known in assessing children's early literacy progress and planning an intervention programme. Ease of use and low production costs were a key factor. It was therefore designed to be simple, inexpensive and self-contained. All of the instructions for use, including the wording used by the tester, were clearly provided on the test paper itself, so that there was no requirement to have a separate user's manual. Production costs comprised the provision and photocopying of a few sheets of paper, with no special equipment, packaging or supplementary items. The word reading test, although long out of use as a standardised instrument, was already available within the authority in the form of many hundreds of re-usable tests printed on card. These had once been used routinely with almost every child referred to the psychological service, and had lain there, most of them unopened, after the test was overtaken by newer instruments and when reading tests were no longer the routine response to referrals.

The simplicity of the baseline assessment also meant that training in its use was very manageable. This was provided on the basis of a single half-day session attended by groups of up to 40 testers at a time. These sessions were provided each year for any new testers, and back-up support was immediately available to deal with any difficulties or queries. The scheme was also so completely embedded in the early literacy curriculum that it was fully acceptable to teachers and others in early education. Also, resourcing was committed to ensure that it could be carried out with every child each year, with data processing arrangements in place immediately after the testing period.

Six years of use have fully proved the utility of the baseline assessment scheme. Between 3,000 and 4,000 children have been individually tested each year. Schools have welcomed it as a useful assessment instrument that gives them valuable information on their children irrespective of its need in evaluation of the project. It has become an essential tool in raising teachers' awareness of their pupils' progress, in planning appropriate curricular activities, in providing key assessment measures of early reading skills and in identifying the needs of children who are experiencing difficulties in literacy. Its simplicity has been such that training sessions routinely raise almost no questions about its use and there has been no real demand made on the back-up support for answering queries about its administration. Also, constant use by literally hundreds of testers has pointed to no significant hitches in its design, and only the most minimal changes have been made to the original format to assist ease of recording scores.

Directness

Direct measures essentially are those that most clearly approximate to, accurately reflect or actually are the real situation, problem or behaviour of concern. They involve making no inferences, or only minimal inferences, about the data. Indirect measures, which are common in psychology in general, such as personality

inventories, are primarily defined by the fact that inferences by the tester are required in order to interpret results. The benefits of direct measures where they can be constructed, and especially in relation to repeated testing of performance in specific areas, has been reviewed by Hersen and Barlow (1976). Direct measures tend to have better validity, to be more accurate and to provide more useful data for planning and evaluating interventions.

The measures used in the baseline assessment were extremely direct. They all measured the areas of performance that the intervention was designed to enhance. They either provided representative samples of these areas, such as the non-word and word reading tests, or they contained the entire population of testable items, as on lower case letter sounds, letter names and the alphabet.

Reactivity

Reactivity refers to the changes that come about in the child's performance through the act of measurement itself. In other words, the process of assessment itself leads to changes in performance without any other intervention taking place. Generally speaking therefore, the aim of any assessment instrument is to have low reactivity. In general this aim applied to the design of the baseline assessment, except in one important respect, as detailed below.

A review of the types of reactivity affecting test design and administration is provided by Webb, Campbell, Schwartz, Sechrest and Grove (1981). These include: the guinea-pig effect, when the measurement process itself engenders responses or behaviours that apply only in the test situation; role selection, where those being tested assume a role different from what they do in real life situations; response set, where the design of the assessment leads to a response bias, for example, of tending to select answers in the same position each time; and interviewer effects, where the results of the assessment are affected by the characteristics of the interviewer.

In assessment of educational performance it is only the latter two of the above-mentioned types of reactivity that are likely to apply, and steps were taken to ensure low reactivity in these areas. The layout of items generally avoided response set; testers were trained to adopt a standard approach, as clearly detailed on the test papers; and the practice at times of splitting classes between two different testers, together with the feedback from a large number of testers over several years, has given confidence that reactivity in these areas is low, and that the results give an authentic demonstration of children's performance.

One other type of reactivity listed by Webb et al. (1981) is in a somewhat different category, namely, measurement as a change agent. That is, the process of measurement brings about change in the behaviour of the person tested or of the system. In this project, the use of measurement as an agent of change was quite deliberate. In short, the assessment procedures represented not only part of the evaluation but part of the intervention. The idea was that the introduction of this formal structure of individual assessment would do two things. First, it would raise awareness in the child of how important and how central the reading process was,

and would bring the key aspects of literacy in the test to the forefront. Second, it would also raise teacher awareness and would alter teacher behaviour, by turning attention to the very areas requiring to be taught as the core of the reading curriculum. Thus, one of the 10 strands of intervention detailed elsewhere in this work was ‘to raise teacher and pupil awareness through focused assessment’.

One of the criticisms that may be made of a test of this kind – especially in circumstances where there is close monitoring of results and detailed feedback to schools regarding their own performance and that of the authority overall – is that teachers will start ‘teaching to the test’. There are two respects in which this would be undesirable. First, it was not wished that schools would so divert attention from the rest of the curriculum that the central question became one of how to get good test scores. Second, it was obvious that the test would not remain objective if, for example, the specific blends selected for the non-word reading test were taught rather than blending in general, or if the children were coached in reading and remembering the exact words in the reading test. One experiment reported by Gough and Juel (1991) illustrates this. Children learned to recognise a word placed on a card with a thumbprint on it. When the thumbprint was absent so was the recognition. Similarly, the children could have learnt words from the reading test in the baseline assessment by rote memory. Careful monitoring procedures were in place to deal with this.

However, there were other ways in which the whole idea was to teach to the test. For example, with the lower case letter sounds, so crucial to early reading development, the test comprised all 26 letters. A teaching focus on learning these 26 letters was therefore encouraged as being fundamental to future progress. Assessment, therefore, was one of the agents of change.

Sensitivity

It is essential for the purposes of evaluation that the measures used should be sensitive to change. Many measures used widely for educational assessment do not have the sensitivity to meet this criterion. For example, the national testing within the 5-14 curriculum gives a broad range of ‘levels’ in the different core subject areas such as reading and writing. While measures like this can be used to show differences between schools or across years, the actual levels are too global in nature to be responsive to the detailed changes taking place as the result of an intervention.

The baseline assessment was designed to be a very sensitive instrument. It had 10 core tests, and each of these gave carefully graded scores on the number of items passed. It was designed with very low floors and high ceilings so that the same format could cover the whole period from the pre-school year to the end of Primary 2, with every child achieving some success and no child scoring on every item. During its six years of use it has demonstrated in precise terms a clear and consistent pattern of annual change.

Feasibility

The last criterion, feasibility, is seen by Bloom and Fisher (1982) as a summary of the others, as it touches to some extent on all of them. It addresses issues such as: is the instrument sufficiently stable over a period of time that it does not show changes just through repeated use? Does it take an excessively long time, leading to boredom or decreasing returns on the part of testers or participants? Does it give the potential of providing assessment information for evaluation purposes? Is it a meaningful indicator of the area being addressed, so that the information it gives can be utilised? Will it become out of date over the time of the intervention and need revision? What are the implications for large-scale data processing when used year on year with thousands of cases? Is it likely to have the continuing commitment of funders to resource its use over a long period of time?

All of these questions have been answered in the course of very extensive use of the baseline assessment by large numbers of people administering it every year. It has not required revision with the passage of time and the information it provides has come to be seen as a routine and necessary part of ongoing assessment in the early years. It has not only had the continuing commitment of funders but has been endorsed as a valuable assessment tool in its own right, and not simply as an evaluation measure for the effects of the intervention.

SUMMARY

This chapter describes the development and design of the baseline assessment scheme that served as the main evaluation instrument for this study. It emphasises the crucial importance of taking account of the wider context of policy and practice nationally and locally in designing any scheme of this kind, and reports a survey of all 32 Scottish local authorities undertaken to inform the way in which baseline assessment should be developed. The rationale for the content of the scheme is outlined, and the content described in terms of the three main sections of the scheme: concepts of print, phonological awareness and early reading skills. The properties and design characteristics of the scheme are covered under the seven headings of reliability, validity, utility, directness, reactivity, sensitivity and feasibility. It is concluded that the scheme has demonstrated its usefulness not only for the primary purpose of evaluation of the study but also in being endorsed by schools as a valuable assessment measure in its own right.

Chapter 8

The Main Study: Rationale and Design

Introduction

The main study took the form of the design, implementation and evaluation of an early intervention programme for all pupils in the pre-school year, Primary 1 and Primary 2 in all of the pre-five establishments and primary schools in West Dunbartonshire. The ages of these pupils ranged from under four years at the beginning of their pre-school year to just under seven years at the end of Primary 2. This provided a total sample each year of between 3,000 and 4,000 pupils at these stages. After the first year, one third of this number left the early intervention programme annually as they progressed from Primary 2 to Primary 3, to be replaced by a similar number of new cases entering the programme in the pre-school year. All of these pupils were individually tested at the end of each calendar year using the baseline assessment scheme designed for the programme (Chapter 7).

The sources of funding for the study from 1997 were the Scottish Office Education and Industry Department (later becoming the Scottish Executive Education Department) through its early intervention programme and its excellence fund, and West Dunbartonshire Council. This funding supported all aspects of the study, but most particularly the key human resources to provide extra help in the classroom in the form of 10 early intervention teachers, plus home support for which four home-link teachers were employed. It also provided for an extensive training programme for staff working on the project at all levels.

Aims of study

The aims of this study were:

- to raise the literacy levels of all children in the pre-school year, Primary 1 and Primary 2 in all schools throughout the authority using a multiple-component intervention
- to provide a basis for long-term improvements in literacy levels in the later years of schooling
- to reduce the numbers of children experiencing reading failure as a basis for eradicating illiteracy throughout the authority.

Hypotheses

The following hypotheses were proposed:

- 1 that the year groups of children receiving the intervention programme would have higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level

- 2 that these gains would be sustained after the children left the programme at the end of Primary 2
- 3 that both high and low achievers would show gains, but that in particular there would be significant reductions in the numbers of children experiencing reading failure in the early years.

Rationale: the '10 strands' of intervention

From the research literature on effective interventions in literacy, 10 strands were selected as the basis for the programme. All of them were of central importance to the intervention. Of these 10 areas, seven were identified as 'key strands' – that is, they were planned and structured in a formal way as the basis on which the programme would operate. These were: phonological awareness and the alphabet; a strong and structured phonics emphasis; extra classroom help in the early years; raising teacher awareness through focused assessment; increased time spent on key aspects of reading; identification of and support for children who are failing; and home support for encouraging literacy. The other three strands, while also being viewed as crucial to successful intervention, were promoted in a less formal and structured way, and were identified as being 'supporting strands'. These were: fostering a 'literacy environment' in school and community; lessons from research in interactive learning; and changing attitudes, values and expectations.

An overview of the research basis for interventions in literacy has been provided in Chapter 6. This is considered further here in specific relation to each of the 10 strands. Additional literature relating to three of the strands is discussed in the chapters describing the subsidiary studies. Strand 2, a strong and structured phonics emphasis, is developed further in Chapter 10; strand 7, identification and support for children who are failing, is developed further in Chapter 15; and strand 10, changing attitudes, values and expectations, is developed further in Chapters 12 and 13.

Strand 1: Phonological awareness and the alphabet

The importance of phonological awareness as a predictor of success in learning to read has been noted by many researchers. Over the last two decades there has been an emerging consensus about the significance of phonological processing abilities in the acquisition of early reading skills (Torgesen, Wagner & Rashotte, 1994). MacDonald & Cornwall (1995) carried out an 11 year follow up of 24 children who were first assessed at school entry. The results indicated that phonological awareness was a relatively stable predictor of word identification and spelling skills.

As well as predicting success in reading, it has been shown by many researchers that training in phonological awareness at pre-school level, and in particular rhyme and alliteration, leads to significant improvements in reading in the early stages of school (Ball & Blachman, 1988; Bryant & Bradley, 1985; Olofsson & Lundberg, 1985; Tunmer, Herriman & Nesdale, 1988; Wagner & Torgesen, 1987; Williams, 1986). Phonological training has also contributed to the remediation of problems for children who have had difficulty in learning to read (Olson, Wise, Ring & Johnson, 1997).

This was therefore built in as a key strand in the main intervention, especially at the pre-school stage. Staff in nurseries and the early primary stages were provided with materials and training for raising phonological awareness, supported by the early intervention team.

Strand 2: A strong and structured phonics emphasis

The importance of systematic phonics instruction has been emphasised in the overview in Chapter 6 and is developed further in Chapter 10 in the discussion of analytic and synthetic phonics. The evidence in favour of phonics teaching is so overwhelming that it has essentially become an area in which researchers have reached a consensus (Stanovich & Stanovich, 1995). Because of the absolute centrality of this area it was selected as the subject of one of the supporting studies, the synthetic phonics study (Chapter 11).

Phonics teaching was made a key strand of the intervention from the beginning. Strategies for preparing pre-school children in their knowledge of phonics were built into the nursery programme in a manner that was consistent with the agreed aims of the pre-five curriculum. That is, the nurseries were not turned into early classrooms for the systematic teaching of reading skills, but their foundations in play, social co-operation and active learning continued to be paramount. However, in addition to story-reading, the use of rhyme and alliteration and other reading readiness skills, children were encouraged in their routine nursery activities to have an enjoyable introduction to phonics. Considerable use was made of 'environmental print' to identify a wide variety of items, and children learnt the initial letters of their own names and were also exposed to those of other children. These and a variety of other strategies were carefully planned to encourage phonics knowledge so that they would have a suitable start when they went to primary. In primary school, good teaching of phonics was essential to the project, whether the schools were using traditional or synthetic methods.

Strand 3: Extra classroom help in the early years

In low socio-economic areas, pupils are likely to begin school having missed out on perhaps thousands of hours of literacy-related activities that take place in higher-income homes (Adams, 1990). Often the level of instruction they need has to be geared to those early activities to do with establishing basic concepts of print and phonemic awareness that already are second nature to other children. It is almost axiomatic to say that this process will be enhanced with increased levels of support in the classroom. For disadvantaged pupils the link between explicit instruction and achievement is notable (Brophy & Evertson, 1978; Good & Grouws, 1975; Stallings, Robbins, Presbrey & Scott, 1986).

Providing additional classroom support in the early years was the key intervention strategy in the Pilton project (Lothian Regional Council, 1995; McMillan, Fox & Wood, 1994). This was done by deploying classroom assistants in all early years classes to assist teachers in literacy instruction, either by working directly with

groups of children or individual children, or by supervising pupils to free the teacher to work with small groups. It was also a central strand in the early intervention work of MacKay and Watson (1996, 1999), as outlined in the preparatory studies in Chapter 5.

However, it is clear from the conflicting evidence in this area that the question of extra help in the classroom is one that must be considered carefully, as there is not a basis for believing that the simple expedient of providing extra help will automatically bring benefits. This may be illustrated by detailed consideration of one study that examined carefully the long-term impact on reading performance of one potentially cost-effective classroom support – volunteer assistants.

The University of Sunderland's *Time for Reading Project* commenced a short time before the main study reported here had been designed, although the results, which included a three-year follow up, were not published until 2000 (Elliott, Arthurs & Williams, 2000). The project employed a volunteer reading programme designed to operate with children in reception classes (age 4-5 years) in areas of socio-economic disadvantage. It was hypothesised that the provision of volunteer support in the classroom would have an impact in reducing the experience of reading failure. The focus of the volunteer inputs was development of phonological awareness, letter knowledge, gaining experience of a wide range of reading-related behaviours and heightening the enjoyment of stories. The project was supported by a detailed training programme, a specially-prepared manual for volunteers and teachers and ongoing supervision, monitoring and management by the research team in collaboration with senior staff in the volunteer schools.

Volunteers were recruited in response to newspaper advertisements and were selected following interviews to ensure suitability and commitment. At the end of six hours training over a three-week period, 31 volunteers were attached to one of three participating schools, with a commitment from each volunteer to be available for two half days each week. Each experimental class had a parallel control class, and in each of the experimental classes two volunteers were present at all times of normal class contact. Control classes received no volunteer support. The key task of the volunteers was to enable the children to develop a range of skills identified in the literature as important for early readers. The content of their support was not viewed as innovative or different from what teachers were already doing, but it was considered that their presence in the classroom released more time for teachers to work with children and increased the overall support available.

At the outset the total sample was 140 (68 experimental, 72 control). Each child was assessed individually at pre-test stage in the December of their reception class year and again after six months of intervention at the end of the session in June, using literacy measures designed for the study. These included letter sounding, phoneme identification and rhyme. Follow-up assessment three years later was undertaken using a standard reading test measuring reading accuracy, reading comprehension and spelling. The aim of the study was not to assess the progress of individual readers but to compare experimental and control groups.

Despite the fact that this study was carefully planned, its outcome was disappointing. No gains were identified for the experimentals at either of the post-test assessment points. Their performance was no different from that of the controls either at the end of the period of implementation or three years later. The authors concluded on a salutary note, confirming that a considerable amount of time and effort to enhance children's reading experience had made no discernible difference, and urging caution in the employment of additional classroom support without careful planning of many factors and rigorous evaluation.

There were many respects in which this study employed good research methodology. The ideas underlying the plan to use volunteers as classroom assistants, and the activities assigned to them, had been carefully researched prior to the intervention. There was also a careful approach to ensuring that recruitment procedures would be as effective as possible, and the training provided appeared to be suitably intensive and appropriate, although short-term. Satisfactory levels of supervision and monitoring were established. In addition, there was a reasonable sample size, divided almost equally between experimentals and non-intervention controls. Both groups were similar in their assessment characteristics at pre-test, and by being in different reception classes in the same schools they were in similar educational settings. No data were provided for the reliability and validity of the assessment measures designed for the first phase of the study. However, the measures were published in an appendix to the study, and their content and construction was very similar to the items in a number of existing baseline schemes with published reliability/validity data.

The authors noted at the end of the intervention period that all concerned (school staff, volunteers, research team) 'considered the project to have proven highly successful' (Elliott et al., 2000, p. 241). Yet there is no evidence that the disappointing outcome of this study represented a type II error. Full assessment data were provided and were analysed using appropriate statistics. Nevertheless, this was a study with many important weaknesses, and it must not be considered as an adequate basis for demonstrating that classroom support using volunteers is ineffective. Most importantly, there is no sense conveyed that this was a coherent intervention that represented part of a more coordinated plan to support literacy in the schools in question. Some of the factors pointing in this direction are only noted in isolation by the authors in their own assessment of why the initiative failed. For example, it is clear that there was no effective link with home, and no involvement of parents appears at any level, despite recognition of the likely impact of home factors. There was also superficial liaison between volunteers and teachers, with 'little communication about the purposes and goals of reading activities' (page 240). Indeed, the volunteers seemed to be working largely independently of the teachers and the curriculum, and to have limited understanding of how their role might be effective. There is also considerable evidence of slippage between what the researchers saw as a coherent intervention programme and what was actually taking place in the classroom. Indeed, the indications are that the presence of two volunteers in a classroom was distracting, leading to some volunteers opting to work with individual children elsewhere.

The overall programme described for the study, while containing expected and appropriate elements, seemed to lack a focus on capitalising on strands of intervention that might be most effective. Indeed, the authors felt that the programme might be ‘overly comprehensive’. It might be better expressed as being under-targeted. Classroom volunteers should have had clearly-defined tasks that supplemented the work of the teachers in relation to the reading curriculum and that focused on the key elements of developing literacy. The ‘comprehensiveness’ of the programme seemed to be vested in the wrong place, namely, with the volunteers, who had too diffuse a range of potential activities to engage in; rather it should have been mediated through school and project management as part of an agreed strategy in which the contribution of all parties was clearly defined.

The sample wastage over the three-year period (loss of almost one-third of the original sample) was mentioned but its possible implications were not explored. It pointed to the likelihood of a complex pattern of child and family needs in this area of socio-economic disadvantage, requiring consideration of an intervention programme that might take account of a wider range of factors. Finally, there was no consideration of planning any continuity for the intervention. The researchers felt that the support might not have been for a sufficiently long period. Crucially, it was not embedded in longer-term school planning that might take account of what happened in the year after the reception class. In light of all of these considerations it is not unexpected that this initiative did not meet with success.

In the current study, extra classroom help was viewed as being crucial to the success of the project, and it was therefore the key area to which resources were allocated. It was decided to employ early intervention teachers rather than classroom assistants, and the model and scale on which this was done is outlined elsewhere in this chapter.

Strand 4: Fostering a ‘literacy environment’ in school and community

An important supporting strand was fostering a ‘literacy environment’. It was viewed as important to this study that reading should not be seen in isolation as a mechanical skill acquired through direct instruction, but rather as a meaningful, culturally-relevant activity forming part of the fabric of children’s lives. While the absolute necessity of promoting phonological awareness and providing systematic phonics instruction has been emphasised, it would be a misrepresentation of this emphasis to construe it as being the sum and substance of the approach taken to literacy throughout this project. A general but unfounded criticism of the phonics approach was that it focused on mechanical skills in isolation from the wider reading context. This was never the case, as noted by many advocates of phonics (Adams, 1990; Chall, 1967, 1983). Chall in particular has emphasised that teaching only phonics in isolation was not an option at any time, and that the wider areas of ‘literature, writing and thinking are not exclusive properties of any one approach to beginning reading’ (1989, p. 531).

It was central to this study to foster a literacy environment for children, ‘engaging them regularly and interactively in the enjoyment and exploration of print’ (Adams, 1990, p. 411). There was therefore a strong emphasis on stories, listening, making

libraries accessible both in the school and in the community and on many other areas related to a focus on reading in the children's natural environment.

Strand 5: Raising teacher awareness through focused assessment

The importance of assessment for the project extended beyond its place in evaluation of the intervention. It served also as an essential feedback mechanism for teachers on the needs of their pupils and the effects of their teaching. This was a key strand and served two related purposes: raising awareness and heightening accountability. Teachers and nursery staff not only had the opportunity to assess their children's progress with sensitive measures but also operated within a context in which good results were expected throughout the education authority. This supported a clear focus on the need for systematic teaching in those key areas measured by the baseline tests. While the dangers of encouraging teachers to 'teach to the test' are recognised, and have been outlined in the discussion of the baseline assessment scheme in Chapter 7, the advantages of a focus on fostering the skills to be tested has also been discussed. For example, if children are taught to recognise letter sounds thoroughly so that they will score 26 out of 26 when assessed, then they will have mastered one of the most basic skills for developing later competence in reading.

Lindsay (1998) has outlined seven purposes of baseline assessment, and has highlighted its potential value in supporting learning and not just evaluating outcomes. The purposes are: early identification of children with difficulties; identification of the nature of difficulties and consequent needs; monitoring progress; identification of goals and of children's steps to learning; resource planning; accountability; and budget determination.

All of these purposes were envisaged when the baseline assessment scheme was designed for the project, and they have clear implications for supporting the intervention. The value of the tests was recognised when discussions were held with head teachers on their views about the future of baseline assessment once it had served the main research purposes of the project. They showed a continuing commitment to the use of the scheme because of the way it raised teacher awareness and gave very detailed information on each pupil. Thus, the scheme has provided an essential tool in planning for effective teaching.

Strand 6: Increased time spent on key aspects of reading

As previously noted, pupils who enter primary school from low-income homes have spent very much less time on literacy-related activities than pupils from higher socio-economic groups, and the gap may be so great that it is almost inconceivable how it could then be remedied (Adams, 1990). For this reason it is clear that for such pupils the pivotal area of developing literacy skills must be made a priority in terms of the time allocated to it. However, it has been demonstrated that schools in low socio-economic areas spend not more but less time in providing reading instruction than comparison schools (Birman et al., 1987). In addition, during the time that reading is taking place, in the low socio-economic schools less material is covered and pupils

are rarely challenged to think about the meaning or structure of the text (Allington, 1989).

The amount of time actually spent on reading is crucial to the development of competent literacy skills (Clay, 1979b; Stallings, 1976; Stallings, Robbins, Presbey & Scott, 1986). For this reason the question of the time to be spent on reading in the primary schools in this study was not treated informally at the level of simply encouraging schools to increase the emphasis on this area. Instead it was treated with such importance that it was formalised as a key strand of intervention. The education authority agreed with the view that reading could not be sufficiently increased within the Scottish Office guidelines for the structure and balance of the 5-14 curriculum. In discussion with the Scottish Office agreement was therefore reached to change the balance within the authority so that reading could be given higher priority. While it is recognised that increasing the time for one curricular areas decreases it for other areas, the position adopted was that literacy is so fundamental to future access to every part of the school curriculum that it must be made the top priority for achieving success.

Strand 7: Identification of and support for children who are failing

This was another of the key strands in the project. Through the baseline assessments, and by other means in the later years of schooling, it was possible to identify every child whose progress in reading was a cause for concern, and therefore to focus additional support appropriately. It was never assumed that children who were failing would somehow catch up at a later stage, and extensive experience of assessing pupils at all stages throughout the authority over a considerable period of time had confirmed that those who fail in the early years were likely to experience accentuated failure later – an observation widely supported in the research literature (Stanovich, 1986).

The extra support provided in the classroom allowed children with difficulties to be targeted for additional help. Those who continued to fail in the early period covered by the main study were again identified in later assessments, and steps taken to remedy their difficulties. This strand of intervention was supported by the individual support study, as described in Chapter 15.

Strand 8: Lessons from research in interactive learning

The research in interactive learning provided both a general context for informing the intervention and also a range of specific supplementary strategies. The general context was a recognition that effective, motivated learning is likely to be enhanced by a social and interactive setting which is meaningful to the child and which involves both peer collaboration and adult support (Hughes & Greenhough, 1994). The specific supplementary strategies included the promotion of paired reading and computer-assisted learning. As outlined in Chapter 6, paired reading has one of the most extensive evidence bases for effectiveness of any intervention (Topping, 2001; Topping & Lindsay, 1992). It also has the advantage that it can be used anywhere and at any time, without the need for specified teaching materials, and can utilise

available adults or peers. Computer-assisted learning, particularly in a collaborative social context, has also shown benefits to various aspects of children's learning (Issroff, Jones & Scanlon, 1994; Tolmie & Howe, 1994). These strategies were promoted throughout the schools, together with advice and support to staff.

Other aspects of social and interactive learning were evident in the studies carried out in support of the main study. The synthetic phonics study (Chapter 11), while focusing specifically on direct phonics instruction, did so in a very interactive social context in which whole classes or groups learnt and physically demonstrated the actions for the different phonemes together. This social learning context was seen as one of the benefits of the intervention, and not just the synthetic approach in and of itself. In addition, the declaration study (Chapters 13 and 14) included active learning through group and class chanting and various other activities involving group participation.

Strand 9: Home support for encouraging literacy

The importance of the home in encouraging children's literacy has been emphasised throughout this study, and this was therefore identified as one of the key strands for intervention. The home is not only a major early determinant at the pre-school stage of how children will progress (Adams, 1990), but it has ongoing significance during the years of schooling (Hannon, 1995; Hewison & Tizard, 1980; Tizard, Schofield & Hewison, 1982; Topping & Wolfendale, 1995; Wolfendale & Topping, 1996). In the early work by Tizard et al. (1982) in disadvantaged areas in London, there was a clear impact of parental involvement in literacy, and test gains were sustained three years later at follow up (Hewison, 1988).

Families from socially disadvantaged areas can be successfully engaged in supporting children's literacy (Topping & Lindsay, 1991). At the same time the challenges in these circumstances of effectively engaging parents, many of whom themselves have significant levels of illiteracy, have also been recognised (MacKay & Watson, 1999; McMillan, Fox & Wood, 1994). To support family involvement in literacy in this project, four home-link teachers were appointed.

Strand 10: Changing attitudes, values and expectations

The significance of attitudes, values and expectations was a major consideration in the preparatory studies outlined in Chapter 5. It was the central thrust of the randomised control trial conducted by MacKay (1995a), and was also one of the key ingredients in the early intervention study by MacKay and Watson (1999).

There were two ways in which this dimension was incorporated within the study as a supporting strand. First, the importance of attitudes, values and expectations was highlighted in the training programmes that supported the intervention. The idea was that teachers should recognise that their pupils would make better progress if they were motivated with positive attitudes about the value of literacy to their own situation. It was essential therefore for reading to be an enjoyable, meaningful and

relevant experience. Also, the need for high teacher and pupil expectations was constantly emphasised.

Second, the key context variables underpinning the whole intervention process and driving the project through its various cycles acted as a strong support for this dimension. These were vision, profile, commitment, ownership and declaration. The entire project was predicated on a context in which everyone was fully committed to it and ownership was enjoyed at every level from researcher to the youngest pupil in nursery. Declarations were constantly made that this project was going to succeed, that nothing would stand in its way, that reading attainment would increase markedly and that even the child with the greatest difficulties would still achieve success. In a context like this where there is no scope for entertaining the idea of failure, positive attitudes are anticipated and expectations are high.

Two studies were conducted to support this strand of the main study. These were the attitudes study (Chapter 12) and the declaration study (Chapters 13 and 14).

Method

Design

This was a long-term, multiple-component intervention study, using a cross-lagged design in which pre-intervention population cohorts served as controls for subsequent intervention cohorts at the same age levels.

Sample

The sample was every child in the pre-school year in all nurseries (N = 23 establishments) and in Primary 1 and Primary 2 in all schools (N = 35 schools) throughout West Dunbartonshire. This provided an intervention sample of 19,327 from 1997 to the end of 2003, plus a pre-intervention control population of 3,659, making a total of 22,986, all tested individually. The breakdown of the sample is shown in Table 8-1.

Table 8-1 Early intervention sample: individual assessments

	Pre-school	Primary 1	Primary 2	Total
1997 Pre-intervention controls	1083	1307	1269	3659
1998	1177	1185	1260	3622
1999	1039	1160	1140	3339
2000	1021	1090	1152	3263
2001	986	1100	1097	3183
2002	893	1054	1127	3074
2003	798	1001	1047	2846
Total	6997	7897	8092	22986

To assess the effects of the intervention into the later years of primary school, the pre-intervention populations in Primaries 3, 4 and 7 were assessed each year on group tests, allowing comparison of same-age cohorts when the early intervention sample progressed through to these stages. This provided a further sample of 20,384 from 1998 to 2003, of whom 8,948 were pre-intervention controls and 11,436 had received all or part of the intervention programme in earlier years. The breakdown in Table 8-2 shows which of the sample had been exposed to one, two or three years of the early intervention programme.

Table 8-2 Group reading tests: Primaries 3, 4 and 7

	Primary 3	Primary 4	Primary 7	Total
1998	1078	1026	1022	3126
1999	*1139	1108	1187	3434
2000	**1221	*1247	1267	3735
2001	***1068	**1128	1057	3253
2002	***1130	***1128	1203	3461
2003	***1047	***1115	*1213	3375
Total	6683	6752	6949	20384

Early intervention sample: * one year ** two years *** three years

The cut-off point for reporting outcomes in this ongoing, long-term study was December 2003 for individual baseline assessments, and May 2003 for group tests. This allowed year-on-year comparisons for the early intervention. It also allowed comparisons in P3 and P4 for children who had been through the entire early intervention programme from their pre-school year onwards, together with the first comparison in 2003 for P7 children who had received early intervention for one year only when they were in Primary 2.

The total number of tests conducted for the main study from pre-intervention baseline to the cut-off dates for reporting was as follows: 22,986 individual tests and 20,384 group tests were carried out, making a grand total of 43,370 tests. Not all of this vast data set needed to be utilised in testing the hypotheses for the main study. However, as discussed and demonstrated elsewhere in the text, the total extent of the available data proved useful in relation to other aspects of the main study and also to the other studies. In relation to the main study it allowed general trends to be checked year on year, calibrating and confirming the data reported for the key years that have been selected for detailed analysis. It also served more general purposes apart from its role in data analysis, and these are discussed in Chapter 9. In relation to the other studies, the range of data available from the main study supported specific analyses, such as those reported for the synthetic phonics study in Chapter 11. The data set

also informed the discussion of the test characteristics of the baseline assessment as discussed in Chapter 7.

Measures used

Individual testing in the pre-school year, Primary 1 and Primary 2 was carried out using the baseline assessment tests designed for the study. This provided essential data on concepts of print, phonological awareness and early reading skills, as described in Chapter 7. Testing was carried out each year in November-December, with the 1997 sample serving as pre-intervention controls against which later population cohorts would be measured.

Testing took approximately 20 minutes per child, and was conducted by classroom teachers and by early intervention teachers employed for the project. All testers were provided with training. Training covered not only the correct administration and recording procedures but also the practical elements of individual assessment for staff not familiar with testing. This included consideration of establishing rapport, making the child feel confident and at ease and choosing an appropriate physical setting to ensure comfort and privacy. About 120 testers were required each year, and training was ongoing to ensure that any new staff joining the programme were trained before carrying out the assessments.

The group tests chosen for pupils in Primaries 3, 4 and 7 were the Norman France Reading Tests, and these were conducted in May each year. These tests had good levels of reliability and validity (France, 1978, 1981). Test-retest reliability coefficients were reported in the range 0.84 to 0.86, with correlations for validity in the range 0.73 to 0.89 in comparison with other established reading tests and teacher assessments. All Primary 3 and Primary 4 pupils took Level 1 (standard form) of the test, while all Primary 7 pupils took Level 2 (standard form).

The group reading tests were administered by class teachers. They took the form of a multiple choice test in which the child looked at a picture or read a sentence and inserted the correct word from a range of possibilities. They yielded a single raw score which could be converted to a reading age based on a normative sample. Although they lacked the sensitivity of the individually administered baseline assessment, and also clearly had a high loading in terms of cognitive and attentional skills, they nevertheless provided an overall measure of change in the cohorts at each of these stages.

Implementation

Implementation of an intervention on this scale over a period of years was a significant logistical exercise. It required co-ordinating the work of some 400 staff, all of whom had to be adequately trained, informed and supported to ensure high fidelity in the delivery of the programme. The key factors in the implementation process were training, resources, monitoring and support, dissemination and review.

In terms of training, an extensive, ongoing programme was planned and carried out. It provided all staff involved in the project with opportunities for developing their skills and knowledge in relation to best practice in teaching literacy, with a particular focus on the key strands of the intervention programme. The programme included series of training sessions for different key groups, such as the staff working in nurseries, or all those teaching a particular level in primary schools. These sessions included key contributions from the authority's psychological service. In addition to these opportunities, an intensive training programme was provided for the early intervention teachers employed for the project. They were a key resource not only in supporting children in developing their literacy but in supporting the teachers and other staff in the implementation of the programme. Their knowledge and experience was therefore crucial to moderating the entire intervention throughout the authority.

In terms of resources, the overwhelming investment made in the project took the form of human resources. While the regular nursery staff and the primary teachers remained the central figures for the delivery of the curriculum, the early intervention teachers referred to above were the key element in terms of special staff dedicated totally to the study. A team of 10 experienced teachers was employed for this purpose, and these were led by a head teacher for early intervention, later supported by the appointment of an assistant head teacher from the team. The early intervention team worked full-time on every aspect of the programme. They devised materials to assist implementation in all of the schools; they supported the regular staff in carrying out the baseline assessments; they worked directly with children, in liaison with their own teachers; and in general they acted as an indispensable resource in monitoring and supporting the programme.

Every nursery and school in the authority had dedicated time from a named member of the early intervention team. A regular pattern of school visits was established, and the total amount of time allocated to each establishment was a function of the number of children and the level of socio-economic disadvantage. Therefore the neediest areas had the highest level of support.

In addition to the early intervention team, four home-link teachers were appointed. Their remit was implicit in their designation, as they acted as link personnel between home and school. They carried out a wide range of functions in relation to home support and the promotion of family literacy, working in liaison with regular school staff. Again, they were allocated according to levels of need, so that a particular focus could be developed for children with the most difficult home background circumstances.

In terms of monitoring and support, the role already outlined for the early intervention team and other staff described above was part of a carefully planned chain of support structures. Full support and encouragement were provided by the council's education committee, which was kept regularly briefed by reports from the directorate and presentations from key personnel. The directorate in turn were supported by a project steering group, of which the core membership comprised a nominated member of the directorate, the head teacher for early intervention, the officer responsible for pre-school education and the author as researcher and

consultant. The steering group had additional members attending as required, such as the principal psychologist, the officer responsible for community education or other directorate members. The group was essential to ensuring the maintenance of the study through all its stages, and to solving any difficulties encountered in its operation.

In operational terms, frequent meetings with the head teacher for early intervention allowed detailed consideration of all issues for taking the study forward successfully, and also provided a forum for the generation and implementation of new ideas as the project developed. The head teacher had a high profile 'hands-on' role, not only in managing the early intervention team and ensuring the quality of programme delivery but also in liaising with the heads of pre-school establishments and the primary head teachers. Overall, these monitoring and support structures offered a very satisfactory means of operating the project effectively at all times.

In terms of dissemination, the highest importance was attached to effective communication throughout the entire system. Councillors, directorate, head teachers, school staff and the general community had to be informed of the details of the study and of progress as it developed. This was done at many levels. In addition to the reports and presentations to councillors and the directorate, the study was a regular agenda item at head teacher meetings. Dissemination was also promoted by major conferences involving not only those directly participating in the project, but also heads of secondary schools, heads of other council services, members of specialist teams and representatives of Her Majesty's Inspectorate, together with the press and key figures from the community. These events were high profile occasions. They were chaired by a member of the educational directorate, with contributions from the director and the chair of the education committee, and keynote addresses from the head of early intervention, the author and others, including members of the psychological service. At times there were speeches not only from teachers involved in the day-to-day delivery of the programme but also from pupils speaking from their own experience of how it had affected them.

Other methods of dissemination included leaflets distributed to schools and parents giving full details of progress. The authority was totally committed to the view that if a major research study was taking place, with 10 strands of intervention and encouraging results from the constant evaluation of the project, then everyone should know what these 10 strands were, why they were there and what changes were taking place as a result.

Finally, in terms of review, it was considered essential that the project should be subject to ongoing assessment and analysis to ensure that it was developing successfully through every stage of its life cycle. There had to be flexibility to respond quickly to any aspects that were not working well, with a mixture of consistency to ensure stability in delivery and planned innovation to ensure freshness and maintain impetus at all times.

From the start of the intervention the factors affecting educational change, as highlighted throughout this work, were articulated. The key context variables of

vision, profile, ownership, commitment and declaration were promoted as being of the highest importance. These concepts were constantly and deliberately mediated at all levels of the education authority – councillors, educational directorate, quality assurance personnel, head teachers, class teachers, other school staff, the early intervention team – indeed all who had an involvement with the project.

Further information on materials and procedures used in the main study is provided in Chapter 16. However, it will be clear from the nature of the study as described that its essence is difficult to encapsulate in terms of specifying particular sets of curricular materials or teaching procedures. There are several reasons for this. First, the intervention that supported the main study was essentially a vehicle for mediating and enhancing a comprehensive literacy programme. It was therefore based to a very large extent on process and structure rather than on specific techniques and materials. The content of the intervention was therefore effectively the content of the pre-school and primary school curriculum as already established in national curricular guidance and supported by existing sets of reading materials used within the education authority or within particular schools. It was the way in which this curriculum was delivered by teachers, and the way in which it was supported by specialist input that constituted the process of the intervention. Thus, central factors in the intervention were the process factors previously outlined – training, resources, monitoring and support, dissemination and review.

Second, the main study was dynamic rather than static and reflected an interactive and evolving process. It did not involve the application of a specific set of materials or procedures but a continuous development of quality in curriculum delivery, supported not only by training but by the lessons, materials and techniques developed in the subsidiary studies. Therefore, for example, as the main study developed, it incorporated into its methods and materials the resources and procedures first used in the synthetic phonics study, as well as ideas and techniques arising from the declaration study.

Third, the key context variables of vision, profile, ownership, commitment and declaration that have been described above as being ‘constantly and deliberately mediated’ at every level and in every forum of intervention represent a model in which teacher behaviour changes – again through processes of information exchange, training and support rather than through using particular curricular materials. Thus, for example, teacher and pupil expectations are raised and literacy comes to the forefront in high profile.

SUMMARY

This chapter describes the rationale and design of the main study in taking forward its key aims of raising literacy levels, providing a basis for lasting improvements and reducing numbers of children failing, with a view to addressing the widespread problem of illiteracy. It develops a rationale for a multiple-component strategy with 10 strands, carried out through the six years of the first phase of the study. The 10 strands were: phonological awareness and the alphabet; a strong and structured

phonics emphasis; extra classroom help in the early years; fostering a 'literacy environment' in school and community; raising teacher awareness through focused assessment; increased time spent on key aspects of reading; identification of and support for children who are failing; lessons from research in interactive learning; home support for encouraging literacy; and changing attitudes, values and expectations. It describes a sample of approximately 23,000 children in their pre-school year, Primary 1 and Primary 2 who received individual assessments, and a further 20,000 in Primaries 3, 4 and 7 who undertook group assessments. The intervention and its implementation are described, and the place of training, resources, monitoring and support, dissemination and review are highlighted as central requirements for its effectiveness.

Chapter 9

The Main Study: Results

Introduction: 'significance' and meaning

The main study was the centre piece of this research. All of the other studies – synthetic phonics, attitudes, declaration and individual support – were designed to support it, and to provide strategies for strengthening its implementation. It was therefore of the most crucial importance to the intervention that it should not only have significant results, but also that these results should have meaning beyond statistical significance in and of itself. It was necessary that the results should be seen as making an impact, and as providing a basis for confidence that long-term and meaningful changes could be effected and sustained, and that these would have actual importance in the lives of the population served by the project.

It was for this reason that a decision was made at the start of the intervention to report results not only in terms of statistical significance using probability values, but also in terms of effect sizes. This decision was guided by the over-arching principles governing the entire research project. It was a project based on a commitment to the scientific method of enquiry, but a science that enshrined in its methods and its priorities a commitment to values. It was a piece of research involving a large amount of public funding applied to vulnerable children and young people most of whom lived in areas of significant socio-economic disadvantage. Statistical significance in and of itself might have served the purposes of a researcher, but might have made no real and lasting impact on the lives of those who were participants in the research.

The governing principles that led to the adoption of a values framework dictated that statistical significance must be viewed in terms of wider questions that were primarily social, cultural and political rather than scientific – questions about whether lives were being changed as a result of the intervention; questions about whether children would leave school with the skills needed for a successful career in a knowledge society; questions about whether 'significant' results actually meant significant to the participants in the research or only to the researcher.

It was this main study more than any other that brought the matter back to the fundamental issue of the values framework previously outlined (Chapter 2). It was absolutely necessary that an intervention in literacy should meet orthodox scientific criteria in determining the validity of its outcomes, but for the declared purposes of this research these criteria must be compatible with promoting health, caring and compassion, self-determination and participation, human diversity and social justice (Prilleltensky & Nelson, 1997).

This issue of seeking real meaning as well as statistical significance has recently come very much to the fore in psychology. It was given centre stage as the lead article in *The Psychologist*, December 2003, in a paper entitled 'Effect size – the

missing piece in the jigsaw', with the sub-text, 'Why you just can't carry on reporting statistical significance alone' (Clark-Carter, 2003). Failure to report effect sizes has been described by the American Psychological Association (2001) as one of the 'defects in the design and reporting of research' (p. 5). This has now been incorporated in the British Psychological Society's notes for contributors to all its journals, namely, that effect sizes should be included wherever possible.

Scope of results reported

As the baseline assessments comprised 10 main tests, and these were administered individually to three separate groups, pre-school, P1 and P2, across six years, with a sample of approximately 23,000 children, supplemented by over 20,000 group tests at later primary stages, the main study generated a vast amount of data in this first phase. These included the possibility of many analyses that are not of direct relevance to this study – for example, a consideration of gender differences, or of differences that became apparent between denominational and non-denominational schools even after matching for socio-economic status (MacKay, in preparation b).

As well as providing data on many questions beyond the scope of the current work, the vast amount of data collected for the main study served additional purposes of direct relevance to the aims of the study. While the key research questions could have been answered with a smaller data set, a policy decision was taken at the outset to ensure that every child in the study was individually assessed annually on the baseline assessments, and that all of the children in the selected years in the later primary stages were assessed using the group tests.

The aims were very much wider than having sufficient data to establish whether reading achievement was rising. The requirement to carry out such extensive and regular assessment for the project was supportive of the key context variables referred to throughout the study. For example, it confirmed a very high level of commitment to the intervention from the Council, the educational directorate, the head teachers, the class teachers and others in terms of effort and resources. It also kept the intervention in high profile. It was seen as something very important involving major assessment programmes. In addition, it fostered a sense of ownership at all levels of involvement from the level of the Council to the level of the classroom.

The extensiveness of the assessment programme, especially in relation to the individual baseline assessments, also supported two of the 10 strands of the intervention – strand 5, raising teacher awareness through focused assessment, and strand 7, identification of and support for children who are failing. This therefore provided data that helped teachers to monitor the progress of individual pupils in their classes and to plan appropriate curricular strategies for them.

The questions of central interest to this study, as detailed in the hypotheses stated in Chapter 8, were: whether the year groups of children receiving the intervention programme would have higher literacy scores than the non-intervention population cohorts at the same age level; whether the gains were maintained after the conclusion of the early intervention; and whether there were significant benefits for children at

the bottom and top of the literacy spectrum – those experiencing difficulty in reading and those with high levels of reading achievement. In summary, was the intervention effective, did the effects last and did it benefit all groups?

The overview of data presented here has therefore been selected to illustrate the results in relation to these key questions. First, to answer the question of whether the intervention was effective, the baseline assessment results for each year group throughout the intervention period were compared with the results achieved by the same year group cohorts prior to intervention. The pre-intervention cohorts therefore served as the controls for the subsequent cohorts. Second, to answer the question of whether the effects lasted, the group tests given each year to children in the later primary years were analysed to show comparisons of cohort scores before and after intervention. Third, to answer the question of whether all groups benefited, separate comparisons were made for the children with the lowest and highest scores. The tests used throughout these analyses were independent two-sample *t* tests, calculated using the data analysis tools on Microsoft Excel Version 8.0.

Effectiveness of the intervention: the baseline assessment results

The first key question was whether the intervention was effective. It addressed the hypothesis that the year groups of children receiving the intervention programme would have higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level. Throughout the six-year intervention period the baseline assessment results showed a marked and consistent trend, which may be summarised as demonstrating a systematic enhancement of scores on virtually every test for every group and across every year. On the more elementary tests, such as concepts of print and nursery rhymes, especially in the older age groups, overall enhancement levels were largely dictated by the ceiling of the tests. On tests with a high ceiling – in particular the early literacy skills tests in the younger age groups and the word reading test in the older groups – the results continued to show an upward trend. The implications of floor and ceiling effects on a number of the baseline assessment tests have been recognised and are considered in the discussion section below.

Table 9-1 summarises the results for the key tests applicable to each year group. The tests of word reading and other more formal literacy skills were not applicable to the pre-school sample as far as overall comparisons are concerned, since most children were not able to score on these, although every child had the opportunity to attempt them if able to do so. Similarly, the early phonological tests were not useful for overall comparisons for the older groups, as most children by the early primary years had reached the ceiling on items such as concepts of print. The summary shows comparisons with the pre-intervention baseline of 1997 in relation to two years, namely, 1998, the first year of intervention, and 2003, the sixth year of intervention. The tests used were independent two-sample *t* tests, calculated using the data analysis tools on Microsoft Excel Version 8.0. By the end of the first year all results had shown a significant increase except the alphabet for pre-school, and letter names for Primary 1. By the end of the sixth year all results had risen significantly.

Table 9-1 Summary of results on key baseline assessment tests

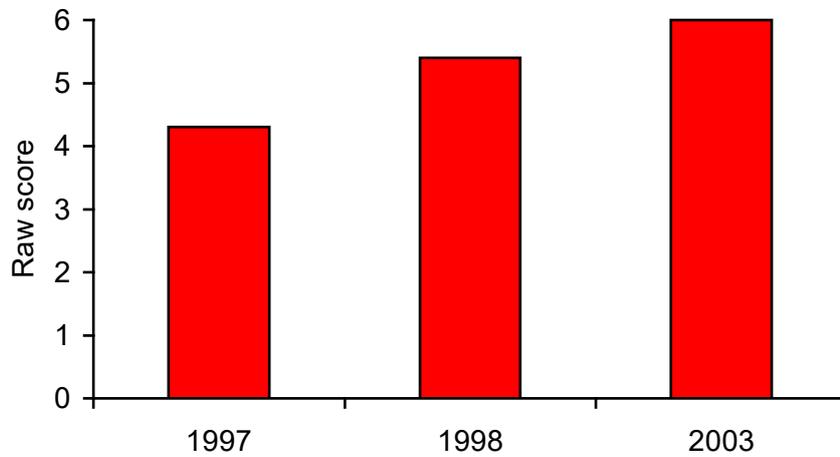
Test	Pre-school		Primary 1		Primary 2	
	1998 N=1177	2003 N=1185	1998 N=1260	2003 N=798	1998 N=1001	2003 N=1047
Concepts of print	***	***				
Nursery rhymes	***	***				
Initial letter sounds	*	***				
Rhyme detection	***	***	***	***		
Rhyme production	***	***	***	***	***	***
The alphabet	ns	***	***	***	***	***
Lower case letter sounds			***	***		
Letter names			ns	***	***	***
Non-word reading test			***	***	***	***
Word reading test			***	***	***	***

Significance (compared with 1997 pre-intervention baseline, one-tailed tests):

* p<0.05 ** p<0.01 *** p<0.001

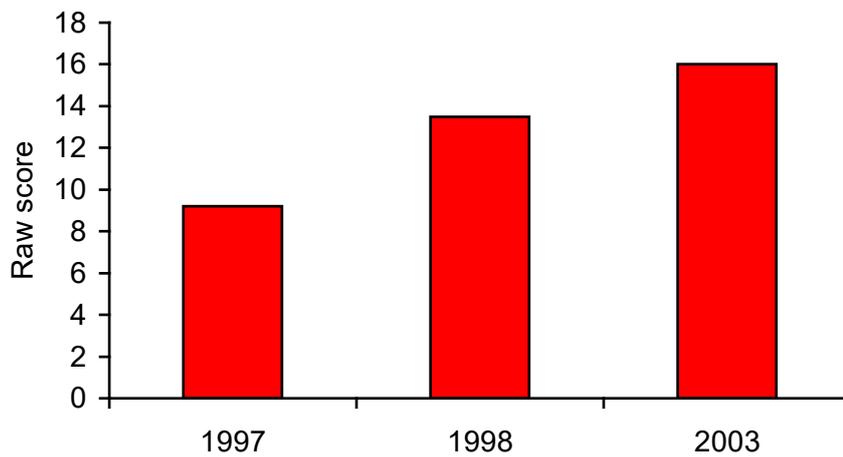
These results are illustrated in Figures 9-1 to 9-20, with effect sizes shown. The baseline assessments were normed on the pre-intervention cohort (N = 3,659). As a non-intervention population, this cohort served as the control group for the intervention samples in subsequent years. This provided the standard deviations that served as the basis for the calculation of effect sizes throughout the period of intervention. The effect sizes therefore show the raw score gains expressed as a proportion of the pre-intervention standard deviation for each test. It was considered appropriate to show the effect sizes for the two intervention years in the comparison (1998 and 2003) on exactly the same basis, that is, to make the comparison with the pre-intervention cohort. This was to answer the question, ‘How does the intervention sample in any given year compare with the data available for the same age group without intervention?’

Figure 9-1 Baseline assessments: concepts of print – pre-school (mean raw scores)



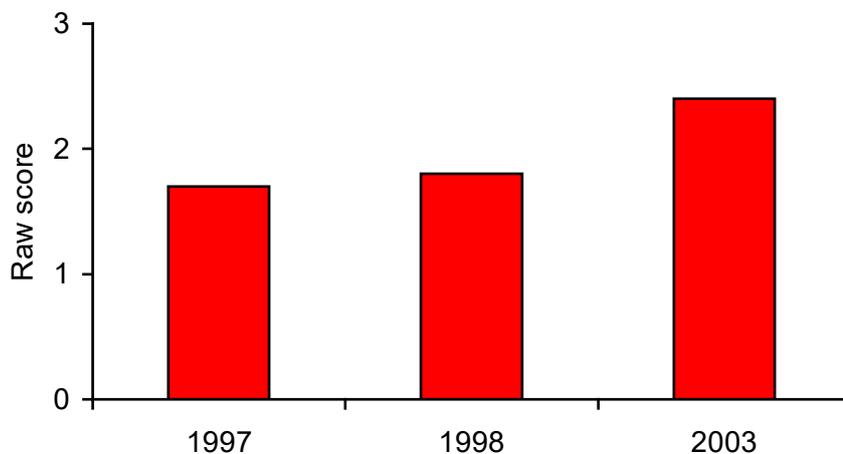
Effect size 1997-1998: 0.58 (t = 11.02***)
 1997-2003: 0.92 (t = 21.74***)

Figure 9-2 Baseline assessments: nursery rhymes – pre-school (mean raw scores)



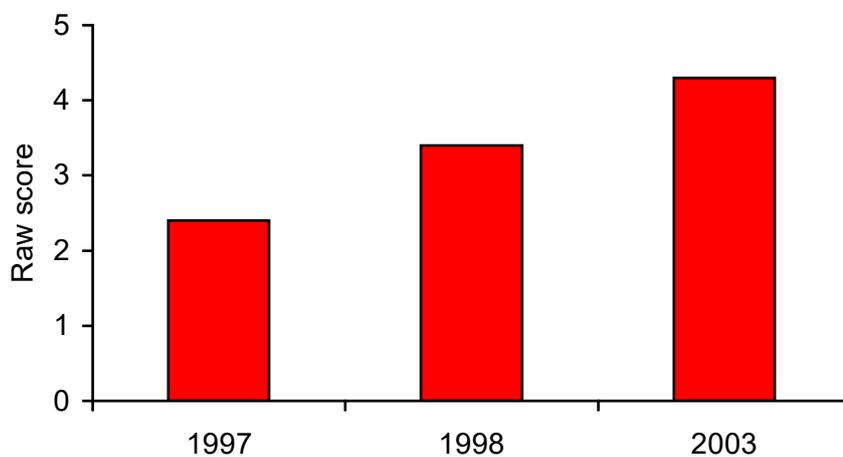
Effect size 1997-1998: 0.89 (t = 17.53***)
 1997-2003: 1.42 (t = 34.18***)

Figure 9-3 Baseline assessments: initial letter sounds – pre-school (mean raw scores)



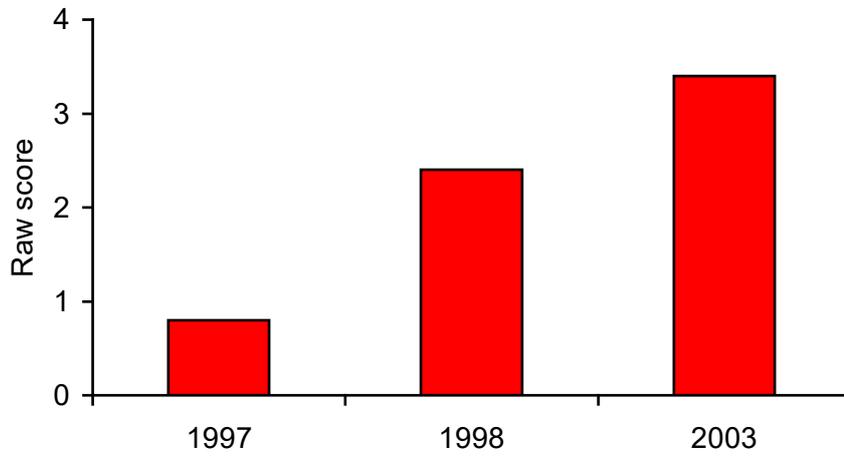
Effect size 1997-1998: 0.07 ($t = 1.74^*$)
 1997-2003: 0.41 ($t = 8.21^{***}$)

Figure 9-4 Baseline assessments: rhyme detection – pre-school (mean raw scores)



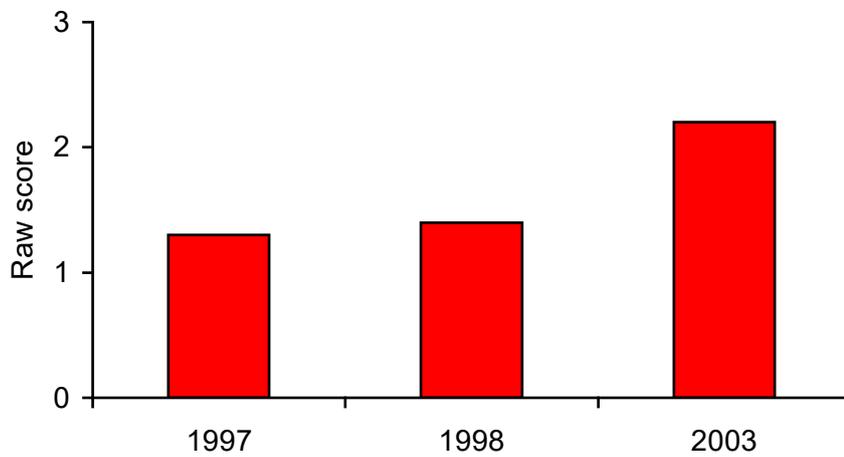
Effect size 1997-1998: 0.49 ($t = 11.38^{***}$)
 1997-2003: 0.93 ($t = 19.73^{***}$)

Figure 9-5 Baseline assessments: rhyme production – pre-school (mean raw scores)



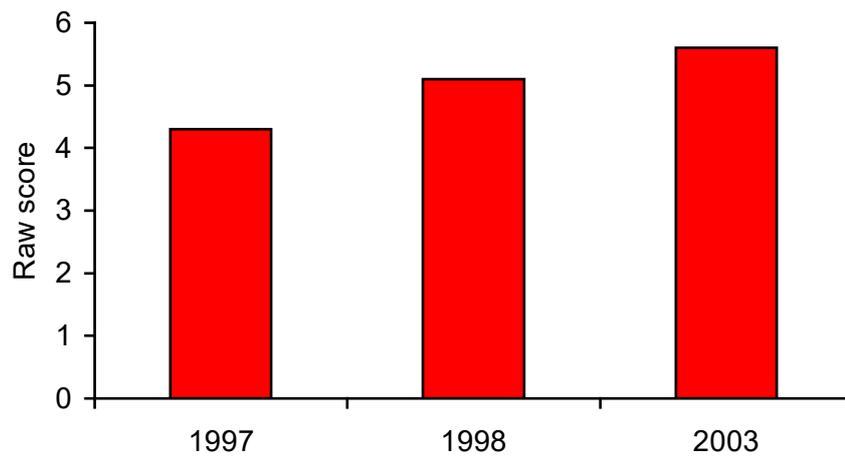
Effect size 1997-1998: 0.93 ($t = 17.11^{***}$)
 1997-2003: 1.54 ($t = 25.78^{***}$)

Figure 9-6 Baseline assessments: alphabet – pre-school (mean raw scores)



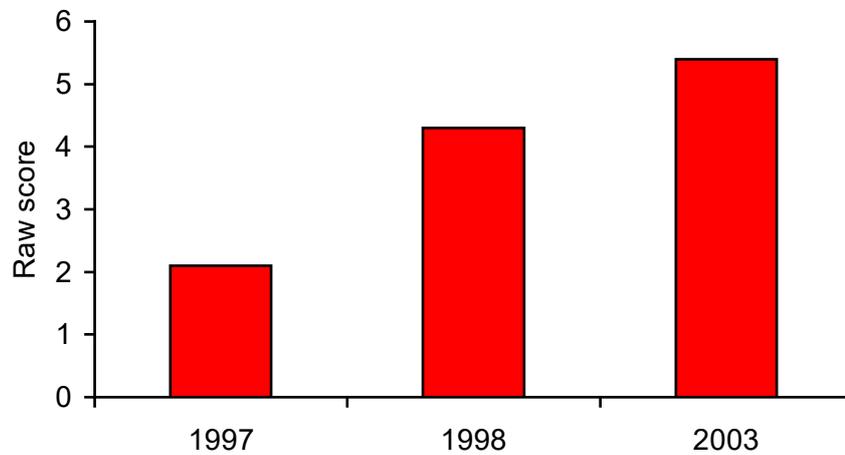
Effect size 1997-1998: 0.03 ($t = 0.58 \text{ ns}$)
 1997-2003: 1.11 ($t = 17.42^{***}$)

Figure 9-7 Baseline assessments: rhyme detection – P1 (mean raw scores)



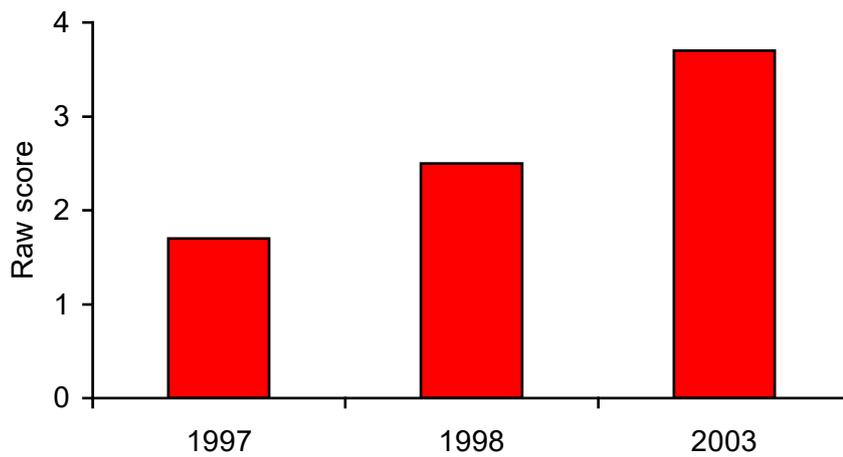
Effect size 1997-1998: 0.43 (t = 12.07***)
 1997-2003: 0.75 (t = 21.47***)

Figure 9-8 Baseline assessments: rhyme production – P1 (mean raw scores)



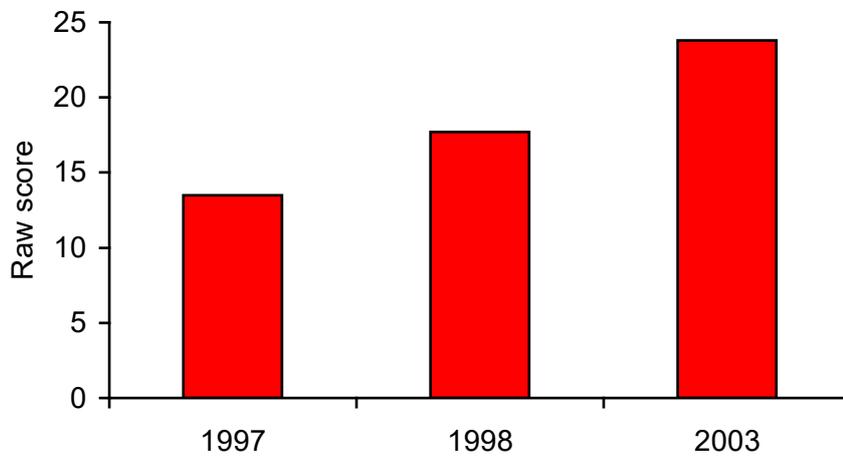
Effect size 1997-1998: 0.89 (t = 22.83***)
 1997-2003: 1.37 (t = 38.12***)

Figure 9-9 Baseline assessments: alphabet – P1 (mean raw scores)



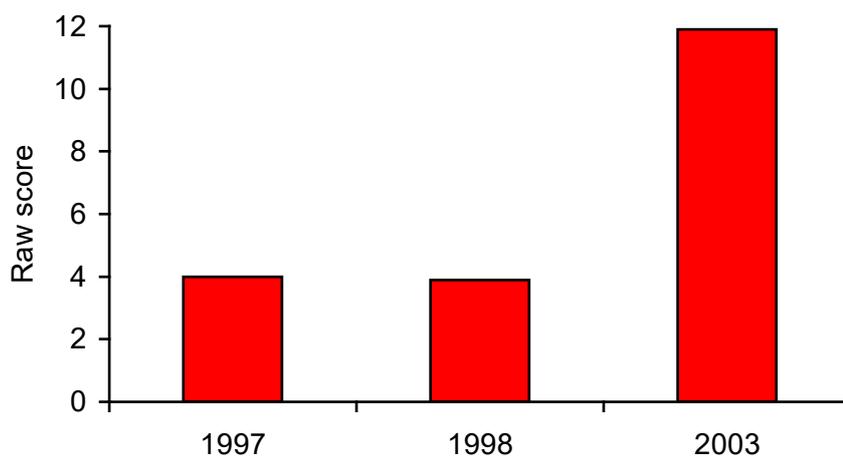
Effect size 1997-1998: 0.69 (t = 15.37***)
 1997-2003: 1.76 (t = 48.97***)

Figure 9-10 Baseline assessments: lower case letter sounds – P1



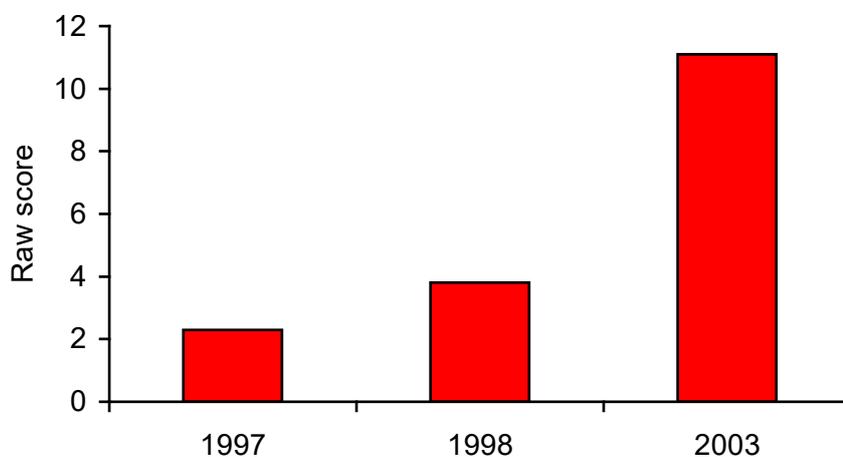
Effect size 1997-1998: 0.55 (t = 14.51***)
 1997-2003: 1.29 (t = 38.07***)

Figure 9-11 Baseline assessments: letter names – P1 (mean raw scores)



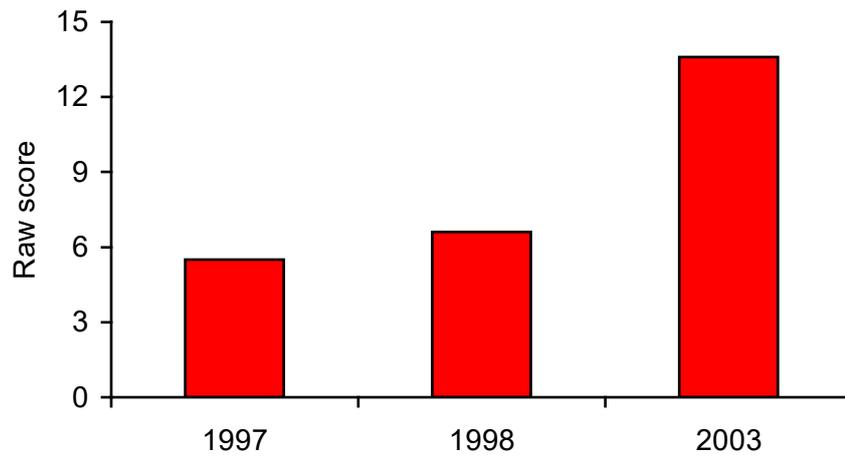
Effect size 1997-1998: -0.01 ($t = -1.65$ *ns*)
 1997-2003: 1.08 ($t = 22.90$ ***)

Figure 9-12 Baseline assessments: non-word reading test – P1 (mean raw scores)



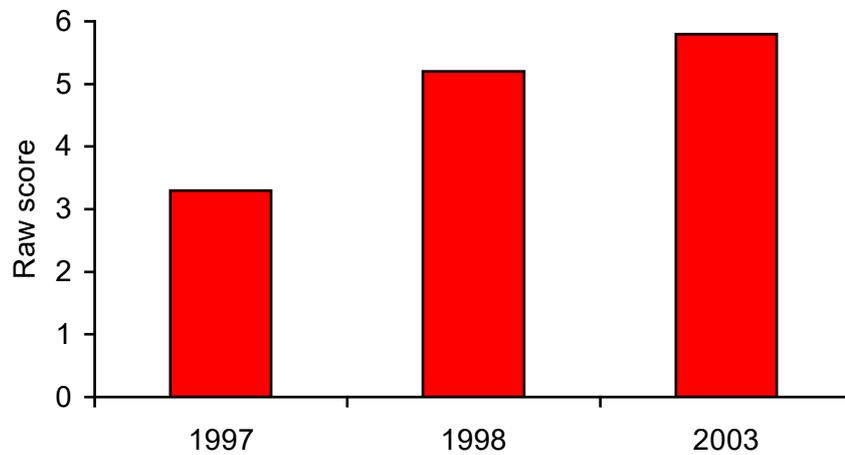
Effect size 1997-1998: 0.29 ($t = 6.55$ ***)
 1997-2003: 1.76 ($t = 35.20$ ***)

Figure 9-13 Baseline assessments: word reading test – P1 (mean raw scores)



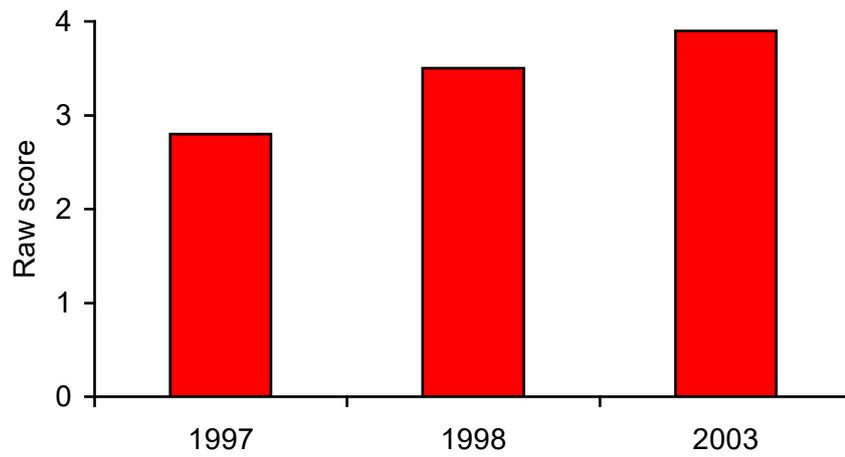
Effect size 1997-1998: 0.15 (t = 3.66***)
 1997-2003: 1.20 (t = 25.71***)

Figure 9-14 Baseline assessments: rhyme production – P2 (mean raw scores)



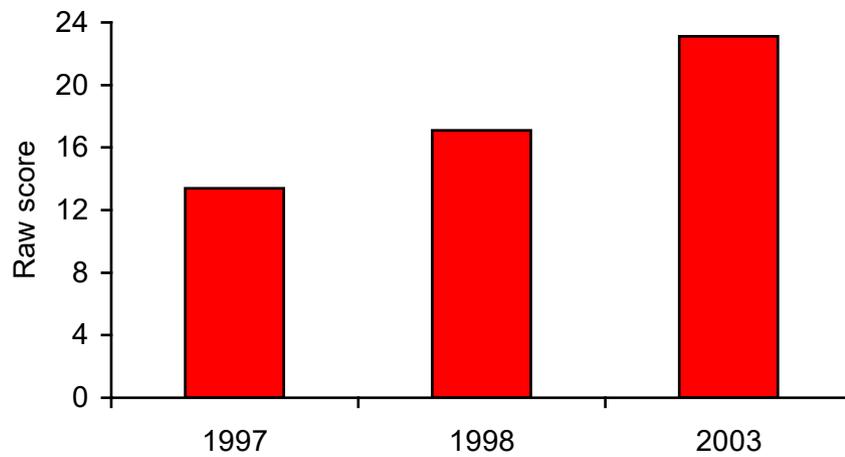
Effect size 1997-1998: 0.79 (t = 23.44***)
 1997-2003: 1.04 (t = 32.36***)

Figure 9-15 Baseline assessments: alphabet – P2 (mean raw scores)



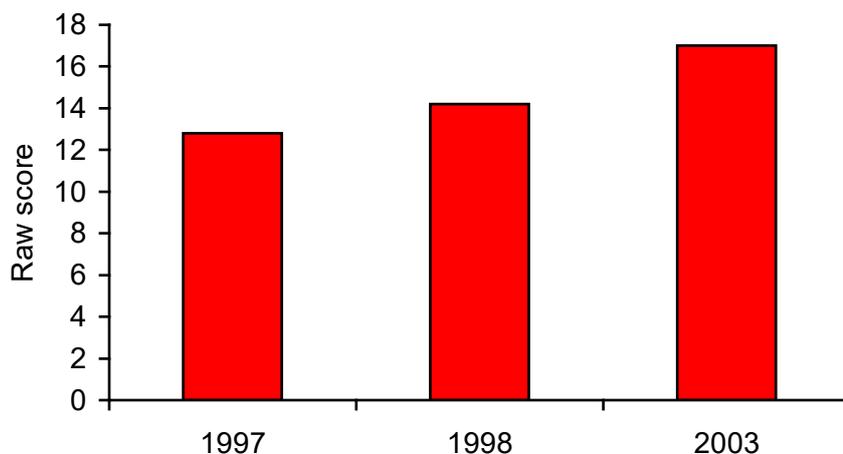
Effect size 1997-1998: 0.54 (t = 12.41***)
 1997-2003: 0.89 (t = 28.06***)

Figure 9-16 Baseline assessments: letter names – P2 (mean raw scores)



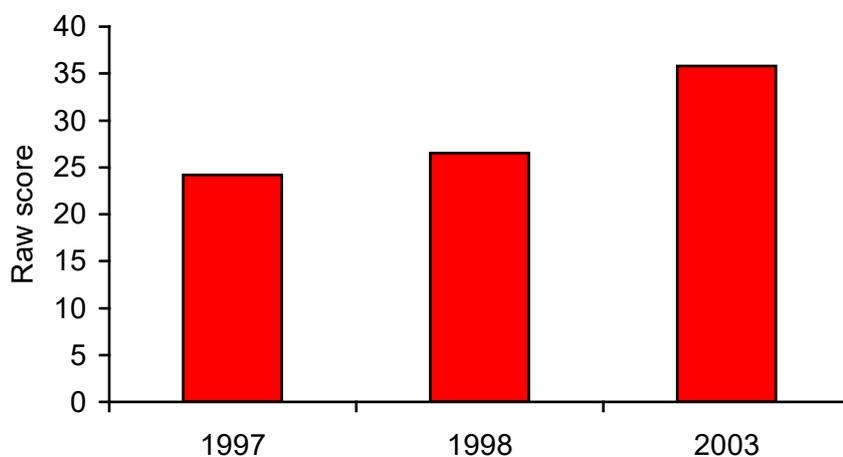
Effect size 1997-1998: 0.39 (t = 9.18***)
 1997-2003: 1.01 (t = 29.09***)

Figure 9-17 Baseline assessments: non-word reading test – P2 (mean raw scores)



Effect size 1997-1998: 0.23 (t = 5.20***)
 1997-2003: 0.65 (t = 19.38***)

Figure 9-18 Baseline assessments: word reading test – P2 (mean raw scores)



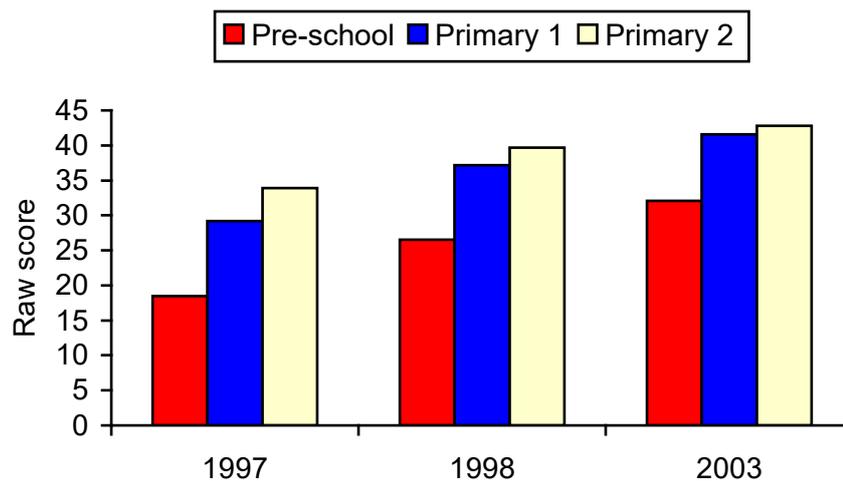
Effect size 1997-1998: 0.16 (t = 4.04***)
 1997-2003: 0.83 (t = 19.26***)

The great majority of the 36 effect sizes shown in Figures 9-1 to 9-18 indicated a moderate to strong effect of the intervention. Only five were very small/negative, and all of these were for 1998, the first year in which the intervention took place. These were initial letter sounds (0.07) and the alphabet (0.03) at pre-school, letter names

(-0.01) and word reading (0.15) at P1, and word reading (0.16) at P2. For the year 2003 the lowest effect size shown was 0.41 (initial letter sounds at pre-school). The average effect size shown across all tests was 0.45 in 1998 and 1.11 in 2003

Figures 9-19 and 9-20 provide an overall summary for each year group in relation to the three years for which results are shown. A combined score for ‘phonological awareness’ has been calculated by adding the score for concepts of print to the combined scores for the actual phonological tests – nursery rhymes, initial letter sounds, rhyme detection and rhyme production – while a combined score for ‘early reading skills’ has been obtained by adding the scores for the alphabet, lower case letter sounds, letter names, non-word reading and word reading. Phonological awareness scores are shown for all age groups, but ceiling effects become apparent in P1 and more so in P2, where many pupils, particularly following the intervention, pass all tests of this type because the skills are fully established. The ceiling for all phonological tests combined was a score of 44, and it will be noted that as the intervention progressed the mean score came ever closer to that figure in P1 and P2. Early reading skills scores are not shown for the pre-school children as these more formal skills are marked by very substantial floor effects at this age.

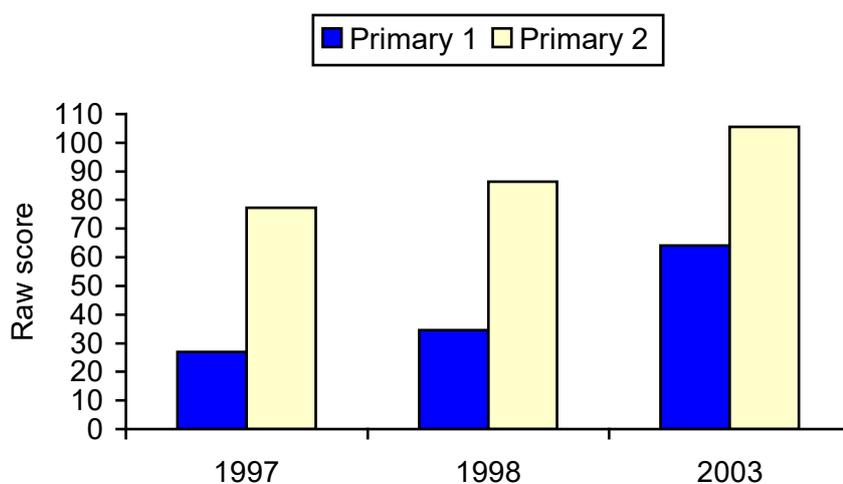
Figure 9-19 Baseline assessments: combined mean raw scores for phonological awareness



Effect sizes:

Pre-school	1997-1998: 0.92	(t = 19.43***)
	1997-2003: 1.57	(t = 33.87***)
Primary 1	1997-1998: 1.00	(t = 27.22***)
	1997-2003: 1.55	(t = 39.30***)
Primary 2	1997-1998: 0.90	(t = 22.46***)
	1997-2003: 1.37	(t = 41.23***)

Figure 9-20 Baseline assessments: combined mean raw scores for early reading skills



Effect sizes:

Primary 1	1997-1998: 0.36	(t = 8.66***)
	1997-2003: 1.78	(t = 39.30***)
Primary 2	1997-1998: 0.31	(t = 8.66***)
	1997-2003: 0.97	(t = 26.22***)

Lasting effects of intervention

The second key question was whether the intervention had lasting effects. It addressed the hypothesis that gains would be sustained after the children left the programme at the end of Primary 2. For this purpose the data collected from the start of the intervention for children in the later primary stages were utilised. The results of the Norman France group reading tests in Primaries 3, 4 and 7 were analysed. These results included pupils who at various stages had either received no intervention on the project, or had received one, two or three years of intervention. The limitations of these comparisons are considered in the discussion section below.

The 1998 Norman France tests were all conducted on pupils who had not been subject at any time to the intervention. This first year of group testing was therefore taken as a pre-intervention baseline for the cohorts at Primaries 3, 4 and 7. By the time of the tests conducted in 2000, the P3 children had had two years of intervention in P1 and P2; the P4 children had had one year in P2; and the P7 children had still not at any time participated in the programme. When the latest available tests were conducted in 2003, the P3 and P4 children had had the full intervention for three years from the pre-school year to P2, while the P7 children had had one year when they were in P2. These three years were selected to provide a snapshot of the trend of reading scores in the later primary years following the implementation of the programme.

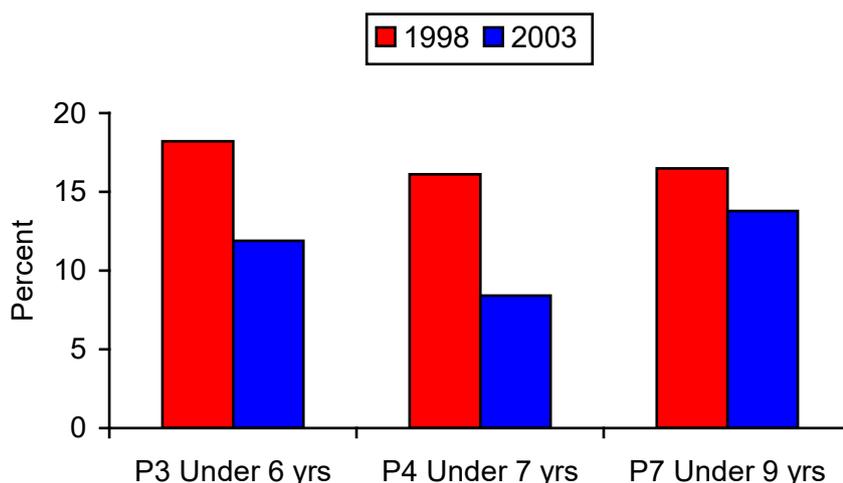
The general trend of these results may be summarised as follows. Pupils who have been on the intervention show a modest but significant increase in group reading test scores compared with the baseline cohorts at the same stages, with the effect being strongest for those who had received most intervention most recently – that is, the P3 pupils with a full three years on the programme. The results of the selected years are shown in Table 9-2.

Table 9-2 Norman France group tests compared with 1998 baseline

Mean reading ages	1998	2000	2003
Primary 3	7y 0m N=1078	7y 4m N=1221 t = 4.92*** Effect size 0.21 (2 yrs on programme)	7y 5m N=1047 t = 6.62*** Effect size 0.28 (3 yrs on programme)
Primary 4	8y 1m N=1026	8y 0m N=1247 t = 1.60 ns (1 yr on programme)	8y 3m N=1115 t = 3.82*** Effect size 0.15 (3 yrs on programme)
Primary 7	10y 6m N=1022	10y 7m N=1267 t = 1.21 ns (not on programme)	11y 0m N=1213 t = 3.82*** Effect size 0.15 (1 yr on programme)

The modest and lasting effects of the programme as discussed below may be illustrated in relation to the number of pupils experiencing reading failure at levels significantly below their age level as in Figure 9-21.

Figure 9-21 Norman France group tests: percent of scores significantly below age level

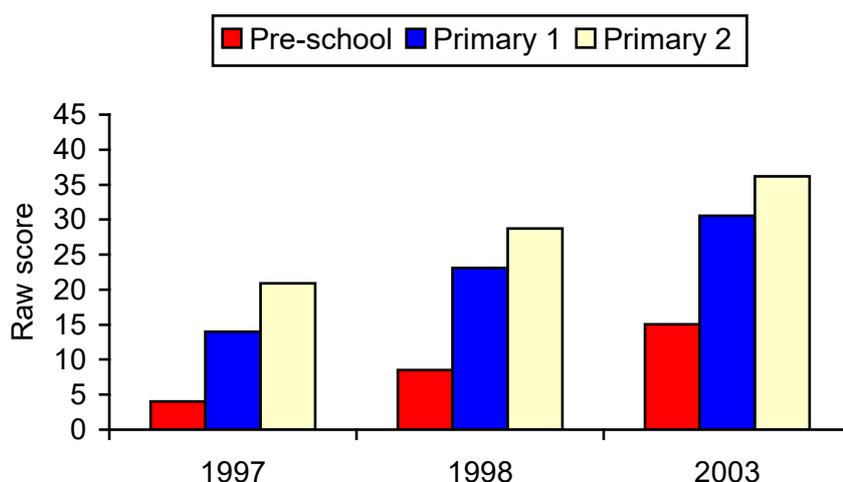


Results for children with lowest and highest scores

The third key question was whether the intervention was of benefit to all groups – to those with the lowest scores and to the high achievers. It addressed the hypothesis that both high and low achievers would show gains, but that in particular there would be significant reductions in the numbers of children experiencing reading failure in the early years. Figures 9-22 and 9-23 show the lowest 10% of scores for the three years under consideration – the 1997 baseline, the first year of intervention in 1998 and the sixth year of intervention in 2003. The data shown are the combined scores for phonological awareness and for early reading skills. Again, early reading skills are not shown for the pre-school children because of the floor effects at this age when most children are not yet learning the more formal skills.

As there was a particular interest in reducing the numbers of children who were failing to acquire basic literacy skills, some illustrative data are provided in Figures 9-24 to 9-26 on the impact of the intervention on some of the key skills it was set up to address at the various stages – concepts of print for pre-school children, letter sounds in P1 and word reading in P2.

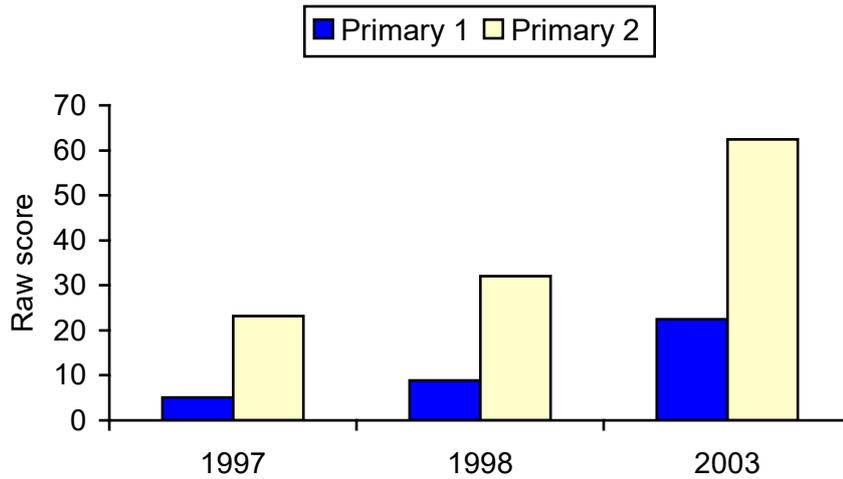
Figure 9-22 Lowest scores: raw score at 10th percentile – phonological awareness



Effect sizes:

Pre-school	1997-1998: 2.40	(t = 12.42***)
	1997-2003: 5.87	(t = 21.90***)
Primary 1	1997-1998: 2.19	(t = 15.79***)
	1997-2003: 3.98	(t = 27.56***)
Primary 2	1997-1998: 2.21	(t = 15.40***)
	1997-2003: 4.33	(t = 31.25***)

Figure 9-23 Lowest scores: raw score at 10th percentile – early reading skills



Effect sizes:

Primary 1	1997-1998: 2.07	(t = 11.05***)
	1997-2003: 9.29	(t = 28.25***)
Primary 2	1997-1998: 1.10	(t = 8.34***)
	1997-2003: 4.82	(t = 25.78***)

Figure 9-24 shows the percentage of the pre-school year passing all or all but one of the items on concepts of print. Figure 9-25 shows the percentage at P1 with a score of 20 or more out of 26 on letter sounds. Figure 9-26 shows the percentage at P2 correctly reading at least 30 words.

Figure 9-24 Pre-school concepts of print: percent of pupils with high scores

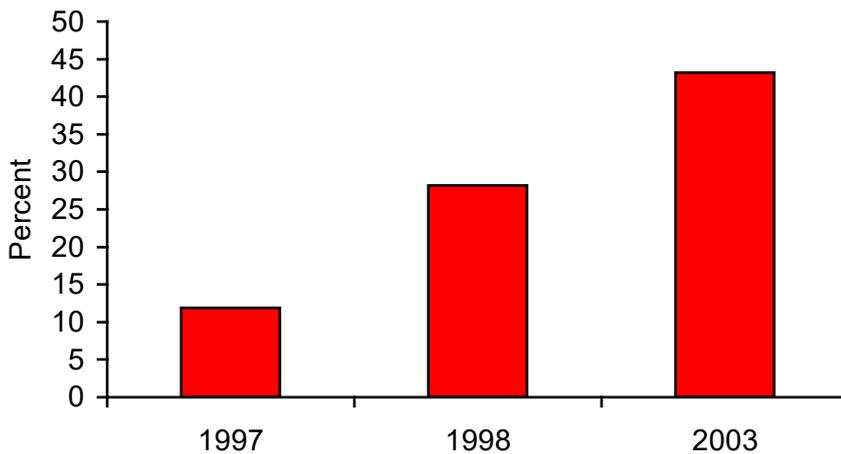


Figure 9-25 P1 letter sounds: percent of pupils scoring 20+

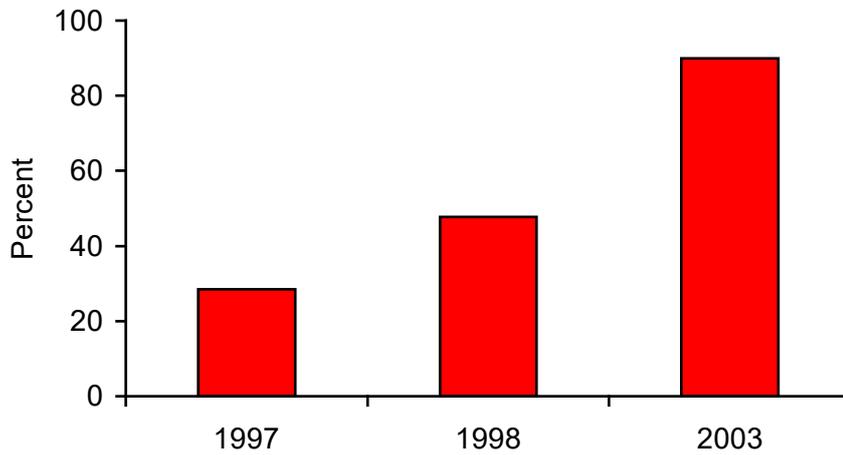
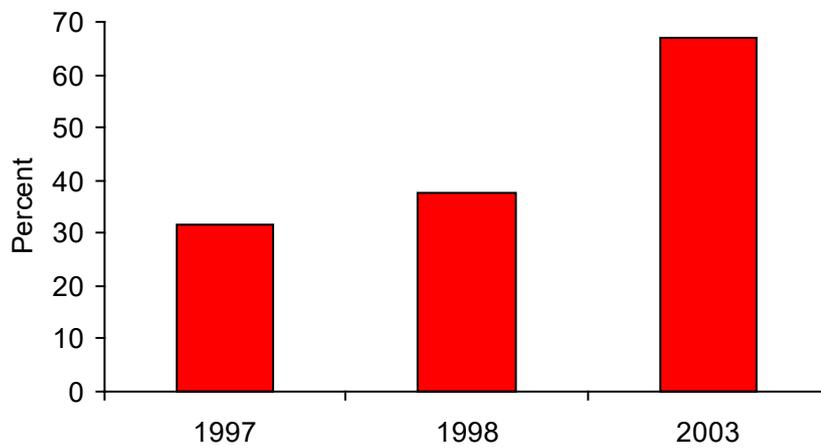


Figure 9-26 P2 word reading: percent of pupils reading 30+ words



Figures 9-27 and 9-28 show the results for the highest scoring 10% of pupils. Again, early reading skills are not shown at the pre-school stage because of the high floors for children not being taught formal skills of this type. Also, phonological awareness scores are not shown for the highest scoring children at P1 and P2 stage, as most of these children were at or near ceiling level for this skill. Even at the pre-school stage

the highest 10% of scores for phonological awareness were close to ceiling levels after the first year of intervention.

Figure 9-27 Highest scores: raw score at 90th percentile – phonological awareness

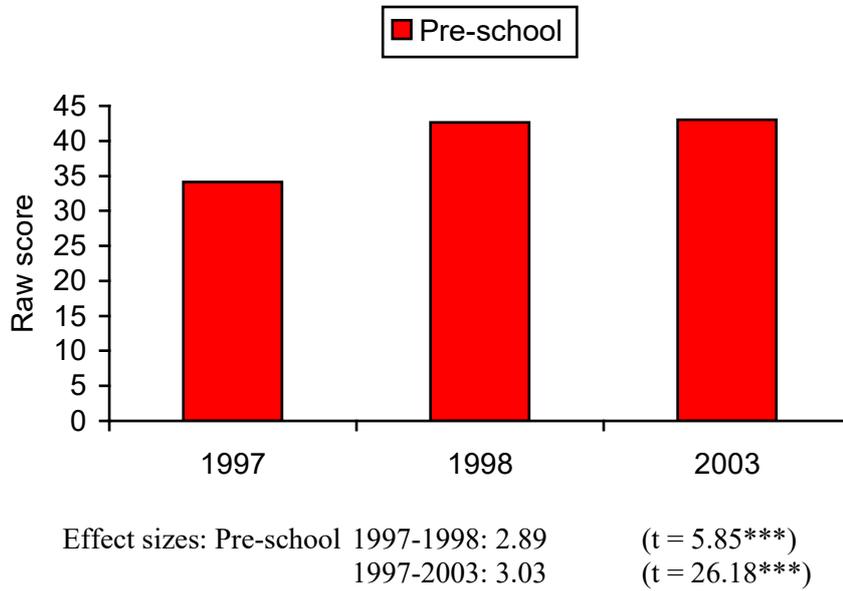
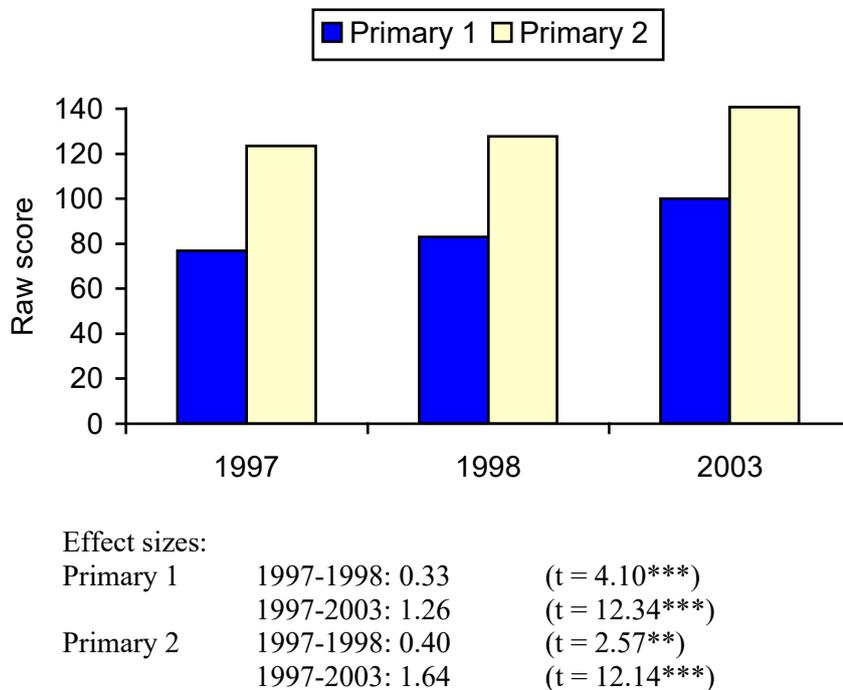


Figure 9-28 Highest scores: raw score at 90th percentile – early reading skills



Discussion

The results of this study indicate that it has achieved its aims. These were to raise the literacy levels of all children in the pre-school year, Primary 1 and Primary 2 in all schools throughout the authority using a multiple-component intervention; to provide a basis for long-term improvements in literacy levels in the later years of schooling; and to reduce the numbers of children experiencing reading failure as a basis for eradicating illiteracy throughout the authority.

Support has been found for each of the three hypotheses. First, the year groups of children receiving the intervention programme had significantly higher scores on all baseline assessment measures than the non-intervention population cohorts at the same age level. Second, there are indications that gains have continued at a reduced level after the children left the programme at the end of Primary 2. Third, both high and low achievers showed gains, and in particular there were significant reductions in the numbers of children experiencing reading failure in the early years. Evidence of reduced numbers of failing readers was also clear in the relevant later years of primary following the intervention.

Limitations and cautionary comments

The limitations in the comparisons of test scores for this study, both statistically and in terms of broader aspects of interpretation of results, have been acknowledged and are discussed here in relation to two areas: first, the comparisons made on the baseline assessment tests, and second, the comparisons made on the Norman France tests administered in the later years of primary school.

First, in terms of statistical limitations it is recognised that it has not been possible with this type of cross-lagged design, as it would have been with some other designs, to use regression analysis to adjust for regression towards the mean when there are significant floor effects. The entire baseline cohort for each test formed the controls for the following cohorts of children at the same stage, and where the floor effects were marked they could do little other than go up in subsequent years if they were going to change at all.

However, this limitation may be viewed in the wider context of the clear and significant trend of the results overall. In general the effects resulting from this study were not subtle. They were large enough that the purpose of the statistics was not so much to determine whether an intervention effect could be identified but rather to confirm its extent. The results were *obvious*. They were clearly discernible in the simplest of descriptive tables or charts used to communicate what was happening. Unmistakable increases in scores were plain across virtually every single baseline test at every age and for every year, including all the in-between years not separately shown for the purposes of this analysis.

Further comment may be made about the statistical limitations acknowledged in the analysis. The tests were not only at times positively skewed because of floor effects,

making it easier to achieve increases, but also at times negatively skewed because of ceiling effects, making it harder to achieve increases. This is clearly illustrated in the phonological awareness tests for the children at P2 stage, many of whom reached the ceiling of the tests with ease because there were no phonological skills left to master. This therefore worked against showing an intervention effect. Nevertheless, it was felt useful to illustrate these results, to conduct the *t* tests and to show the effect sizes, but to acknowledge skewness for the reasons indicated. Large effect sizes have been found throughout, not only where there are floor effects but also where there are ceiling effects, and on many other tests which were not skewed in either direction.

There is therefore an overall robustness in the consistency of the wide range of results shown. Indeed, it is the robustness inherent in the size and consistency of the gains achieved across all tests and age groups that overtakes the general range of statistical limitations associated with different methods of calculating the basis for effect sizes (Glass, McGaw & Smith, 1981) and gives confidence in the outcomes of the intervention. In keeping with the hypotheses, the comparisons of scores between baseline and the selected intervention years were one-tailed. However, in almost every case, the gains (whether calculated as one-tailed or two-tailed) were far in excess of any level at which their actual significance values would be reported.

Second, there are limitations inherent in the comparisons made using the Norman France group reading tests at Primaries 3, 4 and 7. Group tests are subject to various factors other than the ability being tested that may affect scores. They involve multiple choice approaches where several abilities other than pure reading ability are required, such as good comprehension of verbal instructions provided at group level, independent working and sustained attention. They are also more prone to children copying from someone else close to them. In addition, they have a greater focus on reading comprehension, and indeed on general comprehension. While therefore they are useful in providing a measure of real reading, they are not ideal for sensitively tapping into the abilities most fostered by the early intervention programme.

Broader aspects of interpretation of results also affected the Norman France comparisons. The further children were removed from the intervention itself – for example by P4 and especially by P7 – it was difficult to know just what range of factors might be likely to affect cohort scores either positively or negatively. Two of these factors, working in conflicting directions, are noted here.

First, as children progressed beyond Primary 2 the support they received for developing their literacy diminished. Although the overall literacy initiative incorporated a vision to tackle literacy ‘from the cradle to the grave’, the actual level of intervention was much more intense in the age groups up to P2 covered in the main study than in the stages from P3 onwards. Children therefore finished the early intervention and at times went back into settings that did not support their enhanced literacy in the same way. This issue was still being addressed in terms of new initiatives for P3 onwards at the conclusion of the period reported here.

Second, and working in the opposite direction, there were factors likely to affect scores positively over the years in the later primary stages. One of these was the

overall impact of this research. It was designed to touch every part of society throughout an entire Council area. The result, consistent with the key context variables frequently outlined in this work, was that the profile of reading went up, not just where the interventions were formal and intense but also elsewhere. The intensive focus on reading in the early stages, together with the central involvement of all the head teachers, meant that there was a greater ‘buzz’ everywhere about everything to do with reading. This not only influenced class teachers in the later primary stages – who had to adjust curricula to the higher levels of children coming through the intervention – but it also influenced families. Young children excited by the reading programme, not just in the main study but in all the interesting things happening with, for example, the synthetic phonics study, had siblings and parents. Although work in supporting parents is not separately reported on for the purposes of this study, many of the parents commented on the children’s enthusiasm and high achievement, and many examples were cited of older siblings having a heightened interest in reading. Teachers trained on the early intervention programme also at times moved to upper classes, taking with them new strategies for enhancing reading.

Comparisons with P7 are particularly difficult to interpret, and caution would be needed in directly attributing the modest increase in P7 reading ages in 2003 to the intervention. These children had experienced only one year on the programme and it was five years previously when they were in P2. The effect might well have been small and might have largely washed out by P7 without supporting structures in place at that stage to maintain it. Also relevant is the fact that a highly structured initiative directed at reading, spelling and maths was in operation for this year group, although it was taking place in only one cluster of schools in the authority. This was a study using the computer-based programme *SuccessMaker*. This was separately evaluated by the author (MacKay, 2002b), and showed a positive effect on reading performance. It applied to the 2003 P7 pupils in the schools in question when they were in their P5 and P6 years, and continued in P7 following the evaluation.

A final observation is relevant under the general heading of limitations and cautionary comments. Most of the baseline assessments were carried out by the class teachers, and they clearly had an interest in seeing their pupils performing well. Might they therefore in some way have ‘massaged’ the results? This is one of the ‘pitfalls in human research’ (Barber, 1976). Some discussion of this topic has already been raised in Chapter 7 regarding the features of the baseline assessment, and the question of when it is either appropriate or inappropriate to ‘teach to the test’. The overall consistency in the trend of results across the years and across so many schools and nurseries, with hundreds of people involved in testing, militates against this view, as does the consistent qualitative testimony to enhanced literacy from diverse sources, as observed below.

However, it is also answered by the very close monitoring procedures established for the research. Often the assessments were undertaken by members of the early intervention team instead of the class teachers, or else they were split between them. This took place not just for the initial purposes of establishing the reliability of the baseline, but throughout the intervention. Most particularly, if any results looked as if they needed closer monitoring, steps were taken to address this. Schools with scores

that seemed inconsistently high or low compared with the trend of results or with their own known performance were visited for clarification, usually by the head of early intervention and a quality improvement officer. The early intervention team knew very precisely throughout the year which schools were performing at what levels, and were a valuable source in clarifying results. Only two schools were identified where inappropriate use of baseline test materials was suspected. To address this an alternative form of two key tests was designed – the non-word reading test and the word reading test. This was issued without prior notice to these schools and to several other schools for comparison at the time of the baseline, and one establishment was advised on its practice as a result.

Raising achievement and eradicating literacy

Against the background of the limitations and cautionary comments that have been discussed, this study in achieving its aims has significantly advanced the overall vision underlying the research, that of addressing endemic social and educational problems by raising achievement and seeking to eradicate illiteracy in socially disadvantaged populations. The question of whether change might have taken place anyway, even without the intervention, in this type of cross-lagged design has been considered in the conclusions (Chapter 16). However, the size of the gains in the key areas of literacy addressed by the study, and the absence of a basis for these changes occurring for other reasons, points to the success of the programme.

In essence, the changes were marked enough and specific enough that everybody knew that patterns of achievement had been definitively changed. It hardly required even the baseline assessments, far less the statistical analysis, to inform every part of the education system that reading standards were changing. The main study did not gather systematic qualitative commentary from schools on the effects of the intervention – but it was certainly proffered. Head teachers, class teachers, the various types of support staff, parents and in many cases the children themselves knew that new levels of success had been established.

These levels of success are probably demonstrated most clearly in simple terms that hardly need the support of inferential tests. For example, in many cases children were scoring more than a year higher than their pre-intervention cohorts – that is, pupils in the pre-school year were scoring above the P1 controls, and P1 were scoring above the P2 controls. Instances of this were found in tests such as non-word reading and word reading, and for the lowest scoring children on phonological awareness and early reading skills. From the first to the last of the years reported here, the changes were very meaningful in practical terms at classroom level. This is seen, for example, in the rise from 12% to 43% in pre-school pupils attaining a perfect or almost perfect score for concepts of print; in the rise from 28% to 90% in P1 pupils able to recognise 20 or more letter sounds; and in the rise from 31% to 67% in P2 pupils achieving a score of 30 or more on the word reading test. These results were particularly encouraging as the baselines were conducted not at the end of the session for these three stages but half-way through the session.

One of the features of the results over the years was that they continued to rise year on year. Certainly a large rise was anticipated and was obtained in 1998 after the first year of intervention. However, these gains continued to increase each year till the last reported results in 2003, even if the annual rate of increase was not as large as time progressed. The three selected years reported here – baseline in 1997, the first year of intervention in 1998 and the last available year in 2003 – were individual points in a steadily rising graph. Since a new group of pre-intervention children arrived at the pre-school stage every year to begin the programme, it meant that constantly higher outcomes were being attained from a similar starting point each time. Not only therefore were the children doing better, they were doing ‘more better’ each year. This reflects the build up of the programme over time, and the installation or ‘institutionalisation’ of the longer-term processes of educational change as discussed elsewhere in this work (Chapter 2).

Although the eradication of illiteracy throughout the school-age population is a further target with an anticipated timescale outwith the framework of this study, the results reported here have achieved everything hoped for by way of preparation. The numbers of children experiencing reading failure have been systematically reduced. Pupils who were scoring at or near the tail end on the baseline tests in 2003 would in many cases have been viewed as average pupils scoring around the midpoint in 1997. In some instances low scores, once highly prevalent, have almost disappeared. For example, approximately 40% of P1 pupils could recognise only 10 or fewer letter sounds in 1997. In 2003 this had reduced to just over 1.7%. In P2 10% of pupils scored five or less on the word reading test in 1997. By 2003 this was down to 1% – a mere 12 pupils across the 35 schools. This reduction in the numbers failing as they enter the later primary years creates greater scope and economic feasibility for supporting these pupils with intensive individual help to overcome their difficulties.

SUMMARY

This chapter provides a detailed breakdown of the main results of the study. It does so by focusing on baseline assessment comparisons between pre-test cohort controls and subsequent population cohorts at the same school stage following intervention. This is done by providing data for three key years: 1997 (pre-test scores), 1998 (scores after first year of intervention) and 2003 (scores after six years of intervention). The results support three key questions addressed by the main study: first, the intervention was effective, with large improvements in performance in baseline test scores; second, there is evidence that the effects of the intervention, while reducing in size, were making a lasting impact after children left the main programme; third, all groups, both high achievers and low achievers, benefited by the intervention. In particular, the number of children obtaining low scores in key literacy attainments was greatly reduced.

Chapter 10

Developing a Phonics Intervention: Analytic or Synthetic Phonics?

Methods of phonics teaching

Traditionally, phonics teaching in Scotland is analytic. This method starts at the whole word level. Children begin with the whole word and break it down into letter sounds. Johnston and Watson's (1997) study of the method used in this traditional approach indicates that children are typically taught at the rate of one letter sound per week, and are shown a series of alliterative pictures and words starting with that sound, such as *mat*, *man*, *make*, *Mary* and *matter*. Once all 26 initial sounds have been taught in this way, children are introduced to middle sounds – *mat*, *bat*, *ran* – and final sounds – *mat*, *pot*, *sit*. Then initial consonant blends are taught, followed by final consonant blends and then vowel and consonant digraphs, such as *oo*, *ee*, *sh*, *ch*. Finally, rules such as 'silent *e*' in *cake* are taught. In parallel with phonics teaching of this type, and often prior to it, children are introduced to elementary reading books composed of pictures and simple words used repetitively, which they learn by a 'look and say' approach, reinforcing these by frequency of exposure to the same material.

In the synthetic phonics approach, children are taught groups of letters very rapidly, and they learn to combine these in various ways to make simple words. This is done before they are introduced to books or reading. All of the phonemes, approximately 44 in number counting both single letters and digraphs, are taught in this way. Despite this key difference from the traditional approach in methodology, synthetic phonics shows many areas of overlap with analytic phonics. For example, it involves the teaching of rules, and irregular words tend to be taught as whole units.

The ideas underlying synthetic phonics are not new, and the method has been used over many years in Germany and Austria (Feitelson, 1988). It was the work of Lloyd at Woods Loke School in Lowestoft, and in particular the publication of *The Phonics Handbook* (Lloyd, 1992) that marked the beginnings of the approach in the UK. In Lloyd's scheme a systematic course of synthetic phonics teaching began very soon after school entry and before the children were given reading scheme books. The pace was fast, with six letters of the alphabet being taught each week throughout a period of eight weeks. Thus in the first week the letters *s*, *a*, *t*, *p*, *i*, *n* were taught, these being selected as they could combine to make a very large number of simple words already established in young children's vocabulary, such as *pin*, *sat*, *pat*, *tap* and *nip*. These letters were shown in all positions of words – for example, *p* occurs at the beginning of *pin*, in the middle of *taps* and at the end of *nip*. Books were provided showing pictures of words containing the target letter, with the words presented elsewhere on the page.

The whole class was taught the letters at the same time on a daily basis, using distinctive actions and sounds to aid memorisation. A strong emphasis was placed on blending letter sounds together to make words. Irregular words, however, were also

taught on flash cards using the ‘look and say’ method, making it clear that not all words can be read using a simple phonics approach.

Lloyd’s approach was developed by its publishers into the *Jolly Phonics* scheme, comprising a complete course of materials including *The Phonics Handbook*, word books, flash cards, workbooks, wall friezes and ‘finger phonics books’ with cut-out letters shapes, illustrations of the actions and pictures of objects with the sound in the word. (The name *Jolly Phonics*, while having a suitable resonance for a reading scheme designed with the enjoyment of young children in mind, derives from *Jolly Learning*, the publishing company of its founder, Chris Jolly.) The scheme has been so successful as the vehicle by which synthetic phonics is taught in the UK that the term *Jolly Phonics* has come to be used widely as a synonym for the synthetic approach.

From this account it is apparent that several key elements mark the synthetic phonics approach as used in schools, and that as a method it both overlaps with and is distinct from traditional phonics, other than being based on synthetic rather than analytic methods of blending. Five key features are identified here. First, it is a highly systematic scheme of teaching phonics by direct instruction. It follows a precise methodology, and in practice every teacher in every Primary 1 classroom should be at the same stage at the same time. Second, it is clearly based on high expectations of children in terms of achieving success. The pace is fast, and it is expected that at a very early stage children will be motivated by their success in forming large numbers of words from the letters they have learnt. Third, it includes prominent features of social and interactive learning. Whole class teaching is consistent with a cohesive social context for learning to read, with active participation by all children together. Fourth, it is a multi-sensory approach. It explicitly and constantly involves at least three sensory modalities – visual, auditory, and kinaesthetic. The children see the letters, they say them aloud and hear them, and they practise the actions that accompany them. Fifth, multiple methods of teaching are used. This last observation is important as phonics methods, both analytic and synthetic, are criticised at times as being narrow, reductionist and mechanical (for example, Goodman, 1986). The synthetic phonics method explicitly requires the teaching of a sight vocabulary for irregular words, and it also supports other areas of the reading curriculum such as fostering listening skills, story-telling, comprehension and creative writing.

Early evaluation of synthetic phonics: the Scottish context

Early evaluations of synthetic phonics compared with analytic phonics are of considerable relevance to the current study as they were conducted within a Scottish educational context and stem from the work of Johnston and Watson in Clackmannanshire. Between 1992 and 1998 they conducted a research programme in three phases. The first phase explored methods of teaching reading and spelling in primary schools in the authority (Johnston, Connelly & Watson, 1995), while the second and third phases were designed as intervention studies in which phonics teaching was provided both within and outwith the classroom (Johnston & Watson, 1997). These studies are of particular importance not only in the context of the synthetic phonics study reported here, but also in relation to the overall project. They

had a significant political impact through being widely reported in the educational press and through being promoted by the Scottish Office Education and Industry Department as an *Interchange* publication (Watson & Johnston, 1998). They are therefore considered here in detail.

The first study examined how phonics was taught in a sample of 12 primary schools (Johnston et al., 1995). Its purpose was to investigate which aspects of phonics teaching were most effective in producing independent readers. At the start of Primary 1, children were encouraged to read whole words with associated pictures and captions, building up a basic sight vocabulary of key words. The teaching of phonics, while systematic, was designed to follow a gradual analytic approach during the first three years of school. It was introduced half-way through the first term, with the teaching of the 26 initial letter sounds being completed by the start of the third term. In one class, however, there was an accelerated phonics programme that was introduced at school entry. The approach included both analytic and synthetic phonics elements. The conclusions were that the accelerated programme produced better results and that these were maintained at two-year follow up.

The results of this study supported good phonics teaching whether analytic or synthetic, demonstrating, for example, that teaching explicit sounding and blending resulted in gains in reading and spelling. However, the effectiveness of the introduction of a synthetic approach was of particular relevance to the later research findings, and in relation to this it was concluded that it reflected 'the distinctive classroom programme' (Watson & Johnston, 1998). The class teacher described the enjoyment the children experienced through the multi-sensory approach, and noted that 'at the end of the first year, all of the pupils had retained more sounds than in any previous year' (The Jolly Phonics Case Studies, undated). Average reading and spelling ages for the class at the end of Primary 1 were 6 years 8 months and 6 years 11 months respectively, at average chronological age 5 years 9 months.

In the second study all of the children continued with their normal classroom reading activities but an additional training programme was provided outwith the classroom. Three groups received two 15-minute sessions each week. The control group had no extra phonics training but learnt new words using the 'look and say' method. The two experimental groups received accelerated teaching of letter sounds at the rate of two a week, one using analytic and the other using synthetic methods. All three groups were exposed to the same print vocabulary. The results indicated that synthetic phonics teaching led to better reading, spelling and phonemic awareness than analytic phonics teaching and that this superiority was not a result of increased pace (Johnston & Watson, 1997).

The third study compared the effects of phonological awareness training versus synthetic phonics teaching on reading, spelling and phonemic awareness within Primary 1 classes (Johnson & Watson, 1997). Three groups of children received an intervention programme for 20 minutes a day: an analytic phonics control group (four classes), a phonological awareness plus analytic phonics group (four classes, using phoneme-rime awareness training in addition to analytic phonics) and a synthetic phonics group (five classes). The results showed that synthetic phonics

teaching resulted in accelerated reading, spelling and phoneme awareness more rapidly than any other teaching method. Phonological awareness training, while increasing phoneme segmentation ability in comparison with the analytic phonics control group, did not increase reading or spelling skills in relation to this group. This study was also particularly relevant to the current research in relation to underachieving pupils. Gradual analytic phonics teaching produced the highest levels of underachievement, while synthetic phonics teaching resulted in the lowest proportions.

Although the Johnston and Watson studies pointed to encouraging results favouring the use of synthetic phonics, their methodology is marked by a number of significant weaknesses. The first of the three studies (Johnston et al., 1995) drew firm conclusions about phonics teaching that were totally predicated on the methods used, but that did not refer to other possible factors affecting the outcomes observed. Their specific conclusions were: that accelerating the teaching about letter sounds in the middle and final position of words coincided with an increase in word-reading skill; that teaching explicit sounding and blending coincided with gains in reading and spelling; and that gains made in the accelerated phonics/sounding and blending programme were maintained over a two-year period. However, this was not a study that involved any systematic or pre-planned intervention. It was an observational study of phonics teaching in 12 primary schools, during the course of which it was noted that children in one class had noticeably better results, and that in this class the sounds were taught more rapidly and with a greater emphasis on blending.

No account is taken in these conclusions of the fact that these observations related to only one class and that the sample of results being compared with other primary classes was therefore small and possibly unrepresentative. No data were available regarding any pre-test differences that may have been present, and there is no detailed analysis of the meaning of the results cited for this class in relation to the rest of the sample. In addition, teacher characteristics are not taken into account. The fact that one class was using a different approach from all the others to phonics teaching might have been associated with teacher behaviour that differed in ways other than the phonics method adopted. The study did recognise, however, that a crucial factor seemed to be the pace at which phonics teaching took place, but it did not recognise that this might indeed have been the only significant difference affecting outcomes, rather than the distinctive ways in which synthetic and analytic phonics differ in key aspects of method.

The two subsequent studies (Johnston & Watson, 1997; Watson & Johnston, 1998) addressed some of these difficulties, but crucially they did not adequately compare synthetic and analytic phonics on an equal footing. In particular, the methods used in the analytic phonics situations were not necessarily a reflection of best practice, or even of good practice, using this method. The second study took account of the issue of pace of teaching, since the same rate of introduction of new letters was used for both the experimental analytic phonics group and the experimental synthetic phonics group. The specific conclusions were: that synthetic phonics teaching led to better reading, spelling and phonemic awareness than analytic phonics teaching; that this superiority was not due to the fact that in synthetic phonics the letter sounds are

taught at a faster pace, as the fast pace analytic group had less success than the synthetic group; and that the advantage must therefore lie in showing children how to sound and blend letter sounds in order to pronounce unfamiliar words. However, in this study the method of analytic phonics teaching was a limited one. The attention of children was only drawn to letters in the initial position of words and not in the middle or final position, and there is no indication that blending was taught at all. The question must therefore be posed as to whether the advantage of the synthetic approach in this study was its comparison with a limited analytic approach.

The third study had the opportunity of addressing these limitations but did not do so. The same limited approach was used in the analytic groups. However, in addition, pace – which was controlled for in the second study – was not controlled for at all. The four classes in the main analytic group were taught new letter sounds at the rate of one per week, while the five classes in the synthetic group were taught six new sounds in eight days. It is perhaps then hardly surprising that the researchers concluded that ‘synthetic phonics teaching accelerated reading, spelling and phonemic awareness more rapidly than any other teaching method’ (Watson & Johnston, 1998, p. 9).

Despite their methodological limitations, this series of studies highlighted the fact that the actual methods used in phonics teaching affect outcomes, and that synthetic phonics, while not adequately compared with best analytic practice, showed promising potential as an effective method of phonics instruction.

Other studies of synthetic phonics

A number of other studies of synthetic phonics have been carried out. Sumbler and Willows (1996) evaluated the effectiveness of *Jolly Phonics* in senior kindergarten classes in Canada, with children aged just under six years when the intervention started. Children were drawn from 10 experimental and 10 control kindergarten classes from eight schools, giving a sample of 281 comprising 151 experimentals and 130 controls. The experimentals received the *Jolly Phonics* programme, while the control classes had a variety of whole-language based literacy programmes. Pre-test measures consisted of letter names and letter sounds, while post-test instruments included standardised tests of word recognition and spelling, together with a non-word reading test and a non-word spelling test. Experimentals and controls were equivalent on the pre-test measures, but the experimentals were significantly higher on post-test across the full range of measures used. Many of the children were described as having reached ‘a fairly advanced stage of emergent spelling’.

As well as reaching the general conclusion that synthetic phonics is a very effective method to teach beginning reading skills to the typical kindergarten child, the study also considered separately those children viewed as being ‘at risk’. Those whose pre-test scores were in the bottom quartile were selected for this purpose. Again, the experimental children in this group scored significantly higher than the at-risk controls, their results reflecting ‘dramatic improvement’. Indeed, on most tests these children had post-test scores that were higher not only than the at-risk controls but at an equal level of performance with the high-average control group. This aspect of the

study is of particular relevance to the present context of carrying out literacy interventions in populations marked by high levels of socio-economic disadvantage.

Further studies by Willows and her colleagues provided additional support for synthetic phonics as a method for teaching literacy skills to young children. Stornelli and Willows (1998) evaluated the effectiveness of *Jolly Phonics* in a sample of 217 children from public schools in low-income suburban areas of a large city. Three groups were studied – controls, a group who received *Jolly Phonics* in senior kindergarten and a group who received the programme both in junior and in senior kindergarten. Both synthetic phonics groups out-performed the controls, with those who had been on the programme for two years achieving the highest scores.

Kwan and Willows (1998) extended the study of *Jolly Phonics* to at-risk children who were learning English as an additional language. They challenged the assumption that such children are best served in the early stages by restricting literacy development to their native language only. Their findings indicated that experimental children, and especially those who received earlier and more training, were able to overcome their English linguistic weaknesses and significantly out-perform control children whose first language was English.

In a UK context, Stuart (1999) also demonstrated the effectiveness of synthetic phonics using the *Jolly Phonics* approach, again in a population where the majority of pupils (86%) were learning English as an additional language. The 112 pupils in the study were inner-city children who were assigned to one of two intervention groups, an experimental phoneme awareness and phonics group, and a control ‘Big Books’ group. The experimentals received 12 weeks of daily intensive, structured phoneme awareness and phonics teaching during their first year in primary, using *Jolly Phonics*. The controls received 12 weeks of daily teaching for the same amount of time with a more holistic approach. Results at a delayed post-test, 18 months after the end of the intervention indicated that the experimentals had achieved significantly higher scores on tests of initial phoneme identification, phoneme segmentation, letter-sound recognition and recall, word reading, non-word reading, spelling and dictation. On standardised tests the mean difference was 10 months in reading age and 11 months in spelling age.

Follow-up was carried out at 30 months with 101 pupils who could be identified from the original study (Stuart, 2004). Significant differences in favour of the experimentals had been maintained in key literacy skills of reading, spelling and related areas, although no differences in terms of reading comprehension were found. It was also hypothesised that the early experience of the experimentals as successful readers would have led them to develop better attitudes to reading and more positive concepts of themselves as readers. No significant differences, however, were found on a reading self-concept scale.

A number of possible reasons may underlie the demonstrated effectiveness of synthetic phonics programmes other than the basic difference between the analytic and the synthetic approach. Sumbler and Willows (1996) carried out a time management study in which they monitored activities carried out in *Jolly Phonics*

and control kindergarten classes. The key finding was that the children in the *Jolly Phonics* classes spent considerably more time in literacy-related activities than the controls. This applied not just generally but separately to a number of specific activities: letter-sound correspondence, word analysis, auditory phonological awareness, sight vocabulary and basic reading and grammar. While this may indicate that an essential element is the amount of time spent on phonics training and related activities rather than synthetic phonics as such, the multiple regressions carried out for the study pointed to effective features unique to the synthetic approach itself. In particular, the actions associated with learning letter sounds were very strongly related to every outcome measure. This connection of actions to letters and their sounds seemed to have served as a mnemonic device, giving the children an additional means of storing, consolidating and gaining access to the alphabetic coding information.

Most of the above studies had not been published or undertaken at the time that the current study was being planned and implemented. However, the emerging results from the work of the above researchers was of central relevance to the work reported here, and assisted in informing its development as it proceeded.

SUMMARY

This chapter describes two methods of phonics instruction: analytic, or traditional, phonics, in which children are taught letter sounds and use these to break whole words down into their sounds, and synthetic phonics in which they learn to combine letters to make words. As well as differing in this respect, practical differences in the two approaches are also described. Synthetic phonics is normally taught at a much more rapid pace, is associated with high expectations of success and uses methods that involve social and interactive learning. Evaluations of the two approaches are examined, with particular reference to the Scottish context. The emerging results from these evaluations assisted in informing the development of this study.

Chapter 11

The Synthetic Phonics Study

Introduction

The studies of synthetic phonics reported in the literature provided a good basis for considering this method in promoting strand 2 in the intervention strategy, ‘a strong and structured phonics emphasis’. The method also promoted strand 8, ‘lessons from research in interactive learning’. Rather than impose a synthetic phonics programme across the authority (a very significant curricular innovation which not all schools would welcome), or simply make such a programme available to those who wished to take it up, it was decided to use the programme to address a number of key research questions.

First, the reported studies involved interventions with mixed socio-economic groups. It was not known whether similar results could be expected in a large population dominated by high levels of socio-economic disadvantage and educational underachievement. Second, published comparisons between analytic (or traditional) and synthetic phonics did not reflect a context in which good teaching of analytic phonics had already been comprehensively promoted. This is an important consideration since the introduction of any new methodology, such as synthetic phonics, is often associated with a level of interest and enthusiasm absent in the comparison groups. However, in this study all schools were already subject to the excitement created by a high profile programme with a major stress on traditional phonics, supported by a dedicated early intervention team trained to address the key strands in the strategy.

A third question related to longer-term follow up. No longitudinal data were available on the impact of different methods of phonics teaching. The data from the study by Stuart (2004) became available 2001, over a year after the commencement of the study reported here. Although it reported follow up 30 months post-intervention for a synthetic phonics group, the comparison was between phonics and whole language methods and not between different phonics methods. In addition, as 85% of Stuart’s sample were learning English as an additional language, issues arose regarding the comparability of the populations being studied.

Aims of study

The aims of this study were:

- to enhance the systematic teaching of phonics in the early primary years by introducing a more effective methodology than those traditionally used
- to evaluate the impact of the intervention in the longer-term.

Hypotheses

The following hypotheses were proposed in relation to the pupils in the synthetic phonics sample (the experimentals):

- 1 that after intervention they would have higher scores in the key areas addressed by the programme, namely, knowledge of letter sounds and blending skills, and that this would be reflected in word reading abilities
- 2 that they would score higher on tests of spelling
- 3 that these gains would be reflected in the longer term in the mid-primary school years.

The programme

The synthetic phonics method used for the study was the *Jolly Phonics* programme, developed by Lloyd (1992) in *The Phonics Handbook*, together with all the supporting materials produced by Jolly Learning Ltd. This was the programme on which the studies of synthetic phonics reported in Chapter 10 were based. The programme was individually tailored to the needs of the teachers in the authority and was fully supported by an extensive Teacher's Book prepared by the head of the early intervention service. This included detailed assessment sheets for use throughout implementation, and was an essential adjunct to the entire plan of support, monitoring and evaluation. A separate book was available for P1 and P2. Subsequently, to support the curriculum for pupils who had already been taught using this approach, a Primary 3 programme was developed and a separate Teacher's Book provided for this stage.

The available materials provided a complete resource for teaching reading, writing and spelling. The programme addressed five areas for the comprehensive development of key literacy skills: learning the letter sounds, learning letter formation, blending, identifying sounds in words and spelling 'tricky' words. Of these it was essentially letter sounds and blending that made the synthetic phonics approach distinctive from other methods.

For learning the letter sounds, all of the main sounds were taught. These numbered 42 and were adapted from the standard materials, but only to the extent necessary to meet the linguistic context of the population in question. For example, there was not an exact correspondence between Scottish phonemes and the received pronunciation (RP) English phonemes in the programme, such as 'wh' (two phonemes in Scotland, one in RP) and 'oo' (one phoneme in Scotland, two in RP). These sounds were introduced in seven groups of six sounds each. A key aspect of the programme was its multi-sensory approach to learning. Each sound was taught with its own accompanying action to reinforce learning.

For blending, the focus was on the synthetic approach to word building rather than the analytic approach to word dissection, as described in Chapter 10. Once a range of letter sounds was learnt these sounds were then used to build up words. The first group of letters – *s, a, t, i, p, n* – were central to the process because they could be used to make more simple three-letter words than any other six letters. The encouragement of these encoding abilities provided the foundation for enhancing the

decoding skills required to break down whole words into their component phonemes and blend these together so that they could be correctly identified.

It is important to recognise that this was a comprehensive programme for the teaching of early literacy skills. The programme did not exist only in its distinctives but also in its commonalities with other methodologies. Thus, like other programmes, it was addressing the universals of learning how to write the individual letters and later join them to form cursive writing, of how to learn irregular word recognition by 'look and say' methods, of the development of listening skills and of fostering interest in books and stories.

Method

Design

This was a quasi-experimental study using a range of both quantitative and qualitative measures, with pre-post testing carried out through a four-year period.

Sample

Primary 1 classes in 18 primary schools were selected for the study (9 experimental, 9 controls). This gave a total sample size of 590 pupils in Primary 1 classes, comprising 315 experimentals and 275 controls. A quasi-experimental approach was required as randomised selection of establishments would not have been a practicable possibility. The introduction of a new and highly structured phonics system requiring significant staff training represented a major curriculum change. It could therefore be expected to work successfully only on a volunteer basis to ensure appropriate levels of co-operation and high fidelity of programme delivery.

To this end, all primary schools in the West Dunbartonshire authority were approached during Session 1997-98 and given background information on the Jolly Phonics programme. This included a presentation to all head teachers, together with supporting literature and a video produced by the publishers. Nine primaries out of the 35 came forward as volunteers and these were accepted into the study as experimental schools for commencement of the programme at the start of Session 1998-99. Each of these was matched as closely as possible with a control school in terms of socio-economic status and available information on reading levels. The latter was based on existing knowledge of how these schools were generally performing in literacy in the early years, together with comparisons of key scores for the current Primary 1 classes on the 1997 baseline assessment tests.

For SES matching the criterion used was FSE (eligibility for free school meals). For the experimental schools the mean weighted FSE was 44% (range 34%-62%), while for the controls it was 42% (range 20%-57%). The matching for reading levels using the 1997 baseline assessments (that is, the Primary 1 cohort for the year group prior to the intervention) showed no significant differences between experimental and control schools in the three key areas of letter sounds, non-word reading and word reading. However, it was noted that the trend of scores tended to favour the control

schools in relation to letter sounds ($t = 1.788$, $p = 0.07$, two-tailed) and word reading ($t = 1.552$, $p = 0.12$, two-tailed). These, were independent two-sample t tests calculated using the data analysis tools on Microsoft Excel Version 8.0. This trend applied across all baseline tests, an observation that was of importance later in establishing closer comparisons between the two groups.

Implementation

The programme commenced for Primary 1 pupils in August 1998 at the start of the new school session. All of the teachers involved were provided with initial training in the Jolly Phonics approach, together with all required materials. In addition, a rolling programme of follow-up training was provided. This covered areas such as planning, assessment, recording, monitoring and evaluation. Support structures were also put in place to ensure that the overall programme was carefully monitored at all times, and that teachers could have an immediate response to any enquiries about operational procedures. In addition, a group meeting was held once a term for the experimental staff. This allowed them to provide feedback on the operation of the programme in their own schools, to raise any difficulties being encountered and in general to share experiences.

When general problems were noted in the group meetings, agreement was reached regarding any further modifications to how the programme would be implemented. The most recurrent concern related to pacing, as it was generally felt that the rate at which the phonemes were being introduced was too fast for this population. The programme had been developed in a much more advantaged area where reading difficulties did not have the same prominence. Appropriate adjustments were therefore made to this and to other aspects.

The group meetings also served another very important function. Although the schools selected for the study were volunteers, this often meant that the volunteer was actually the head teacher or the school management team and not necessarily the class teachers themselves. It was apparent from the start that the teachers varied considerably in their commitment to this method of teaching phonics, and some attended the initial training with considerable scepticism. Had their views been general within the group it is most unlikely that the project could have proceeded successfully. However, these issues were dealt with not so much by the reassurances of the researcher and project leader but by the other group members themselves. Many of the teachers found almost immediately that their pupils were responding to Jolly Phonics with more enthusiasm and success than they had ever known before. When these teachers encountered difficulties in the programme they found effective means of tackling them, and were not distracted from the main task of continuing with the strategy. Their experiences were crucial in supporting those who had begun with low commitment, and in due course the whole group was viewed as a highly motivated and committed one.

Implementation continued throughout Primary 1. The whole process of training and support was undertaken again for the Primary 2 teachers when these children moved on in Session 1999-2000. It was considered essential that, having begun with

synthetic phonics, the phonic instruction for this group should continue by the same method. Further comparisons among schools, however, became increasingly difficult because of the extent of contamination of the control schools with the experimental method. The experience of the experimental teachers, together with the dissemination of the early results half-way through the first year of implementation, meant that West Dunbartonshire was unstoppably becoming a Jolly Phonics authority. Within a short space of time virtually every primary school was teaching phonics by this method.

Measures used

For the first series of comparisons the entire sample was used. That is, the 18 schools were compared on the basis of the performance of all pupils in Primary 1 for the first year of implementation, then of all pupils in Primary 2, Primary 3 and ultimately Primary 4 in the succeeding years. In Primaries 1 and 2 the baseline assessment results were used, while the Norman France results were used at Primaries 3 and 4.

Two additional sets of comparisons were made using different sampling methods. First, since the baseline assessments throughout the authority were conducted mid-way through the session, it was felt useful to make some further comparisons of progress at the end of the first complete year of the study for the Primary 1 children. For this purpose, individual tests were carried out with all of the experimentals (N = 315) and a systematic sample of controls (N = 80). This was obtained by selecting every *n*th pupil from the classes in the control schools to give a sample of 10 from each. As the smallest school had only seven P1 pupils, and the school indicated that they were very unrepresentative of their normal P1 pupils, they were not included in this comparison. This sample was tested using two baseline tests, non-word reading and word reading, plus a spelling test constructed for the purpose.

The spelling test is shown at Appendix 3. Just as the overall baseline assessment scheme was designed to meet the requirements of the study, so also the spelling test was designed to meet the needs of the narrow age group of young children for whom it was being used. In particular, it was constructed to give a low floor and a high ceiling for Primary 1. Although not designed as a standardised instrument for wider use, the indications were that it met good reliability and validity criteria. In terms of reliability, the test forms of a sample of 354 children pointed to high internal consistency, with a split-half reliability co-efficient of 0.90, using alternate items. It seemed clear that the test also had good face and content validity. It directly measured the self-evident skills that the intervention was designed to enhance in relation to spelling. Before use, the test was assessed as to its suitability by several teachers with general and specialist experience of working with the age group in question. In addition, it was carefully graded using the difficulty levels in a number of existing spelling tests as a comparison.

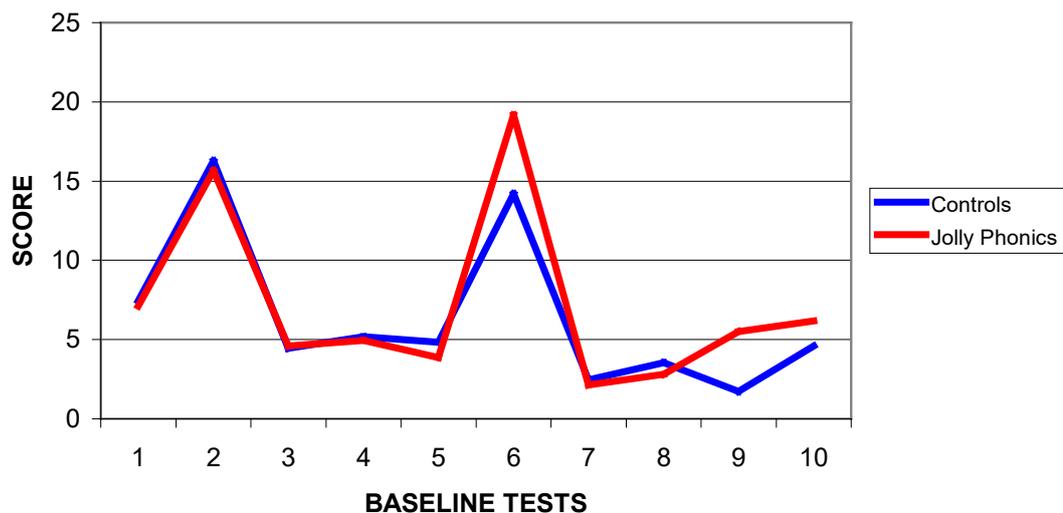
Second, the mismatch between teachers' perceptions of the progress of the Jolly Phonics sample through the years and some of the actual test results (as discussed later in the summary of the results) led to a further set of more precise comparisons being made. It was realised that the original sample had suffered from considerable

dilution by the time they reached P3 and P4. In some of the schools a significant number of pupils had moved in or out of the area, so that many of the children ultimately were not part of the original Jolly Phonics sample. To control for this, the 1997 baseline assessments were searched to identify every known child who became part of the study in the experimental and control schools in the following year. These pupils were then followed through to identify those for whom scores were available in P3 in 2001 and P4 in 2002. This reduced the sample from 590 in the original P1 classes to 286 identified at pre-school, P3 and P4 (146 experimentals, 140 controls). These pupils were then matched in pairs by their total baseline assessment scores prior to the intervention in nursery school. Only those who could be exactly matched in this way were included. This reduced the sample further to 180 (90 matched pairs), and allowed a direct comparison to be made in P3 and P4 for pupils who had identical scores at pre-test in nursery and who were then exposed to two different methodologies.

Results

Experimentals and controls were first compared on the Primary 1 baseline assessments in 1998. The tests were conducted during November and December, meaning that the synthetic phonics sample had been on the programme at that stage only for three to four months, that is, from the start of the session in August. The previous year's trend for the control schools to be scoring slightly higher across all tests, as already noted, had not changed except in relation to the three key tests that were the subject of the first hypothesis, namely, letter sounds, non-word reading and word reading. At this early stage the experimentals were significantly ahead of the controls in these three tests alone (independent two-sample *t* tests): letter sounds ($t = 10.845$, $p < 0.001$, one-tailed, effect size = 0.62), non-word reading ($t = 7.933$, $p < 0.001$, one-tailed, effect size = 0.75) and word reading ($t = 2.567$, $p < 0.01$, one-tailed, effect size = 0.23). The results are shown in Figure 11-1.

Figure 11-1 P1 baseline assessments: Jolly Phonics v controls



Tests: 1 Concepts of print 2 Nursery rhymes 3 Initial sounds 4 Rhyme detection 5 Rhyme production

* p<0.05

** p<0.01

*** p<0.001

Of the three tests highlighted above, the one that most centrally reflects the whole aim of synthetic phonics is non-word reading. It is the aim of fostering this blending skill that principally distinguishes the programme, and it is in relation to this area that the results may be illustrated most graphically, as shown in Figure 11-2.

It was possible to make a further comparison of the Primary 1 classes at the end of the session in June 1999 to determine whether these early gains were being maintained throughout the first year of intervention. For this purpose the pupils in the experimental classes were compared with the 80 pupils in the stratified sample described above on tests of non-word reading, spelling and word reading. The results are shown in Table 11-1. The experimentals were still significantly ahead of the controls for non-word reading and spelling. However, the scores on the word reading test were almost identical. It should also be noted that the large size of the sample allowed modest differences in scores to give large statistical significance.

Figure 11-2 P1 baseline assessments: non-word reading scores

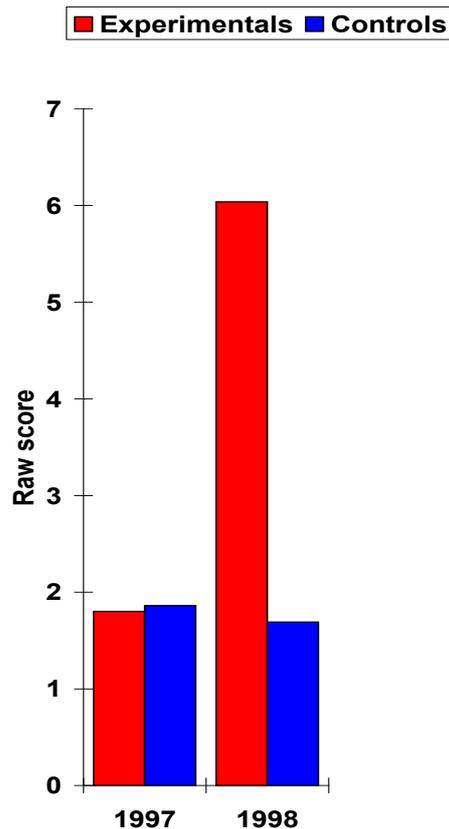


Table 11-1 Non-word reading, word reading and spelling after one year (end of P1)

	Experimentals	Controls	Significance
Non-word reading	M = 15.10 (N = 309)	M = 12.73 (N = 80)	t = 8.586 *** SD = 4.96 Effect size = 0.48
Word reading	M = 22.38 (N = 309)	M = 22.93 (N = 80)	t = -0.848 <i>ns</i>
Spelling	M = 8.72 (N = 309)	M = 7.69 (N = 80)	t = 2.424** SD = 3.38 Effect size = 0.30

* p<0.05 ** p<0.01 *** p<0.001 (one-tailed tests)
Independent two-sample *t* tests

In the comparisons using the baseline assessment results of the entire sample at Primary 2, only two of the tests, non-word reading and word reading, allowed useful comparisons because of ceiling effects on the easier tests at this stage. By Primary 2 most pupils were scoring near the ceiling on all the tests of phonological awareness, and the overall impact of the early intervention meant that with the letter sounds also almost every child was achieving virtually the maximum score. The results continued to favour the synthetic phonics schools for non-word reading. However, there were significant ceiling effects limiting the possibility of larger differences between the groups as both groups on average were by this time scoring almost all of the items. Again the differences, although significant, were modest, as seen in the small effect size. Once more, it was on the word reading test that no significant differences were found between the groups at this stage. The results are shown in Table 11-2.

Table 11-2 Non-word reading and word reading in Primary 2

	Experimentals (N = 291)	Controls (N = 269)	Significance
Non-word reading	M = 16.87	M = 15.90	t = 3.045 ** SD = 6.35 Effect size = 0.15
Word reading	M = 32.39	M = 32.47	t = -0.067 <i>ns</i>

** p<0.01, one-tailed
Independent two-sample *t* tests

The hypothesised difference between the groups in longer-term reading gains in the mid-primary school years was not demonstrated in the results of the Norman France Reading Tests either at Primary 3 in 2001 or at Primary 4 in 2002. The mean experimental P3 score was a reading age of 7y 2m compared with 7y 4m for controls ($t = -1.870$, *ns*), while at P4 the figures were 8y 1m and 8y 2m respectively ($t = -1.002$, *ns*).

It was this lack of a longer-term difference between experimental and control schools that led to the final set of comparisons being made, using the 180 children from the original sample who could be exactly matched in pairs by their pre-school baseline assessment results and for whom subsequent results were available in P3 and P4. The results are shown in Table 11-3.

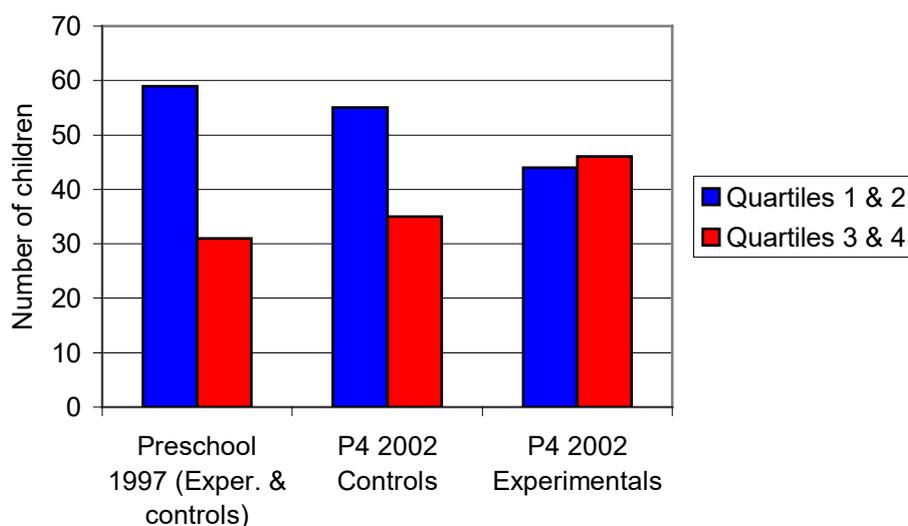
Table 11-3 Pre-school results compared with reading ages in P3 and P4 (N = 90 matched pairs)

	Experimentals	Controls	Significance
Matched pre-school Baseline assessments 1997 (total scores)	18.37	18.37	N/A
P3 reading age 2001 (Norman France)	7y 3m	7y 4m	$t = -0.294$ <i>ns</i>
P4 reading age 2002 (Norman France)	8y 4m (Raw score 32.6)	8y 0m (Raw Score 31.0)	$t = 1.945$ * SD = 5.6 Effect size = 0.46

* $p < 0.05$, one-tailed
Related two-sample *t* tests

It is noted that while no significant differences were found for the sample at the Primary 3 stage, by Primary 4 the original groups had again diverged in their scores in favour of the experimentals. When these P4 results were broken down by quartiles, the pattern shown in Figure 11-3 emerged. The distribution of the P4 controls in relation to the results for all schools in the authority was very closely matched to the way these same children's scores were distributed at pre-school ($p = 0.374$ *ns*, chi-square). However, the same comparisons for the experimentals showed that there had been a significant re-distribution of their scores from being in the 1st and 2nd quartiles at pre-school to being in the 3rd and 4th quartiles at P4 ($p < 0.001$, chi-square).

Figure 11-3 Distribution of sample by quartiles: pre-school and P4 (N = 90 matched pairs)



Qualitative results

Extensive qualitative data were also obtained from the staff in these schools. This was done in two ways. First, during the early years of the project feedback was obtained from all involved teachers at the regular group meetings. Second, a questionnaire was completed by teachers early in session 2003-04, five years after the original study commenced.

The ongoing feedback from the Primary 1 and Primary 2 teachers provided very strong support for the effectiveness of the programme, since teachers were virtually unanimous in asserting that their pupils were working at higher levels of skill than had ever been known before. This view was expressed so universally that the nine volunteer schools were joined within a year or so by virtually every other primary, so that the area became, in effect, a ‘synthetic phonics’ authority. Box 11-1 shows a small sample of the comments that were typically made both in P1 and in P2.

Box 11-1 Synthetic phonics: feedback from meetings of P1 and P2 teachers

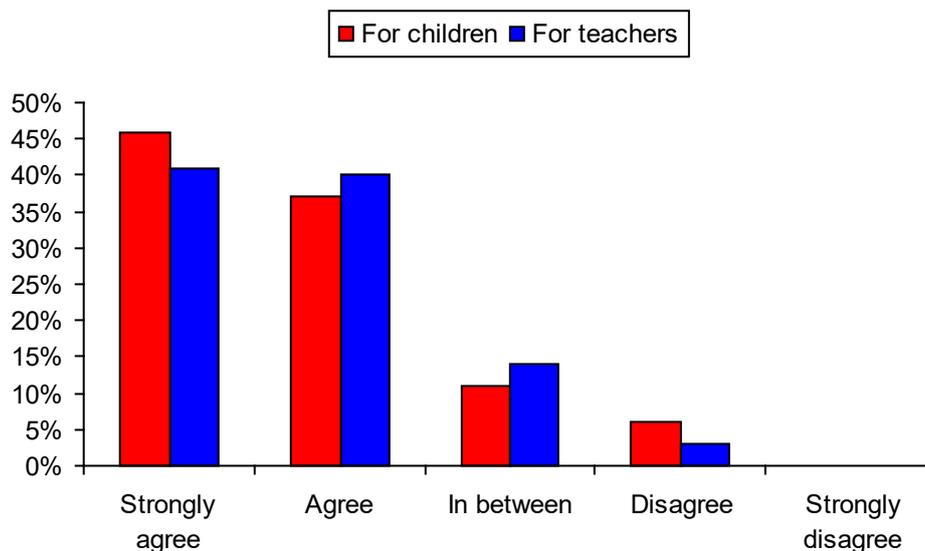
‘Even the lowest child did far better than could ever have been done before’
‘Absolutely no comparison. Even the very poorest are achieving’
‘Results are still good after the holidays when it was thought it would have gone’
‘Parents said they saw a big difference’
‘All parents felt the children were highly motivated’
‘We are very surprised at how quickly children have learned – and have retained what they have learned’
‘Children love it’
‘We couldn’t have stopped Jolly Phonics – the parents wouldn’t have let us’
‘The children are writing a lot more now compared with the other children’
‘There is positive feedback from parents and a smaller tail of children struggling’
‘Because I’m feeling so confident it now rubs off on the children’

‘They’re so enthusiastic – always asking for another sound’
‘There’s now a great buzz around the place’

It was recognition of the ‘buzz’ that Jolly Phonics had created that determined the need to obtain further qualitative feedback at a later stage, once the programme was so well bedded down in all participating schools that it ceased to be viewed as new or different. Any well-resourced new initiative that pupils and teachers find enjoyable is likely to experience a considerable Hawthorne effect in the early stages of operation. The period selected for obtaining further feedback was a full five years after the introduction of the programme, in the first half of session 2003-04. All participating teachers were sent a questionnaire asking them to rate the programme on a five-point scale in comparison with other methods. The criteria selected included teacher and pupil enjoyment, effects on progress in writing and in other subjects and value of the programme for higher or lower achievers, together with questions on what was liked best or least about the programme, or found most or least helpful. A copy of the questionnaire is shown in Appendix 4.

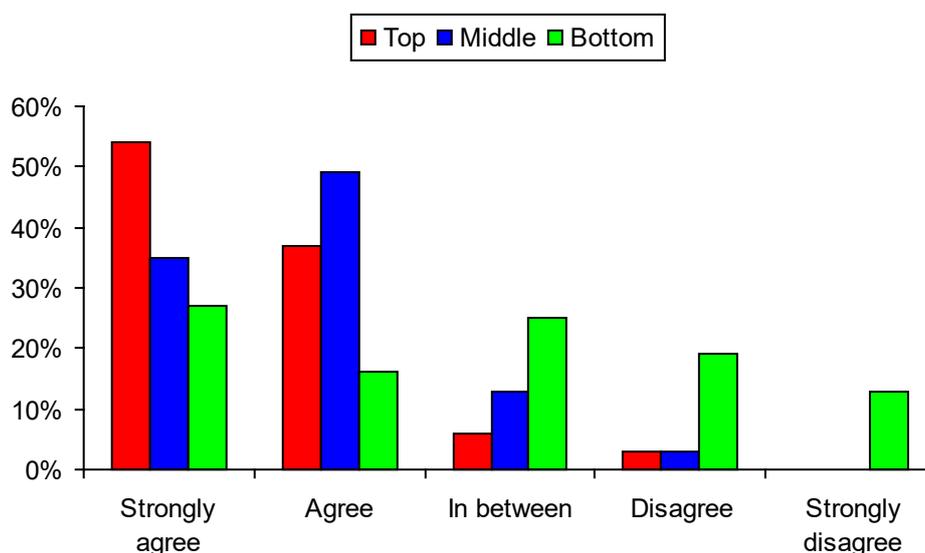
A total of 64 teachers of Primary 1 and Primary 2 classes responded. The following proportions of respondents agreed or strongly agreed with these statements regarding Jolly Phonics: that children enjoy it more than usual methods (83%); that teachers enjoy it more (82%); that it is good for children at the top end of the class (91%) and in the middle range (84%); and that children become better in their writing (75%). What was particularly noticeable on these items was that almost every respondent who did not actually agree with the statement chose the ‘in between’ category, with only a very small number selecting ‘disagree’ and no one selecting ‘strongly disagree’. In addition, most teachers believed that they would not get as good results using any other phonics approach (70%), and just over half believed that Jolly Phonics led to children becoming better in other school subjects also (52%). Figure 11-4 shows teachers’ views of the programme as being enjoyable for children and teachers in comparison with other methods.

Figure 11-4 Jolly Phonics is more enjoyable for children/for teachers



The questions that received a more mixed or negative response were those relating to pupils who were likely to fail. Just over half were of the opinion that with Jolly Phonics fewer children went on from P1 to P2 with poor reading skills (56%), and 14% disagreed with this statement, none of them strongly. Most believed that those who did fail with Jolly Phonics in P1 would find it even more difficult to keep up in P2 (77%). There was a very mixed response to the statement, ‘Compared with usual methods, Jolly Phonics is good for children at the bottom end of the class’. While 43% agreed with this (27% of them strongly), 32% disagreed (13% of them strongly). Figure 11-5 shows teachers’ views of the comparative benefits of Jolly Phonics for pupils at different levels of ability.

Figure 11-5 Jolly Phonics is good for children at the top/middle/bottom of the class



The overall views of teachers who had been using Jolly Phonics therefore provided strong support for the method over other approaches in terms of enjoyment and

effectiveness, but with significant concerns over the poorest children, especially of their chances of catching up if they had failed in the early stages. These views were broadly similar whether the respondents were teaching Primary 1 or Primary 2, and whether they had been teaching Jolly Phonics from the start of the project or only for a year or two.

Of the 64 teachers who responded, 61 answered the questions on what they liked best and least about Jolly Phonics (or found most helpful or unhelpful). The structure and timetable of the programme were cited most frequently as being liked best, and the fast pace, especially for poorer children, was cited most frequently as being liked least. Four respondents wrote that there was nothing about the programme they did not like, while another 10 left this question blank. Thus, about a quarter of the respondents to this section did not identify anything they did not like or found unhelpful. Table 11-4 shows the most frequently cited aspects of the programme for these questions. Some respondents identified more than one item for each question.

Table 11-4 Most and least liked (or helpful) features of Jolly Phonics (N = 61)

<i>Most liked/helpful (number)</i>	<i>Least liked/helpful (number)</i>
Structure and timetable (22)	Pace too fast (especially for poor readers) (35)
Actions/multi-sensory approach (16)	Assessments (too time consuming) (12)
Attractive materials/resources (13)	
Fun/enjoyment (10)	

The structure and timetable of the Jolly Phonics programme had been carefully adapted by the early intervention head teacher to meet the agreed needs of the project. This was particularly appreciated by teachers, whose comments frequently reflected the fact that the programme books prepared for them were well structured and easy to follow.

Discussion

Within a context in which a structured phonics approach, whether analytic or synthetic, has played a crucial role in a multiple-component intervention programme, this study provided a confident basis for supporting synthetic phonics as an effective strategy for enhancing early literacy skills. The initial results obtained in favour of the experimentals, together with consistent qualitative feedback from teachers, demonstrated that the synthetic phonics programme was doing exactly the job it was designed to do. It was accelerating early reading skills in the key areas of knowledge of letter sounds and blending and was enhancing spelling abilities.

These results are consistent with the trend of research studies reported in Chapter 10. In the current study there was an opportunity not only to consider synthetic phonics in its own right as an effective method of phonics instruction, but also to evaluate its

impact in a context where traditional phonics teaching had already been enhanced as part of a literacy intervention. The programme was therefore being compared with good models of analytic phonics instruction, and was still showing superior outcomes.

It was in the early stages of literacy that the comparative advantages of the programme were most conspicuous, with significantly better blending abilities being demonstrated by the experimentals in the middle of their P2 year, that is, about a year and a half after the introduction of the programme. More caution is required in drawing conclusions about the differential impact of the programme on reading skills in the later years of primary school, and its effects on word recognition. The pattern here was that the experimentals were significantly ahead of the controls in word reading at the time of baseline assessment in Primary 1, but that this advantage was not sustained in Primary 2. It was also noted that a rather confusing pattern of results had emerged in Primaries 3 and 4. The experimentals showed no advantage over controls in Norman France Reading Test scores in Primary 3, even when the comparison was made with the most closely matched sample available. Nevertheless, on the same test in Primary 4, a significantly better outcome for the experimentals had again asserted itself.

These results merit further comment. First, it is not unknown for a real advantage to be present but to remain undetected in a sample of children at one stage, but to show itself when the same sample are tested on the same skills at a later stage. This was apparent in the study by Muter, Hulme, Snowling and Taylor (1998) of two different aspects of phonological awareness, segmentation and rhyming, as predictors of later reading and spelling progress. Segmentation was strongly correlated with attainment in reading and spelling at the end of the first year at school, while rhyming was not. By the end of the second year, however, rhyming had started to exert a predictive effect on spelling. The authors concluded that segmentation and rhyming are probably predictive of quite different aspects of the spelling process. Abilities with rhyming do not particularly support the approaches to spelling used in the earlier stages, but are more relevant to approaches children use later. They cite similar finding by Wimmer, Landerl and Schneider (1991) in their longitudinal study of early literacy development in German-speaking Austrian children. They found that rhyme awareness was only minimally predictive of reading and spelling achievement at the end of Grade 1, but that it gained substantially in predictive importance in Grades 3 and 4.

Further study would be required to ascertain how this might relate to the pattern of results obtained on the Norman France tests, but at least these studies provide additional illustrations of underlying differences that are not reflected in scores at one point but are then found at a later point. It has been noted in the discussion of the main study results, for example, that the Norman France tests are administered as 'comprehension tests', and that there is a prominent cognitive element in addition to testing reading skills as such. It may be that these comprehension factors play a more significant role with the younger children in Primary 3, and that at this stage they are not necessarily the best measures for tapping into real differences in the key skills addressed in the synthetic phonics programme. These differences may be obscured at

P3, but may come to the fore again with the older and more cognitively mature P4 children.

Another key factor, however, was the early contamination of the programme, as mentioned above, since it became clear that the dramatic changes about which teachers were enthusing led to diffusion of the synthetic phonics methodology across the authority in ways that the study could not reasonably control. This meant that the first set of comparisons gave a pure measure of two methods, but that subsequent comparisons became weaker. Certainly, the baseline tests in Primary 1, being carried out just three or four months into the programme, came at a point when the experimental and control schools were quite distinct in having either a synthetic or an analytic methodology. As time went on these distinctions became less certain.

Finally, it became clear that the teachers in P3 and P4 were not prepared for the enhanced abilities of the children who came to them from the Jolly Phonics programme in P2, and it was likely that the differences became considerably diluted at that stage through not capitalising on the gains already made. The Primary 3 reading curriculum in its structure and content remained as it had been in the preceding years. Some of the P3 teachers made specific comment that they were receiving children from the P2 Jolly Phonics classes who were throwing their normal curriculum into chaos, because they had a quite different starting point from what was normally expected. A planned response to this changing context did not take place until a later date, with the revision of the approaches taken in P3 and the introduction of an upwards extension to the Jolly Phonics programme.

The overall combination of the quantitative results of this study, cautiously interpreted for the reasons indicated, and the qualitative feedback from staff provided good ground for confidence that a synthetic phonics approach would result in enhanced outcomes compared with traditional phonics methods. The fact that Primary 3 teachers, although not part of the study, volunteered their views on the impact of the programme on the pupils coming up from Primary 2 was in itself a fairly obvious indicator that this programme was having successful results. The enjoyment of the children in following the programme, and its widespread endorsement by parents, many of whom were able to make direct comparison with the experience of older siblings using traditional methods, similarly combined to support this approach.

The hypotheses that children on the programme would make gains in key areas, and that these would continue to be reflected in the longer term, were generally supported, although highlighting areas for further consideration. The conclusion was that this was a well-structured and highly successful synthetic phonics programme. It provided for the teaching of key phonic skills at a fast pace using a multi-sensory approach, and in a way that capitalised on social interaction in learning. The result was that almost all of the primary schools in the authority adopted the synthetic phonics approach.

SUMMARY

This chapter presents a quasi-experimental subsidiary study designed to take forward a key strand of the intervention, ‘a strong and structured phonics emphasis’. Its aim was to enhance the systematic teaching of phonics in the early primary years by introducing a more effective methodology than those traditionally used. The programme adapted for the study, *Jolly Phonics*, is described. The sample was 590 Primary 1 pupils in 18 schools (9 experimental, 9 control). The study was supported by a comprehensive programme of training, monitoring and support. Baseline assessment tests and a spelling test designed for the study were used to evaluate the intervention. For follow-up, baseline assessments results the following year in P2 were used, and Norman France group tests in P3 and P4. Qualitative feedback from teachers using the new method was obtained during monitoring and support sessions, and questionnaires were completed by P1 and P2 teachers five years later. The results supported the use of synthetic methods as an effective approach to phonics instruction.

Chapter 12

The Attitudes Study

Introduction

The attitudes study served to develop strand 10 of the intervention strategy designed for the main study – ‘enhancing attitudes, values and expectations’. It represented a five-year follow up to the Edinbarnet Reading Project, a full account of which is provided in the preparatory studies outlined in Chapter 5. The original project, a randomised control trial carried out in one of the schools that became part of the main study, was designed to address endemic problems of reading failure. It was based on a clear rationale underlying the idea of addressing attitudes and values, namely: that there is an established association between social disadvantage and educational underachievement; that children in areas of social disadvantage differ from those in higher socio-economic groups not only in terms of early childhood educational experiences but also in terms of educational attitudes and values; that there is an association between attitudes and values and educational achievement; and that by seeking to change attitudes and values an opportunity may be created to enhance reading achievement in this group of children.

The intervention in the original project involved no extra reading tuition or modifications to the reading curriculum. Instead, it focused on changing attitudes and values towards reading within a groupwork context. At the end of a 10-week intervention, significant changes in attitude in favour of the experimentals were found, with children becoming more positive in their views of school and of reading. At the same time there was a significant enhancement in the reading achievement levels of the experimental groups. These showed increases of approximately one year in reading accuracy scores (range 4-21 months) and one and a half years in reading comprehension (range 3-30 months). Qualitative feedback from teachers, parents and pupils also reported improved attitude, higher confidence levels and more reading taking place at home, with teachers indicating a ‘dramatic’ change.

The conclusion of the study was that the whole area of attitudes, values and expectations was of crucial importance in relation to reading and to reading failure. At the same time, it was recognised that results obtained from short-term interventions in schools to raise achievement must be viewed with caution in the absence of longer-term follow-up data. Gains have often been recorded for brief interventions, but it has been a consistent finding over a very long period of time that in the case, for example, of remedial reading programmes, these have not been sustained after the special help is withdrawn (Carroll, 1972). Nevertheless, the Edinbarnet Reading Project was not based on a remedial programme or any other form of direct help with reading. It aimed instead to bring about changes in values and attitudes and to create a new educational context in which a number of previous obstacles to the learning of skills would be removed. It was considered that this might provide a more lasting basis for long-term change than a short period of direct teaching of reading skills.

It was for the purpose of long-term follow up that the attitudes study was carried out. The children who had taken part in the original project were traced to their secondary schools five and a half years after the initial study had taken place. At the time of the original study they were around the middle years of their primary schooling, and at follow up were around the middle years of secondary schooling. All were individually assessed for their level of literacy skills and for their attitudes to school and to reading, allowing new comparisons of experimentals and controls to be made.

The place of attitudes, values and expectations

The significance of psychological factors such as attitudes, values and expectations, together with the associated areas of self-concept and confidence, has not been to the fore in the study of reading and of reading failure, especially in recent years. In earlier years there was a period of interest in these factors, leading to a number of reviews. Thomas (1980) summarised the research on self-concept and educational achievement. The relationship was complex and difficult to interpret in causal terms. Nevertheless, he concluded that self-esteem is integral to academic performance, and an increase in the one may be expected to lead to an increase in the other. A review by Gurney (1987) examined controlled experimental studies that have sought to enhance children's self-esteem in classroom settings. He reviewed curriculum packages directly aimed at self-esteem enhancement, interventions designed to improve academic performance, counselling interventions and attempts to change teacher or pupil behaviour. Among factors considered to be important he identified the powerful influence of parental contributions, the value of making positive statements about oneself and the advantage of working on both self-esteem and educational skills at the same time.

Two of the earlier studies are of particular interest because they examine self-esteem and attitude specifically in relation to reading. Wattenberg and Clifford (1964) noted the association between poor self-concept and reading disabilities and wished to determine which was the antecedent phenomenon. As this is one of the few available studies relating to this specific area, and as the authors claimed that their 'findings seem to have far-reaching import' (Wattenberg & Clifford, 1964, p. 466), it merits further discussion. The authors put forward a number of hypotheses. The most central of these were: that measures and ratings of self-concept taken at kindergarten entry stage would be predictive of later achievement in reading; that self-concept would be better than intelligence test scores in predicting later reading achievement; and that self-concept would be more important in predicting achievement than achievement would be in predicting changes in self-concept, thus pointing to self-concept as a determinant of achievement rather than vice versa.

Based on assessment carried out at kindergarten entry and then again two and a half year later the authors concluded that the results gave considerable support to their main hypotheses. The study is presented as being 'exploratory', but it is a very complex one with a wide range of findings that are confusing and difficult to interpret as to their real significance. The sample was quite a large one, comprising 185 children, but it is noted that over 30% of the original number had been lost by

the time of follow up. Descriptions are provided of attempts to control as much as possible for sex and socio-economic status, but these descriptions are not supported by presentation of any data that would allow the robustness of the methodology to be assessed.

The measures employed to assess the variables used in the study, other than the intelligence test (the Detroit Beginning First Grade Intelligence Test) also raise issues as to their validity and appropriateness. For reading achievement, instead of using a single standardised instrument, a decision was made to use two different tests produced by the publishers of the two textbook series in use. This meant that the whole sample could not be compared on the same terms. For this reason the sample was divided by textbook series used as well as by sex and socio-economic status, but without data to indicate the numbers or constitution of each of the 'homogeneous groups' arising from this. Defining 'self-concept' was also approached in a manner that led to a complex sub-division into different aspects (for example, 'quantified self-concept – competence', 'quantified self-concept – good/bad' and 'ratings of ego strength').

It is probably because of this complexity in defining concepts and constructing assessment measures that the various sub-groups of the sample, with their unknown numbers, yielded many results that were described only as 'positive' but few that reached significance in relation to the hypotheses being tested. In addition to this, because again there are inadequate published data to support interpretation of the various analyses recorded, it is difficult to assess the strength of the findings. At the same time, the large number of correlations described in favour of the main hypotheses tend to support the authors' conclusion that 'although the levels of significance were far from overwhelming, the consistency in results was striking', and as an exploratory study this investigation has made one of the few specific contributions to this area.

Lawrence (1971) produced the first in a series of studies of the effects of counselling on children with reading difficulties. While working as a remedial teacher he had concluded that the child's emotional life is the most crucial factor in determining progress in reading. He clearly articulated the view that direct reading instruction is not the only consideration in dealing with reading failure, and indeed that there may be central factors of a psychological nature:

'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. Apart from the recognition that a generally sympathetic and encouraging attitude is desirable, there is rarely, if ever, a systematic plan to deal with the child's emotional state, and more specifically, his self-image' (Lawrence, 1973, p. 11).

His studies aimed to improve self-image through systematic individual counselling (Lawrence, 1973). Although serious criticisms were made of his methodology and

conclusions (Pumfrey, 1979), the importance of recognising emotional and attitudinal factors and seeking to change them was acknowledged.

While the strong relationship between socio-economic status and low achievement in reading and other educational outcomes is clearly established (Chapter 4), and the importance of factors related to attitudes has been recognised, the question of whether children and young people from lower socio-economic backgrounds do in fact have poorer attitudes to school and education requires to be considered further.

A number of studies have addressed this issue over a long period of time. Fogelman (1983) reported on the findings arising from the 1958 Cohort of the National Child Development Study when the children were 16 in 1974. Over 8,000 pupils responded to the question, 'I feel school is largely a waste of time'. While 80% of all pupils rejected this statement, there were significant differences between those from higher and lower socio-economic levels, with more of those of higher SES rejecting it. Similar findings were reported by Cox (1983) in a longitudinal study of disadvantaged pupils. He followed the school progress of approximately 50 children from disadvantaged homes and compared it with a socially advantaged group attending the same schools and matched for age, sex and non-verbal intelligence. In the follow-up phase of the study when the children were age 15+ they completed an attitude to school scale and other measures of self-esteem and self-concept. The disadvantaged pupils scored significantly lower on all measures. They had less positive attitudes to school and learning and lower levels of academic self-concept and self-esteem than their more advantaged peers. However, the differences in attitude were not as pronounced as the differences in actual levels of educational attainment.

In a national study of attitudes to school among secondary school pupils, Keys and Fernandes (1993) examined the responses of approximately 2,000 pupils to written attitude questionnaires. They concluded that there was a relationship between attitude and SES but that it was relatively weak. However, they used a limited range of SES indicators, namely, the cultural level of the home based on the pupils' own estimates of the number of books at home, and the type of catchment area served by the school. Further evidence of a relationship between SES and attitude was found in a MORI survey of attitudes to learning in a sample of over 1,000 young people and adults in England and Wales aged 16 years and onwards (Campaign for Learning, 1998). Those from lower class households were less likely to show a desire to be involved in learning and were also less likely to rate learning as being either important or enjoyable.

These studies show a consistent trend of findings over the course of about a generation. They may be summarised as follows: there is a relationship between socio-economic status and attitudes, with those from lower socio-economic groups showing less positive attitudes to school and to learning and poorer self-concept; the relationship is not as strong as that between low SES and actual learning outcomes; and although the processes involved are interactive, in that poor results at school adversely affect attitudes and self-concept, there is evidence that attitudes and self-

concept are predictors of later outcomes in reading and general learning even when intelligence is controlled for.

Hypotheses

The hypotheses for this study were:

- 1 that the experimentals would have higher levels of reading accuracy and reading comprehension than the controls
- 2 that the experimentals would have more positive attitudes to school and to reading than the controls.

Method

The method adopted for this follow-up study was very simple. The 24 children who had participated in the Edinbarnet Reading Project in late 1993 and early 1994 were traced to their various secondary schools in May 1999. Of the original sample, 20 were identified in three secondary schools and one special unit. Of the remaining four, three had left the area and one could not be traced. All of the pupils were in Primary 4 or Primary 5 when the original study took place, and they were now in their second or third year of secondary schooling.

Of the 20 pupils who could be traced, 19 (11 experimentals, 8 controls) were tested again on the Neale Analysis of Reading Ability, both for Accuracy and Comprehension scores. The child in the special unit was excluded because of 'caseness'. He had shown significant problems in his development and adjustment and had been admitted to a unit for pupils with autistic spectrum disorders attached to a secondary school, although the exact nature of his difficulties had not been clearly identified. It was felt appropriate not to include him in the follow-up sample. All but one of the sample also completed the attitude questionnaire, 'What you think about school' (designed by Briggs and MacKay, and published in MacKay, 1995).

Results

The results for reading accuracy and reading comprehension are shown in Table 12-1. The mean pre-test scores for this follow-up sample at the time of the original study were the same both for accuracy (6y 0m) and for comprehension (6y 4m). After the passage of approximately five and a half years, during which time the children in the sample had received no further specific interventions either to enhance reading ability or to maintain attitude change, the experimentals were reading 1y 8m ahead of the controls for accuracy ($t = 1.873$, $p < 0.05$, one-tailed test), and 1y 9m ahead for comprehension. ($t = 1.752$, $p < 0.05$, one-tailed test). (independent two-sample t tests using the Microsoft Excel Version 8.0 data analysis tools.)

Table 12-1 Attitude change and reading ability: follow up (N=19)

	Experimentals (N=11)	Controls (N=8)	
Reading accuracy	10y 11m (Raw score 71)	9y 3m (Raw score 55)	p < 0.05 SD = 19.4* Effect size = 0.82
Reading comprehension	12y 3m (Raw score 25)	10y 6m (Raw score 31)	SD = 8.6* Effect size = 0.70

* Very few standard deviations are cited in the Neale Analysis standardisation, and those cited closest to this age group are for age 11.0-11.11. For both accuracy and comprehension the SD varies little across the age groups.

At the time of follow up, eight out of the 11 experimentals had a reading age above a 'functional literacy' level of nine and a half years, and three of these had a reading age equivalent to their chronological age. By comparison, only two out of the eight controls had a reading age above functional literacy level. One of these was reading at chronological age level, and had therefore made better progress than most of the experimentals. This pupil served as a good example of the unexpected effects that may occur through being part of a control group. He was asked if he knew how he had made such good progress as he had a significant reading difficulty when he was tested for the original study in 1993, but had not received any extra help. He said it was because he had decided to do a lot of reading at home. When asked, 'When did you decide to do that?', he replied: 'On the day I was tested in 1993!'. He realised at that time that he had a reading problem, and was determined he would overcome it.

On the attitudes questionnaire, no significant differences were found between experimentals and controls. This applied both to the total scores for the 14 relevant questions relating to attitudes, and also to the score for the three statements that related specifically to attitudes towards reading: 'It's a good thing to have lots of books at home', 'I'm not very good at reading' and 'I do not like reading very much'.

Conclusions

This study has indicated that, for pupils who had a significant reading difficulty in their mid-primary school years and were subject to an intervention designed to change their attitudes and values concerning reading, the gains they made after a short 10-week programme were still apparent in levels of reading achievement over five years later. In contrast with the controls, most of whom had not achieved functional literacy levels, most of the experimentals had achieved this level, with some reaching the level appropriate to their chronological age.

Although no significant differences were apparent on the questionnaire exploring attitudes to school and to reading, this study highlights the importance of addressing issues of attitudes and values as part of any intervention designed to tackle reading failure and to raise literacy levels. Cox (2000), in a review of pupils' perspectives on their education, has concluded that:

'It is... important for children to be helped to develop positive attitudes to themselves as learners, ie in terms of their self-concepts, self-esteem, and their levels of aspiration in learning, so that they can approach the challenges of learning with confidence' (pp. 136-137).

This attitudes study provided further support for recognising the importance of attitudes to learning and of the promotion of positive self-concept. These features were recognised throughout the main study. They permeated all aspects of the intervention, and were constantly reinforced through the key context variables of vision, profile, commitment, ownership and declaration. The attitudes study also provided a basis from which the following study was developed – the declaration study as reported in Chapters 13 and 14.

SUMMARY

This chapter describes a five and a half year follow up to a randomised control trial described in the preparatory studies. Its purpose was to support one of the strands in the intervention, 'enhancing attitudes, values and expectations'. Twenty of the original primary school sample of 24 pupils were traced at the mid-point of their secondary school career. Of these, assessments were carried out on 19 using both the accuracy and comprehension tests on the Neale Analysis of Reading Ability. The results indicated that the experimentals, who had received a programme aimed at changing attitudes and values towards reading, were still scoring significantly above controls on both tests, but no significant differences were found on an attitudes questionnaire. The literature relating reading achievement to factors such as attitudes and self-esteem is considered. This factor is highlighted as an important component in literacy intervention.

Chapter 13

The Declaration Study: Rationale and Design

Introduction

This study served to develop further strand 10 of the main study, ‘enhancing attitudes, values and expectations’. The study was carried out in 12 primary and nursery schools (six experimental, six control) in East Renfrewshire during Session 1999-2000. The aim was to change children’s expectations regarding their achievement in literacy, and to assess the impact of such change on actual reading scores. A total of 565 pupils participated in the six experimental establishments – 320 at pre-school level and 245 in Primary 1, with 27 teachers plus school management involved in implementation. Schools were matched for socio-economic status and included establishments with high and low levels of disadvantage.

Staff were trained in a novel intervention strategy based on changing expectations through declarations by pupils regarding future achievement, and this was implemented daily throughout a period of approximately nine weeks. A systematic sample of 60 children was assessed individually before and after the intervention, using the baseline assessment designed for the main study. Pre-post measures of attitudes to reading were also obtained. These quantitative measures were supported by qualitative indicators obtained from children who were assessed individually and also from staff in relation to all of the participating pupils in the experimental establishments.

This study, together with the attitudes study (Chapter 12) strengthened the base for ensuring that addressing attitudes, values and expectations should be built into any literacy intervention programme as a variable which could affect outcomes but which was essentially separate from the literacy content of the programme itself. Both studies addressed this variable in its own right, while keeping the content of the reading curriculum constant. The declaration study resulted not only in significantly higher scores on early literacy skills for the experimentals but also in significant shifts towards more positive attitudes and expectations regarding reading. As a study of children in their pre-school year and Primary 1 it was particularly relevant to informing a large-scale early intervention.

Aims of study

The aims of this study were:

- to change children’s expectations regarding their future levels of reading achievement
- to enhance their attitudes towards reading
- to raise their levels of reading achievement through changing their attitudes and expectations.

Hypotheses

The following hypotheses were proposed in relation to the pupils in the declaration sample (the experimentals):

- 1 that after the intervention they would have higher scores in tests of early reading skills
- 2 that they would show more positive attitudes to reading.

Rationale

The idea that children's reading levels would be enhanced by getting them to do nothing other than make bold declarations about their future levels of reading achievement may seem not only novel but improbable. This study, nevertheless, involved nothing other than this. It included no curricular innovations and no alternative teaching methodologies. It required nothing other than that children – whether individually, in groups or in whole classes – should declare what good readers they were going to become in the future.

It was intended that this study should be highly innovative and for this reason it does not draw from an existing body of literature specific to this field. Nevertheless, any idea of such a nature would need to be discussed in the light of a clear rationale that would at least illuminate the reasons for devising a strategy of this kind. It would preferably also be embedded in a known body of psychological theory that might provide its practice with a reasonable foundation. The rationale underlying the study, and the body of psychological theory from which it drew, may be stated simply. First, a number of reports relevant to this subject from areas of psychological and other literature that had not been subject to experimental investigation or indeed any level of formal scientific enquiry had been noted. Second, a wide range of topics for which there is an existing base in mainstream psychological theory and practice seemed to provide support for believing in the potential of declarations. It was on this basis that it was decided to make an approach to 'declaration theory' – that is, a schema about declarations that might be supported both by wider psychological thinking and by practice.

Declaration: the informal literature

Within the general arena of the informal literature on psychology and other fields of human interest, some writers have made explicit reference to the idea of declaration. These, however, have tended not to be at levels that are, or perhaps even could be, subject to formal evaluation. Joseph Murray, in a best-seller of the 1960s, *The Power of Your Subconscious Mind*, began by observing that 'your habitual thinking and imagery mould, fashion and create your destiny' (Murray, 1963, p. 5). His work was not of an experimental kind, but he did outline examples of personal declaration in ways that pointed to this area as a potentially fruitful line of enquiry. He used the concepts of 'affirmation' and 'decreeing' – essentially a form of making bold declarations of positive outcomes.

Outwith psychology, Tracy Goss made reference explicitly to declaration in her work in transforming businesses and multi-national corporations (Goss, 1996). She defined a declaration as ‘an act of speaking that brings forth a future the moment it is spoken’ (p. 116). The concept expressed by Murray and Goss was not fanciful or mystical, but was based informally on a psychological theory that declaration would shape the actions required to achieve the stated goal. ‘This declaration lays the groundwork for action in a new realm...Said another way: a declaration of possibility bring “what is not” into existence as a possibility’ (Goss, 1996, p. 117).

While the above examples may be defined as unscientific in that they have not been subject to an experimental approach involving experimental and control groups, they have nevertheless been based on enquiry which has been systematic and which has built on cumulative successful experience.

Declaration: pointers from mainstream psychology

The ideas leading to this study, and the methodology devised for it, were consistent with many established areas of mainstream psychological theory and practice. Nine such areas, most of them closely inter-related, are highlighted here:

- attitudes, self-concept and self-esteem
- expectations or ‘expectancy’
- cognitive dissonance
- social and interactive learning
- motivation
- attributions
- goal setting
- self-efficacy
- visual imagery.

Each area has its own extensive literature, and it would be altogether outwith the scope of this work to seek to review it. However, it may be of relevance to provide a few key pointers for each area and to relate these to the concept of declaration.

Attitudes, self-concept and self-esteem

The area of attitudes, self-concept and self-esteem, together with the following discussion of expectations, was central to the declaration study, since it was an investigation specifically designed with reference to the broad concept of ‘enhancing attitudes, values and expectations’. The literature in this area in relation to reading and to educational attainment has been discussed in Chapter 12 with particular reference to socio-economic status. It has been summarised by noting that in disadvantaged populations negative attitudes to school and education correspond with lower levels of academic self-concept and self-esteem (Cox, 1983). Self-esteem has been described as possibly having ‘a directive, pivotal role in learning’, with poor self-concept portrayed as ‘a primary cause of academic underachievement’

(Kershner, 1990). It is in this context that the value of interventions aimed at changing attitudes has been demonstrated (MacKay, 1995a, 1999b).

With reference to declarations, it has also been noted from the study by Cox (1983) and the review by Gurney (1987) that among factors considered to be important were the value of making positive statements about oneself. Indeed, 'positive thinking' in general has been shown to have beneficial effects on well-being. Burton and King (2004) studied the health benefits of writing about positive experiences. They recruited 90 young adults to participate in a writing exercise for 20 minutes a day over three days. Half of the participants were to write about intensely positive experiences they had had, while the other half were to write about mundane things such as their plans for the day, their shoes or their bedroom contents. For three months prior to the exercise both groups were indistinguishable in the number of visits they made to doctors, while in the following three months the positive writing group made significantly fewer visits. The authors' view was that the positive approach was a way of 'obtaining self-understanding, of gaining a more clearly articulated sense of self and of discovering and creating one's life goals'.

Positive thinking and the making of positive statements are at the heart of 'declaration theory' – it is based on positive assertions about oneself and about future levels of achievement. The expectation is that making these positive statements will mould attitudes, and children will have a concept of themselves as future good readers.

Expectations or 'expectancy'

The study of expectations is equally pivotal to the issue of asking children to make declarations. Indeed, this particular piece of research was contracted to East Renfrewshire Council under the broad heading of studying 'expectations'. It was felt that whatever else was happening with declarations, it was certainly changing expectations both on the part of children and on the part of teachers. Both are of central importance.

The broad field of 'expectations' has occupied an important position within social psychology over a long period of time. Within education the subject became of central interest with the work of Rosenthal and his colleagues. Rosenthal and Fode (1963) found that students who were given groups of 'bright' and 'dull' rats produced results consistent with the label of their rats, even though the rats were in fact randomly mixed on the basis of maze learning abilities. This study was used to show that experimenter expectancy can affect the behaviour even of rats. It was, however, in the human arena that the Rosenthal experiments attained a degree of notoriety with the publication of *Pygmalion in the Classroom* (Rosenthal & Jacobsen, 1968). It was reported that randomly selected children whose teachers heard that they were expected to make late gains in their academic development actually did make more gains than controls. While these results have engendered considerable controversy, the meta-analysis by Rosnow and Rosenthal (1997) of hundreds of expectancy studies in diverse areas points overall to significant expectancy effects.

Experimenter expectancy studies have frequently been conducted to provide a commentary on research methodology by highlighting contaminating variables. However, in the case of the effects teachers' expectations have on children's learning they have also provided a valuable basis for promoting high expectations (Saracho, 1991). A similar comment may be made in relation to participant expectancy. This may be demonstrated by the 'Hawthorne effect', a topic so widely cited that it has become a standard part of the vocabulary of psychology and requires little expansion. The occupational psychology studies carried out from around 1927 to 1932 at the Western Electric Hawthorne Works in Chicago found, after varying factors such as level of illumination and timing of rest breaks, that productivity increased *whatever* the change – including returning to the conditions operating at the outset (Mayo, 1933; Roethlisberger & Dickson, 1939). The conclusion was that the effect was simply the result of being the focus of interest and attention. The effect may reflect changed levels of expectancy, or indeed raised motivation. In any event, it clearly belongs under one of the headings being considered in this overall discussion. While the Hawthorne effect provides a caveat to researchers regarding contaminating variables, it may also be seen as a potentially valuable tool in raising expectations.

High expectations are one of the important and consistent factors in successful schools (Sammons, Hillman & Mortimore, 1995). The very nature of a declaration about academic success is that it incorporates an expectation of fulfilment. To say, for example, *I will become a good reader*, is to make a statement about something the person who utters it expects to happen. To have a teacher promoting this statement is an endorsement that the teacher's expectation is in line with that of the child. Thus, if children make bold declarations regarding high levels of achievement, and if these are reinforced by teachers, the whole context of high expectations is likely to be supported.

Cognitive dissonance

Festinger's (1957) theory of cognitive dissonance is so thoroughly established as a landmark in mainstream psychology that it requires little comment or explanatory background. Essentially, there is a tendency for individuals to seek consistency among their cognitions and behaviours. Where there is inconsistency or dissonance, something must change to eliminate it. This is likely to be done in one of three ways: first, by reducing the importance of the dissonant beliefs, second, by acquiring new consonant beliefs that change the balance, or third, by changing the conflicting attitude or behaviour.

In relation to declarations, if children are boldly declaring – and most frequently aloud in the company of others – that they will become good readers, together with other and perhaps quite detailed statements of a similar vein, then they will be likely to experience dissonance if they do not engage in behaviours that are consonant with such declarations. Therefore, it may be expected that their behaviour will adapt to achieve consonance, and they will engage more with those attitudes, interests, tasks and activities that are associated with becoming a good reader.

Social and interactive learning

Social and interactive learning has been described in Chapter 8 as one of the 10 strands of the overall intervention. The evidence base covering this area is extensive and includes groupwork, peer-assisted or parent-assisted learning and collaboration of students in interaction with computers (Foot et al., 1994). In summary, it is recognised that effective, motivated learning is likely to be enhanced by a social and interactive setting which is meaningful to the child and which involves both peer collaboration and adult support (Hughes & Greenhough, 1994). Social and interactive learning has also been identified in Chapter 10 as one of the features of the synthetic phonics programme, with children learning letter sounds and actions at the same pace as part of a whole class activity.

The activities involved in making declarations did not specifically require that learning should be of a social or interactive nature. The whole process was capable of being conducted at individual level. However, as might have been expected of a classroom programme that could be carried out with individuals, groups or whole classes, the main method used was the whole class approach. That is, the declarations were generally made in an interactive social context. It was for this reason that press reports of the intervention included frequent description of the programme as ‘chanting’ (see Appendix 6), and radio and television coverage tended to centre on entire classes of children enthusiastically making their declarations in unison. For the individual child, this meant being part of a novel situation which involved both social support and peer accountability, and in which everyone was acting together in a common purpose.

Motivation

Motivation and its relation to learning and achievement is so firmly established in psychology as to be an integral part of its basic subject matter and of any comprehensive learning theory. It has long been recognised as being central to children’s learning. In cognitive theory, it serves to create intentions and goal-seeking acts (Ames & Ames, 1989). The concept of ‘achievement motivation’ has become a well-developed area of research in its own right and is highly relevant to learning (Atkinson & Raynor, 1974; McClelland, 1985; McClelland, Atkinson, Clark & Lowell, 1953; Weiner, 1990). Motivation to achieve is a function of the individual’s desire for success, the expectancy of success and the incentives provided.

The potential relevance of declarations to motivation is clear. For example, children who perceive themselves as having a low level of ability, and who believe there is little or nothing they can do to change the situation, are likely to experience motivational problems (Dweck, 1991). Declarations of future success are designed to counter these perceptions and to reinforce a self-perception based on ability. Keller’s (1983) model of motivation is also of interest in the context of declarations. It encompasses four components: arousing interest, creating relevance, developing an expectancy of success and producing satisfaction through intrinsic and extrinsic rewards. If children several times each day are focusing their attention on statements

about reading, then it is probable that their interest in reading is likely to be aroused as something relevant to their experience and something that may be rewarded with success.

Attributions

Attribution theory is concerned with how people interpret events and how this relates to their thinking and behaviour. The focus is on the causes to which individuals attribute events or behaviours. A psychological theory of attribution was first proposed by Heider (1958), but it was mainly Weiner and his colleagues who were responsible for developing it as a major social psychology paradigm (Jones et al., 1972) with a principal focus on achievement and motivation (Weiner, 1974, 1986). He identified four general concepts to which one can attribute achievement – skill, or ability, effort, task difficulty and luck. Attributions are classified along three causal dimensions – locus of control, stability and controllability. Locus of control may be internal to oneself or external. Stability refers to whether causes are seen as broadly stable, such as ability or task difficulty, or broadly unstable, such as effort or luck. Controllability indicates whether the causes are within one's own control or, like luck, uncontrollable.

Inherent in the concept of declarations is a confirmation of internal locus of control. A declaration by its nature is likely to begin with statements such as *I will*. The focus is removed from external, uncontrollable factors to definite confirmation of individual action. The work of Dweck and her colleagues on attribution theory is of relevance here. Some children make 'entity attributions' about self-characteristics, while others make 'incremental attributions' (Dweck, Hong & Chiu, 1993). To make an entity attribution is to believe that the talent or skill is a fixed quantity that one possesses or lacks to a greater or lesser extent. To make an incremental attribution is to believe that the skill can be improved with effort, practice or support. Children who hold entity theories have been shown to make more global statements about themselves in response to negative feedback. For example, they might respond to criticism with statements like *I'm just not smart enough* (Dweck et al., 1993). Children with these perceptions of themselves will be less likely to respond to the changes implied by constructive criticism. Since the mental model they hold is based on the possession of a fixed amount of talent, 'criticism indicates that this amount is, immutably, less than what they desire it to be' (Shaughnessy & Self, 1997). Declarations are designed to promote an incremental model of attributions. They reject ideas of the current position being a fixed one and encourage a commitment to being able to become successful in the talent in question.

Goal setting

The importance of goal setting in relation to achievement has for a long time been understood in relation to many groups including school pupils (Gaa, 1973) and college students (Morgan, 1987). In summary, the identification of academic goals is seen to be essential to the successful management of learning (Bandura, 1986; Locke & Latham, 1990) and those who set effective goals tend to achieve at higher levels than others (Locke & Latham, 1990; Zimmerman, 1989). Goals mobilise effort, lead

to task-appropriate study strategies and increase persistence and commitment (Bandura, 1986; Locke & Latham, 1990; Schunk, 1996).

One of the expectations in getting children to make declarations in this study was that it might influence their goal-setting behaviour. It is reasonable to anticipate that people who make daily statements that relate to a future achievement or goal – in this case becoming good readers – may identify with that goal and engage in attitudes and behaviours that are compatible with it. In addition, because the making of declarations involves not only broad, general statements (*I will become a good reader*) but also specific goals (*I will be able to read all the words in my tin*), there is a definite context in which task-appropriate study strategies are being encouraged.

Self-efficacy

‘Perceived self-efficacy is defined as people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives. Self-efficacy beliefs determine how people feel, think, motivate themselves and behave’ (Bandura, 1994, p. 71). The work of Bandura and others in this area has made a significant contribution both to mainstream psychological theory and to professional practice across a range of fields (Bandura, 1986; Maddux & Stanley, 1986). For example, self-efficacy is an important underlying concept both in cognitive-behaviour therapy (for example, Hawton, Salkovskis, Kirk & Clark, 1989) and in solution-focused approaches in which it is assumed that the client wants to change, has the capacity to envisage change and is taking steps to make change happen (for example, de Shazer, 1985). In relation to children, schools and education, self-efficacy is crucial not only for the children themselves but also for their teachers. Children’s belief in their capabilities to master educational activities affects their aspirations, their level of interest and their attainments, and teachers’ own views of their self-efficacy and how to encourage it in others is of the highest importance (Bandura, 1994).

One of the aims of having children make bold declarations about their future reading achievement, and having teachers participate in promoting these declarations, is to develop beliefs about capabilities – to lead children to a position in which they say what they believe and believe what they say. Thus their feelings, thoughts and behaviours are likely to be shaped to support the messages they are constantly asserting when they make their declarations each day.

Visual imagery

Making declarations is also designed with the expectation that those who are making regular statements about becoming good readers may be more likely to form mental images of themselves engaged in the reading process. Visual imagery, or visualisation, occurs when people imagine themselves performing an action in the absence of the actual practice of it. More broadly, it is also called mental imagery, mental rehearsal or mental practice, since while most discussion of imagery focuses on the visual mode, another sensory mode such as the auditory may at times be more central. Several theoretical models of visual or mental imagery have been proposed

(Murphy, 1990), but whatever the psychological mechanisms involved it has frequently been shown to enhance performance.

The research base for this area is growing rapidly in the field of sport and exercise psychology, a branch of the discipline with a sufficient evidence and practice base that in 2004 the British Psychological Society formed a Division of Sport and Exercise Psychology. For example, Isaac (1992) tested 78 subjects and classified them according to whether they were novel or experienced trampolinists and also according to whether they were high or low imagers. An experimental and a control group were trained in a skill. The experimentals followed their training with mental practice of the skill, while the controls followed it with a mental task of an abstract nature, such as maths problems. High imagers in both groups showed more improvement than low imagers, but the experimentals also showed more improvement than the controls, pointing to the benefits of visual imagery, particularly when focused on practice of the skills in question.

Roure et al. (1998) linked mental imagery with specific autonomic nervous system responses, suggesting that it may help in the construction of schema that can be reproduced, without thinking, in actual practice. It has also been shown that those who engage in imagery set higher achievement goals and have higher commitment to practising the requisite skills (Martin & Hall, 1995). While most studies in this area have been with adults, the benefits of imagery have also been demonstrated with children (Orlick, Zitzelsberger, Li-Wei & Qi-Wei, 1992).

In summary therefore of these areas of mainstream psychological theory and research, it may be said that while the idea of declarations to enhance children's literacy is novel, it is not detached from an established evidence base across several areas of psychology. Working from this theoretical and practice base it was envisaged that the project would:

- define and enhance children's expectations about reading
- raise their confidence that they would be good readers
- affirm their enjoyment of books and the reading process
- modify their beliefs about what they could achieve
- shape their internal images of themselves as successful readers
- motivate them in acquiring literacy skills
- emphasise ownership and self-efficacy in relation to their achievements
- set in motion the process of achieving declared goals.

Method

Design

This was a quasi-experimental study using a range of quantitative and qualitative measures, with pre-post testing taking place prior to and following a nine-week intervention period.

Sample

There were two aspects to the sampling process in this study: first, the selection of educational establishments and second, the selection of pupils for individual assessment.

The sample of establishments comprised six experimental and six control schools. The four primaries and two nurseries in the experimental group were matched as closely as possible with the schools in the control group. The three criteria used for matching were socio-economic status, achievement levels as indicated by group reading test results and confirmation by the education authority of comparability in terms of overall variables and perceptions. The experimentals and controls were also balanced for denominational and non-denominational schools.

Matching by achievement levels and by utilising the knowledge of schools possessed by the educational directorate has clear face validity. The significance of socio-economic status in accounting for a high amount of the variance in educational outcomes is already well established and is covered in the literature review relating to socio-economic disadvantage (Chapter 4). Currently unpublished research data gathered by the author also point to the likelihood of systematic differences in literacy levels in the early stages between denominational and non-denominational schools (MacKay, in preparation b).

In terms of free school meals (FSM), footwear and clothing grants (FCG) and standard achievement scores on the NFER-Nelson GRT-II (SAS), the final matching produced two primaries per group in the high socio-economic category (averages: FSM 5%, FCG 9%, SAS 109) and two per group in the low category (averages: FSM 31%, FCG 50%, SAS 99). Out of 24 primary schools in the area covered by the authority, four of the five with the highest FSM and FCG were included in the low socio-economic sample. The school with the lowest socio-economic rank in the authority (FSM 69%, FCG 84%) was not included as it did not have a sufficiently close match among the available primaries. Similarly, the four nurseries represented one high and one low socio-economic establishment in the experimental and control groups, using the criteria of post-codes and with the suitability of matching confirmed by the educational directorate.

For the individual assessments, a small systematic sample of five pupils was selected from each establishment. This generated a sample of 60 children for pre-test and post-test. The close matching of experimental and control establishments ensured

that the children who were assessed individually in the two groups were likely to have had comparable social backgrounds and educational experiences.

The sampling method involved selecting each n th child from the register, where n equalled one-fifth of the number in the primary class or nursery pre-school year group. Therefore, for a group of 20 it was every 4th pupil, while for a group of 60 it was every 12th. Only two qualifications were applied to the selection process. First, to ensure a gender balance, the next boy or girl as appropriate after the n th child was selected in cases where it would otherwise have been almost all boys or girls. Second, children were excluded either if it was known that they were likely to move to a different school during the project or if they had been identified as having such pronounced additional support needs as to render the assessment measures inappropriate for them. Thus, for example, a child would be excluded from the sample on the basis of a severe speech and language disorder or sensory impairment, and the next child on the register selected.

While this selection method did not preclude the possibility of random differences in achievement levels between experimentals and controls in the small sample selected, it did ensure that the sample was unbiased and was not pre-selected by staff on the basis of perceived ability or any other factor. As to random differences in the sample between one establishment and another, this was not an issue as the experimentals and controls were being compared on the basis of their own individual pre-test and post-test scores.

Consultation, training and support

This intervention was supported by a programme of staff consultation and training, by information to all parents in the appropriate classes and nursery groups in both the experimental and the control schools and by ongoing monitoring and support for schools as required throughout the project.

Staff consultation and training was in three stages. Early in Session 1999-2000 a meeting was held for all head teachers in the experimental establishments, with a member of the educational directorate and a representative of psychological services present. The purpose of the meeting was to provide information to school management on the background to the project and to outline the research proposal and its rationale. The meeting also served to orientate the researcher with regard to the needs and expectations of the various schools, and to provide a basis for detailed planning of the intervention.

Following this a meeting was held for all of the staff who were likely to be directly involved in implementation of the project, together with representatives of school management. This took the form of a more detailed presentation of the project aims and rationale, and at the same time served to commence the process of raising staff expectations regarding the successful outcomes that might be anticipated for the children who would take part.

Finally, separate meetings were held in each of the experimental schools for all staff who would be involved, along with school management. These meetings provided the detailed training in carrying out the intervention, and were supported by an information pack that set out both the theoretical and practical aspects of the project, together with weekly record sheets. School management were also encouraged to ensure that staff throughout their establishments were informed of the project so that it would have appropriate profile and be seen as integral to whole school policy in raising achievement.

Information to parents took the form of a letter sent from the educational directorate noting that the research project in the experimental and control schools formed part of the authority's early intervention programme for raising educational achievement. It was noted that schools would be visited for the purposes of the research and also that a number of children picked at random from the register would have their progress assessed before and after the intervention. The psychological service was also kept informed of the project and of the schools in which intervention or assessment would be carried out.

It was important to ensure that both the detailed implementation of the strategies and the enthusiasm of staff were maintained, and throughout the project monitoring and support were available from the researcher whenever required. This included a 'help line' to phone for advice. In practice, because the intervention was designed to be simple to carry out and also enjoyable, staff were confident in implementing it satisfactorily at all times.

Implementation

Although the underlying rationale for this study stemmed from the range of complex theoretical considerations outlined above, implementation of the strategies was based on the premise that effective educational interventions should be simple, easily understood, sustainable and enjoyable for both staff and pupils. All of these elements were therefore built into the specification of the strategies adopted. It was essential that it should not be viewed as an 'extra chore' for teachers, and it had to be seen to support the school's existing aims and curriculum.

The project was based on children making declarations about literacy – particularly about their future levels of reading achievement and their enjoyment of books and reading. The basic requirement was that each child must make a minimum of three positive declarations about future reading achievement every day throughout the project. This could be in an individual, group or class situation. In the individual situation, each child would make a personal declaration, which might be specific to the particular needs or aspirations of that child. With group declarations, all of the children in a group would make a declaration together. For class declarations, the entire class would recite the declaration in unison.

Each of these situations had its own strengths to offer. Individual declarations could be 'customised' to the child, and would therefore involve a personal and unique aspiration or commitment. Group declarations involved social variables of

collaboration and group identity. They also allowed competition among groups to be encouraged – for example, finding which group could shout their declaration loudest. It was the third situation, class declarations, that was most likely to be used by teachers routinely as it was probably both easiest and most enjoyable. This type of ‘class chant’ also offered potentially powerful effects on learning, by promoting social learning, reinforcing class goals and helping the full inclusion of children with difficulties by maximising identification with the whole class and its curriculum.

Prior to the intervention teachers were provided with a range of sample declarations. These varied from the general, *I will become a very good reader*, to the specific, *I will learn to write my name*, *I will be able to say all my letter sounds*, *I will know all the words in my tin*. In addition to declarations regarding future achievement, other examples covered enjoyment of reading, such as, *I like books – books are fun*, *I want to take a book home to read*. Teachers were also encouraged to reinforce these declarations by making positive affirmations such as, *You are all going to become very good readers*.

To support the programme and increase children’s enjoyment in participating, examples of ‘props’ were outlined and demonstrated. In particular, the use of a puppet was recommended, with the proposal that it should become the focus of all work on declarations. The puppet could be used when asking the class or group questions: ‘Crazy Crow says, *I’m going to become a good reader*. Who’s going to become a good reader? *We are*’. Similarly, the puppet could be used to support whatever specific literacy skill was being taught at any given time: *Crazy Crow can say the alphabet*. *Can you say the alphabet?*

The puppet could also be used as a prop for children to tell it what they knew, what they found difficult, what new things they had learnt that day or what they expected to be able to do in the future. *How will you know when you are a good reader? Tell Crazy Crow all the different ways you can of being a good reader*. As well as motivating children, props like this provided an alternative and at times easier method of communication, and also could facilitate interactive learning in peer learning situations.

It was felt that by providing staff with every detail they required to meet the programme’s specification, including ready-made declarations, it maximised the likelihood that the implementation would be carried out by everyone, including any staff who might at times feel the pressures of time or the demands of creativity. At the same time, staff were strongly encouraged to use the examples as a basis for generating their own ideas and also helping the children to propose new ideas and declarations. For example, it was proposed that they might devise ditties or doggerel for class chants, make up declaration games or generate new things to say to the puppet about reading. These proposals were designed with a view to building enthusiasm and fostering a sense of ownership of the project by those participating in it.

To monitor and support high fidelity of programme implementation, all participating staff in the experimental schools completed a weekly record sheet to note whether or

not they had met the basic project specification each day, and to note any additional declarations or activities that had been introduced. These records indicated that implementation of the intervention as specified was carried out to a very high degree. This applied to every establishment, and in very many cases the basic specification for each day of the intervention was supplemented by additional activities. Staff noted only a very few occasions when the daily requirement was not met for very straightforward reasons – ‘on a course, supply teacher in’, or, ‘forgot on first day of week’.

Measures used

A range of quantitative and qualitative measures was used. The main quantitative data were obtained from individual assessment of early literacy skills, and these assessments were supported by data from children on their attitudes to reading, teacher ratings, semi-structured teacher interviews and direct feedback from children. Information was also obtained from weekly record sheets completed by teachers.

The baseline assessment (Chapter 7) was used for the individual assessments. In addition to its sensitivity to change and other characteristics as previously outlined, it was felt to be particularly suitable to the structured approach taken by East Renfrewshire to early assessment. All of the children selected by systematic sampling in the experimental and control establishments were individually tested on the baseline assessment before and after the intervention.

An attempt was made to obtain direct information from the pupils on their attitudes to reading using a strategy designed by the author for this study. Each child in the systematic sample was asked five questions constructed to explore the following areas before and after the intervention: 1 how well they thought they had done on the baseline assessment; 2 whether or not they liked reading; 3 whether they found reading easy or difficult; 4 whether they liked having stories read to them; 5 whether they thought they would be good readers the following year.

Because of a number of established difficulties in assessing the views of young children in this way, such as always responding positively to please the examiner, the children had three glass jars set before them. Each jar contained a large and similar number of small white cards, designed to represent the responses made by other children to the same questions. Each question was presented in the following manner: *These children don't like reading very much* (pointing to jar 1), *these think it's OK* (jar 2) and *these children like reading a lot* (jar 3). *What do you think about reading?* The child was then invited to post a card into the chosen jar, and the choice was recorded. The direction of the questions was varied to avoid response set (for example, always choosing the first jar, or the last one). Posting the cards in this way made the exercise appear more anonymous, and since children have high levels of group conformity (see, for example, MacKay & Watson, 1996) it was clear from the spread of cards among the jars that the answers could safely be deposited wherever the children liked.

While the baseline assessments were used as the objective measure of literacy skills for this study, additional information was gathered from teachers before and after the intervention for the children in the systematic sample. This invited their opinions on various aspects of the children's literacy skills and interests. A seven-point rating scale was used for five simple questions (Table 13-1). This offered the opportunity to note teachers' views of the children in the light of the more formal assessments and also to consider any apparent changes in ratings pre-test and post-test.

At the end of the intervention all staff who participated were interviewed regarding their experience of the project and of their pupils' involvement in it. The interview included: the extent to which staff had carried out the project specification; an account of the specific ways in which they conducted the implementation; notes on spontaneous comments from children or parents; the extent to which they felt that pupils had enjoyed and had benefited from the intervention; and the effects of the project on their own expectations and the confidence, expectations, motivation and interest of the children.

Table 13-1 Teachers' Ratings

What is your present opinion of this pupil's ratings in the following areas of literacy?

	Does not apply at all				Applies very much		
1 This pupil has made very good progress in early literacy skills	1	2	3	4	5	6	7
2 This pupil shows promise of becoming a very good reader	1	2	3	4	5	6	7
3 This pupil is very confident in literacy at the appropriate level	1	2	3	4	5	6	7
4 This pupil shows a high degree of interest in reading	1	2	3	4	5	6	7
5 This pupil takes a great deal of pleasure in literacy-related activities	1	2	3	4	5	6	7

To supplement the qualitative information obtained, all of the experimental children in the selected sample were asked for their views on whether they had enjoyed the project and whether they remembered the declarations they were making. Further information of this kind was obtained on visits to whole classes at Primary 1 level and groups in nursery.

SUMMARY

This chapter introduces a novel intervention aimed at developing further the intervention strand of enhancing attitudes, values and expectations. The sample was 565 pupils, of whom 60 were targeted for individual assessment, in 12 nursery and primary schools (six experimental, six control) in a different Council area. The schools were selected on the basis of high or low levels of socio-economic disadvantage. Staff were trained in a novel intervention strategy based on changing expectations through declarations by pupils regarding future achievement in literacy. The declarations were made for a minimum of three times daily through a nine-week period. Although the concept was novel, aspects supporting it were noted in the informal literature, and a rationale for its use was proposed in reference to nine areas of mainstream psychological theory and practice: attitudes, self-concept and self-

esteem; expectations or 'expectancy'; cognitive dissonance; social and interactive learning; motivation; attributions; goal-setting; self-efficacy; and visual imagery. Pre-post assessments of the targeted sample used the baseline assessment scheme developed for the main study, together with a method for assessing attitudes in young children developed for this subsidiary study. A range of qualitative measures was also used in relation to the whole sample.

Chapter 14

The Declaration Study: Results

Baseline assessment

Of the 60 children selected for the systematic sample, 54 were available for both pre-test and post-test. The remaining six could not be assessed at post-test because of either absence or having left the school. These divided equally between experimentals and controls, giving a final sample of 27 experimentals and 27 controls, with 9 nursery and 18 primary in each group.

For the nursery group, the experimentals showed significantly greater gains than the controls on three tests: nursery rhymes ($t = 3.05$, $p < 0.01$, effect size 0.51), rhyme production ($t = 3.78$, $p < 0.01$, effect size 0.98) and lower case letter sounds ($t = 2.264$, $p < 0.05$, effect size 1.51). The tests used were independent two-sample t tests, one-tailed, calculated using the data analysis tools on Microsoft Excel Version 8.0. The results are shown in Figure 14-1. For the primary group, the experimentals showed significantly greater gains than the controls on five tests: nursery rhymes ($t = 1.88$, $p < 0.001$, effect size 0.44), lower case letter sounds ($t = 3.07$, $p < 0.01$, effect size 0.21), letter names ($t = 2.23$, $p < 0.05$, effect size 0.64), non-word reading ($t = 2.57$, $p < 0.01$, effect size 0.60) and word reading ($t = 2.61$, $p < 0.01$, effect size 0.64). The results are shown in Figure 14-2.

These results followed a similar pattern to the results of baseline assessment tests in the main study, where the tests most applicable to the pre-school group were those that assessed skills typically taught in nursery, such as concepts of print and phonological awareness. The older children in Primary 1 were often at or near the ceiling for these tests, but were able to tackle the more formal tests of early reading skills including word reading.

Figure 14-1 Nursery group: tests showing significant gains (gain scores)

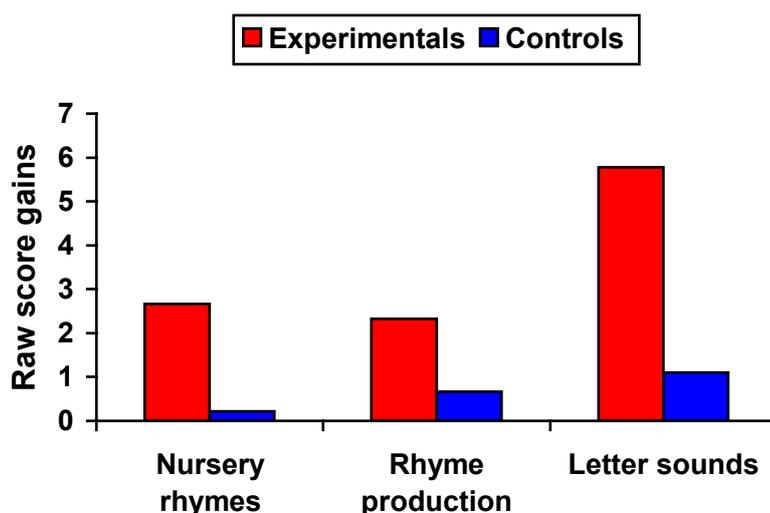
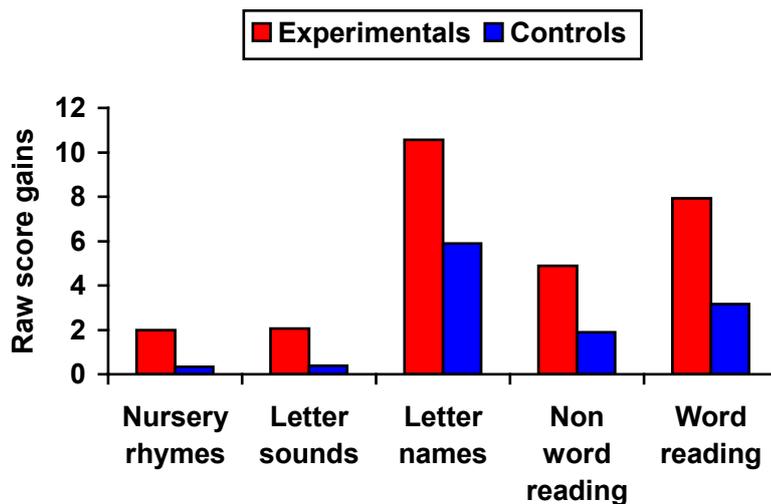
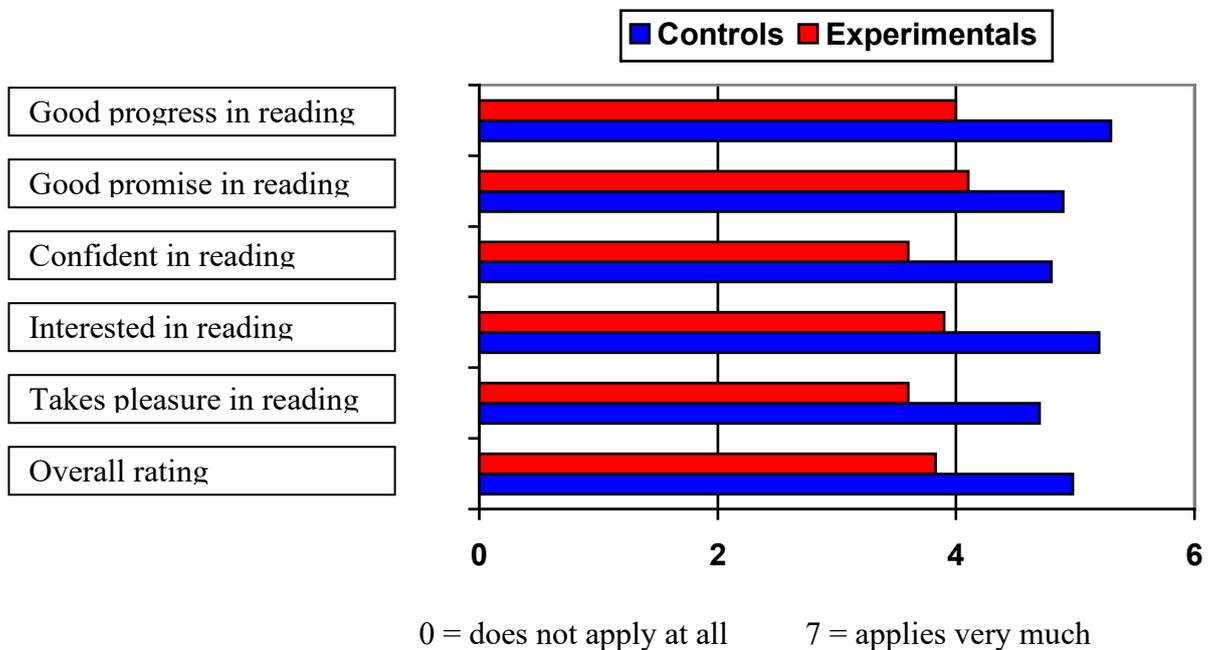


Figure 14-2 Primary group: tests showing significant gains (gain scores)



Because of random variation in the small systematic sample selected for assessment, the pre-test scores of the controls were higher than those of the experimentals on a number of the tests, particularly for the primary group. That is, a larger number of more able children were found among the controls. This observation was supported by the ratings made by class teachers of the children’s literacy abilities at pre-test stage. These ratings were higher for the controls than for the experimentals at primary stage, as shown in Figure 14-3 ($t = 2.42, p < 0.05$, two-tailed test).

Figure 14-3 Primary group: teacher ratings (pre-test)



The data were analysed to ensure that it was not easier for children who started with lower scores to make greater gains, as this would have favoured the experimental group in which more children began with lower scores. In fact, the gains made by low scoring and by high scoring children were similar. In the primary group there were no significant differences on letter names when the sample was split between those with lower starting scores and those with higher starting scores, either for the whole sample, or for experimentals or controls on their own. With word reading, there were no significant differences for experimentals or controls on their own, but in fact for the whole sample there was a tendency for those with higher starting scores to make greater gains ($t = 2.58$, $p < 0.05$, two-tailed test). It was not easier therefore for the experimentals to make greater gains on the basis of starting with lower scores. At the nursery stage there were no significant differences between the experimentals and the controls in terms of pre-test scores or on teachers' ratings.

In summary, following the intervention children in the experimental group made significantly greater gains in key baseline assessment scores than controls at both pre-school (letter sounds) and primary stages (letter naming, word reading).

Attitudes to reading

Of the 54 children seen at both pre-test and post-test, data on attitudes to reading were available for 53. Only one child was unable to engage with the idea of posting the cards into the glass jars.

There is a general tendency in young children to express positive attitudes regarding their abilities and interests, and this was seen for experimentals and controls both pre-test and post-test. A score of 1 was assigned to the most positive of the three choices for each question, with 2 for the middle position and 3 for the least positive. The overall mean at pre-test was 1.5 for both experimentals and controls, pointing to generally positive attitudes in both groups. There was no relationship between attitude and actual performance. This was demonstrated in a comparison of word reading scores and attitudes for the primary age group, both for overall attitude scores ($r = 0.02$) and for views as to whether they had done well on the test ($r = 0.08$).

However, significant differences were found between experimentals and controls at post-test in terms of attitude change. All of the scores that changed up from less to more positive and down from more to less positive were analysed by comparing actual changes with possible changes. This therefore took account of the fact that scores that were already at 1 could not change up, and scores at 3 could not change down. These overall change scores favoured the experimentals in terms of positive attitudes ($p < 0.01$, chi-square test). It may also be noted that, while the raw data for each individual question were too few to analyse separately, the direction of change in every case favoured the experimental group.

Teachers' ratings

It has already been noted that at pre-test the ratings by teachers for controls were higher than for experimentals. There was a substantial relationship between teachers' ratings and actual performance ($r = 0.55$ for ratings compared with word reading test scores, primary group). In terms of change in ratings there were no significant differences between the experimental and control groups from pre-test to post-test. It must be noted, however, that not all schools provided completed ratings and there were also some teacher changes during the course of the intervention. The numbers for analysis were therefore small and the data not comparable. The qualitative data from teachers are therefore more informative in indicating their views as to the effects of the intervention.

Semi-structured interviews

The feedback from staff about the project was uniformly positive. All indicated that the children had greatly enjoyed participating and that they had benefited from it, but to different degrees, with benefits being more apparent for some than for others. There was also strong support for the view that children's confidence, expectations, motivation or interest had been enhanced. Although there was not an opportunity to have separate interviews with parents, all comments made by staff regarding feedback from parents were highly positive. Details of staff responses are covered more fully in the section on qualitative feedback.

Interviews with children

As might be expected with an intervention of this kind, there was a virtually universal response from the experimental children that they had enjoyed making the declarations and that they thought this would help them to become better readers. It is the case, of course, that questions of this kind are unlikely to receive a negative response from young children. However, the sheer enthusiasm with which many of the children replied was itself a very positive indicator of how they had engaged with the programme. Of the 27 children interviewed about the intervention, 20 provided specific examples of declarations they had been making each day, and in many cases two or three different examples were given.

Of the seven children who could not give examples of declarations, one was at nursery and the remaining six at primary. Some of these had not appeared to engage with the programme at all, as not only could they provide no declarations but they did not seem even to remember anything about it when prompted. Two significant comments may be made about this group. First, all but one were low scorers on the baseline assessments. They included three of the four lowest scores obtained at primary stage on the word reading test. Spontaneous comments from their teachers referred in almost every case to being very slow, being unforthcoming or having learning difficulties. However, the more important comment about these 'non-engagers' is that they showed the same pattern of gains as the children who had clearly engaged. Further information from the interviews is covered in the following section.

All of the results summarised above applied across the socio-economic spectrum. For the quantitative analysis it would not have been possible to break down a sample of this size further to look separately at socio-economic factors. However, from inspection of the data it is clear that the whole experimental group benefited from the intervention, and this is supported by the qualitative feedback from high and low socio-economic establishments.

Qualitative feedback

While the quantitative results provided a good foundation for supporting the view that this study resulted in a range of benefits to the children who participated, it is most particularly the qualitative feedback from staff, pupils and occasionally parents that offered richer data on the value of the intervention and the ways in which it was carried out. This has been touched on in the outline of results given above and is expanded below through more specific examples.

Children

The most prominent feature of the children's responses was their enthusiasm. In the individual interviews they not only said they enjoyed the programme but in many cases emphasised this by vigorous nodding of the head and other signs of zeal. At nursery stage, the many examples they gave of declarations were simple but clearly effective:

I like books, books are fun. We'll all be great wee readers when we go to school. I can read by myself. I'm going to be a good reader.

The primary pupils added many others:

Reading is great, let's celebrate. Books and school are cool. I will become a very good reader. I am going to be a very good reader when I grow up.

Some of the declarations were more specific:

We will be neat – at our seat. We can do long sentences. We can work hard. I will read my work very well today. I love stories. I know all my words. I know all my sounds.

Often these were chanted three times by the whole class, and the value of using puppets as props was also apparent:

Different days we say something different to reading rabbit.

The declarations had also spread somewhat beyond literacy:

I will do my sums very well today.

Some of the responses were very thoughtful. One child, when asked if the programme had helped her, replied:

Yes – I'm getting better: we're doing more than we normally would.

Sample primary classes or nursery groups were visited at the end of the intervention and staff were enthusiastic in demonstrating how the programme was being implemented. One strategy that was proving very popular was to encourage groups of children to shout their declarations at the highest possible volume – a strategy that clearly produced maximum response for minimum prompting. In one nursery this reached a great crescendo:

*Rhymes are fun, rhymes are cool
We're listening to sounds because we're going TO SCHOOL!*

Staff

At nursery level, feedback was equally positive from the higher and lower socio-economic groups. In both cases the intervention was extended to all the children in the nursery at levels that went well beyond the basic requirement. Strategies that were developed included:

Collective chanting, getting louder each time. Whispering and shouting. Using signs as well as speech. Doing actions to it – keep fit and hand-clapping. Staff making declarations in confirmation of those made by the children.

Declarations were both general and specific, the latter including:

*Looking for the first letter of their own name and declaring, 'I can find the letter k'.
On library day saying, 'I'm going to take this book home to read – I'm going to be a good reader'.*

The possible scope of the strategy was realised:

We began with declarations about reading, but as time has gone on it has been applied to everything else as well.

At primary level the picture was similar across establishments and was again highly positive. While everyone carried out the programme as planned, different schools and different classes varied, in that some followed the minimum requirement while others developed the implementation more extensively. The significant factor is that everyone met the requirement. This has important implications for project work in schools and these are discussed below. Comments on implementation from the primaries included:

*We used the declarations for other things such as sums and found them good there too.
We 'sneaked' it in everywhere through the whole day.*

Group competition was encouraged.

Let me hear the snails – I wonder if they can do better.

In one school the declarations were made as ‘promises’, the children standing with feet together and arms folded. Success was reinforced by the teacher:

That’s because you made your promise.

The children asked if they could do their own private promises, and so at times they were making quiet, personal declarations.

It was clear for both nurseries and primaries that the project was meeting its aim of supporting delivery of the existing curriculum rather than introducing new curricular activities.

The activities were just the normal – the only difference was the declarations, and they did respond. It provided a medium for reinforcing what we were already doing. It’s a ‘snappier’ way of doing it. It was easy to integrate into the routine – it wasn’t extra work.

One of the nurseries built on work they were already doing – ‘Tell Ted’. Ted (a teddy bear) became the focus of literacy activities. The children liked to correct him when he held the book the wrong way up. Ted also served a function mentioned by other schools in relation to puppets or cuddly toys introduced for declarations. A child could speak through the puppet – for example, to spell a word – and if it was wrong it was the puppet’s fault and not the child’s. This increased confidence to take the risk of making mistakes.

Staff were unanimous when asked if the children enjoyed the intervention:

Absolutely. They got a lot out of the chanting – it definitely built their enthusiasm. Thoroughly enjoyable. If we forgot, they prompted us. When you go to do a story you hear them chanting before you even get the book organised. They’re all desperate to come out and have a try at reading now.

Many responses were given with regard to the benefits experienced by the children:

It pulled them together and motivated them as a whole group. It’s developing confidence and encouraging a positive attitude. It really helps to concentrate on what they’re doing. Definitely more interested and more confident. It made children expect more. They became more aware of all the different things involved in reading. It gave them more power over knowing how to become good readers. One child has the puppet at home and now he likes to do his reading at home too. The more confident ones are surpassing themselves in what they can read. If they chanted before a writing lesson you could guarantee writing would be better that day. It’s definitely making a difference.

Teachers varied in their views as to whether all children received the same benefits, with some questioning in particular whether less able children were being helped:

I feel there was a benefit for the 'better' and the 'middle' children but not for all – there was a slight change for the 'lowest'. The brighter children like it more than the strugglers – for them, over a longer time, I think it would make a difference.

This observation is compatible with the trend of results previously noted, the indications being that the more able children made greater gains. Nevertheless, the experimental children at every level made gains compared with controls. This was highlighted in two of the schools. In one, it was felt that one or two children who were very shy and unforthcoming had become much more confident and independent. In the other the teacher noted:

It was remarkable – it was a random sample, yet it picked up the children who most needed help.

(This is a feature regarding the sampling on which comment has already been made). She continued regarding one of these pupils (who made exceptional gains in his scores):

It was the most amazing thing one day – suddenly he read a lot of words and we said, 'This was because he made his promise today: he said he was going to read well and he did'.

And regarding another:

She is the one who has really taken it to heart and really believes it – she will definitely come on well now.

Several staff commented that the project had either reinforced their views of what they were already doing by giving them a new strategy, or that it had changed their expectations:

It has raised our awareness. Staff have higher expectations – when they see the children progressing their expectations rise. You know it makes sense. Once you're used to it you just do it without thinking. We started feeling they should be reading more and more. I didn't know what to expect at the start and was apprehensive – but once you know that something works it's all worthwhile. It was a very positive experience.

Parents

Although the project did not have the scope for incorporating parent interviews, considerable informal feedback from parents was received spontaneously during interviews with staff. One primary school, which had just had a parents' night commented:

A lot of parents said the children were enjoying reading more.

One parent noted in particular that her child...

...used not to like the reading homework but really enjoyed doing it now.

In one of the nurseries the parents were reporting that the children were chanting their declarations in the car, and the mothers had started the chants too. At the other experimental nursery a non-speaking child started putting words together and her mother came to the teacher and said:

She keeps saying something over and over, and I don't know what it is.

It was the declarations.

A further comment came from the only parent with whom the researcher had a direct (and chance) encounter. She said:

Christina has been saying her chants everywhere. Everybody she meets she's been saying, 'I'm going to be a good reader'. It's really raised her confidence and her expectations, and it's spilled over to other things as well. She has really enjoyed it.

Discussion

The declaration study was very much a novel and exploratory investigation and as such is subject to a number of limitations. It requires replication in both similar and different settings, and the development of a rationale drawing from established areas of mainstream psychology is at an incipient stage. The method devised for assessing attitudes to reading is also a novel one. While it has face validity in containing direct questions on the items it seeks to measure, and while steps have been taken to overcome problems of unreliability previously encountered in this area with young children, further comment on its validity and reliability cannot be made. Reference has also been made to issues regarding the representativeness and size of the sample obtained for direct assessment.

Despite these limitations, all of the data gathered from multiple sources in relation to this project on raising expectations and raising achievement pointed strongly in one direction, namely, that the children who participated gained benefits in terms both of their achievements and of their positive attitudes, motivation and confidence. These findings were particularly encouraging because both of the brevity of the intervention and the recognised difficulty in changing measured levels of literacy attainment in very young children.

Recognition of change processes in real world research

Throughout this declaration study, as in the main study, there was explicit recognition and anticipation of the ‘messiness’ of real world research (see Chapter 2). The implications of this factor were highlighted in initial meetings with the staff involved and steps were taken to ensure that unpredictable difficulties would not undermine the outcome. Aspects of the project itself confirmed the ‘messiness’ of school interventions. The timing of the intervention was susceptible to factors such as staff maternity leave and the difficulties of finding dates when key staff from six establishments could meet together. Situations of staff changes and supply cover arose during the implementation phase. The systematic sample showed pre-test differences between experimentals and controls. An incomplete return of pre-test and post-test ratings from staff was received, and the post-test ratings were not all completed by the same staff as the pre-test ratings. Also, it became clear at the start that the planned arrangements for education authority personnel carrying out the baseline assessments were going to be too demanding for the staff concerned, and this very considerable task had to be organised instead by the researcher. Nevertheless, the outcomes were extremely positive and the project was successfully maintained through its potentially difficult phases.

Recognition was also given to the importance of *context* as well as *content* variables. The question of *how* the project would be successfully delivered was viewed as being crucial as well as the question of *what* was to be delivered. The importance of factors already highlighted such as vision, profile, commitment and ownership was acknowledged. These factors were therefore built in from the beginning: the vision of a novel strategy that could achieve something not achieved before; high profile in the participating classes, with head teachers encouraged to extend the profile throughout the whole establishment; the commitment to fulfilling the project specification throughout the implementation period; and ownership of the project at every level of involvement – the authority, the researcher, the staff, the pupils, the parents. It was clear at the final interview stage how fully the question of ownership had been embraced by teachers and children.

Achieving a high level of commitment by staff was supported by the methodology adopted. Many educational projects fail because they are not actually implemented – they make too many demands on school staff, and when difficulties arise they are abandoned or else carried out with low fidelity. In this project the commitment to making three simple declarations a day was not at all onerous, and helped to guarantee that implementation would take place. A suitable balance was struck between leaving scope in the programme for staff to develop their own declaration strategies, and at the same time ensuring that whatever difficulties or pressures they might face they could fall back on some simple, ready-made ideas. As one of the teachers who ticked the checkbox for carrying out the project ‘at minimum requirement level’ commented: *I kept it simple and to the minimum*. This fulfilled the specification, and her results were good and her response enthusiastic.

Conclusions

The declaration study reinforced further the importance of attitudes and expectations as a factor in enhancing children's literacy. It demonstrated that improvements could be achieved in literacy scores simply by engaging teachers and children in the positive process of making bold declarations about their enjoyment of reading and in particular about their future levels of literacy achievement.

The study was completed at the end of session 1999-2000, and its findings were utilised in informing the implementation of the main study from session 2000-01. More wide-scale formal studies of declaration were not introduced into the main study during the first phase, but the lessons from the declaration study were disseminated to the staff working in the main study through the processes of staff training and information exchange to emphasise further the importance of the strand relating to attitudes and expectations.

SUMMARY

This chapter presents the results of the introduction of a novel strategy in primary and nursery schools – having children make declarations about their future levels of reading achievement. In nursery, experimentals showed significant gains in three baseline assessment tests: nursery rhymes, rhyme production and letter sounds. In Primary 1, gains were significant for five tests: nursery rhymes, letter sounds, letter names, non-word reading and word reading. Teachers also gave the experimentals higher ratings for a range of characteristics in relation to reading: progress, showing promise, confidence, interest and enjoyment. Pre-post changes in attitude also indicated that the experimentals developed more positive attitudes to reading. These results were supported through teacher and pupil interviews. It is concluded that the simple act of making declarations about future levels of reading achievement has a beneficial effect on attitude and performance.

Chapter 15

Individual Support: The Need for Strategies Beyond Early Intervention

Introduction

The individual support study developed another of the 10 strands of literacy intervention described in the main study, strand 7, 'identification of and support for children who are failing'. This was a key strategy for addressing one of the main long-term aims of the study, the eradication of illiteracy throughout the entire school-age population. While the baseline assessments in the pre-school year and Primaries 1 and 2 provided an excellent basis for early identification of reading failure, it was also recognised that two factors relating to the later years of schooling required to be addressed. First, there were many pupils in the upper primary years and in secondary school who were already experiencing reading failure and who were not going to benefit from a literacy intervention focused on the early years. Second, a strategy was needed for the identification and support of children who were still failing even after they had been through the early intervention programme. It was in relation to tackling these factors that the individual support study was developed.

The choice of individual support was a deliberate one. It reflected not only a belief that children who fail to learn to read in the early years require a high level of learning support, but more specifically a view that support provided at group or class level would not offer a sufficiently confident basis for dealing with illiteracy. This view was based both on experience of working with different models of intervention in this population over a period of years and also on the available evidence base.

In relation to lessons learnt by experience, there was no support for the view that children who had failed to learn to read in the early primary stages were going to succeed by the normally available group and class methods in later years. Decades of systematic observation and practice of educational psychologists working in areas of socio-economic disadvantage had supported the opinion that such children could make good progress on one-to-one remedial tuition; that the effects of such tuition were later washed out if programmes were not maintained until the attainment of a functional literacy level; and that group methodologies did not bring about sufficiently fast progress to establish satisfactory outcomes. This had been borne out many times when reading levels at school leaving age were assessed, even after group learning support had been provided.

In relation to the available evidence base, data collected from secondary school learning support teachers regarding the group methods traditionally used by them pointed to the inefficacy of these methods, while the research literature supported good individual instruction. One example of each of these areas is considered further here.

MacKay and Boyle (1994) surveyed the views of primary school head teachers and secondary school learning co-ordinators regarding the actual and ideal contribution of psychologists to pupils with learning difficulties. The results were revealing in illuminating the different attitudes and approaches between primary and secondary in relation to such pupils. While primary schools were emphasising the need for advice on teaching approaches and materials, the concern in secondaries had shifted away from this area to a need for individual support and counselling of pupils with difficulties. This finding was in line with the common experience of psychologists in this area. At primary stage the teaching staff tend to have a view that children who have failed to achieve success in literacy should still have their skills addressed through any strategies as yet untried. On the other hand, in the secondaries there is frequently a view that if pupils have already failed through seven years of primary there is little point in subjecting them to further years of failure on the same set of skills. This is borne out by the negative experience secondary learning support teachers often have of attempting to teach basic literacy skills. There is, however, a recognition that these pupils will often have developed problems in their overall adjustment, and that they will at times need a higher level of psychological support.

In terms of the research literature the value of individual tuition has been highlighted for a very long period. Clay's Reading Recovery programme (Clay, 1979b, 1993a, 1993b) was introduced in the early 1980s following extensive research in the previous decade. It was designed to enable 6-year-old children who had made poor reading progress to acquire essential literacy competences. The programme lasts between 12 and 20 weeks depending on progress, with daily 40-minute individual lessons with a teacher trained for this purpose. Effectiveness studies have demonstrated Reading Recovery to be a successful strategy, with less than 1% of children being unable to complete the programme satisfactorily (Pinnell, De Ford & Lyons 1988), although methodological criticisms of the evidence have been raised (Chapman & Tunmer, 1991; Iversen & Tunmer, 1993). However, Reading Recovery is expensive, being dependent on a lengthy training course, and despite good evaluation in a UK setting by Wright (1992) its lack of economy was the main reason for its being phased out in UK schools after its introduction in 1993 (Nicolson, Fawcett, Moss, Nicolson & Reason, 1999).

The fact that Reading Recovery on the one hand is supported by effectiveness studies pointing to successful outcomes, and on the other hand may prove too expensive to make it an economically viable option, merits more detailed consideration. It points to the need for studies that identify its successful elements, leading to less expensive interventions incorporating these elements. This was the aim of the study by Iversen and Tunmer, 1993, and it is considered in greater detail here.

Iversen and Tunmer questioned aspects both of the method and of the philosophy of Reading Recovery, and asked the fundamental question as to whether its specific procedures and instructional strategies were more effective than any other remedial approach. In particular they questioned Clay's (1985, 1991) argument that instruction in alphabetic coding should normally arise incidentally in the context of reading connected text. To determine whether the Reading Recovery programme would be more effective if systematic instruction in alphabetic coding were incorporated, they

divided first-grade at-risk readers into three matched groups of 32 children each: a modified Reading Recovery group (individual instruction), a standard Reading Recovery group (individual instruction) and a standard intervention group (small group instruction using normally-available learning support structures). The children in the modified Reading Recovery group received explicit alphabetic instruction involving phonograms. Both Reading Recovery groups out-performed the standard intervention group and reached the criterion required for discontinuation of the programme. However, the modified group reached the criterion much more quickly. It was also found that the children selected for the groups were particularly deficient in phonological processing skills and that their progress in the programme was strongly related to the development of these skills.

This study followed a rigorous methodology throughout in terms of assessment and selection for the programmes, matching and allocation to groups, training of teachers in assessment and intervention, reducing possible bias in assessing outcomes and analysing and interpreting results. Triplets for the three groups were exactly matched on the basis of their scores from key subtests of the assessment. All had very low scores on a series of literacy assessment measures. They were assigned to groups on a quasi-randomised basis, but because of geographical constraints, assignment could not be completely randomised. Since the teachers in the modified and standard Reading Recovery groups were trained in different sessions at different locations, they were not aware of any differences in their procedures. They also taught children who were not target pupils in the study, and had no knowledge of who the target pupils were. These steps served to reduce possible teacher bias in judgements about when a child's programme should be discontinued.

One-way analyses of variance of the means of the three groups were carried out for all measures used at discontinuation, and a full breakdown of means and standard deviations was provided. There were no significant differences in test scores between the two Reading Recovery groups, but both performed significantly better than the standard intervention group on each of the 10 measures assessed. The Reading Recovery groups were also compared with children from their own classrooms who were judged by their classroom teachers to be performing at an average reading level. Not only did they perform as well as these classroom controls, but on several measures they performed better.

As research has frequently indicated the advantages of individual over small group instruction for failing readers (Bloom, 1984), it is unsurprising that the two Reading Recovery groups performed better than the standard intervention group. The key significance of this study, however, lies in the mean number of lessons required to reach discontinuation criteria between the two Reading Recovery groups. The mean for the modified group was 41.75 lessons, while the mean for the standard group was 57.3 lessons. That is, the standard Reading Recovery programme took much longer to reach the same point. In this respect it was 37% less efficient than the modified programme involving alphabetic instruction. The measures used at discontinuation were repeated at the end of the school year and no differences in performance had arisen between the two Reading Recovery groups.

This study represents a very positive and rigorous evaluation of Reading Recovery as an effective programme, but points the way to modifications that increase its effectiveness and its economy. It is possible that the extra time spent by children on the standard Reading Recovery programme produced benefits of a kind that could not be assessed using the battery of standard tests adopted for this study. The study did not report any qualitative measures involving the views of teachers or pupils on their participation, attitudes, enjoyment of reading or desire to engage in reading for pleasure. For example, the greater emphasis in the standard programme on story book reading and writing may lead to greater motivation to read and more pleasure in the reading process, or other factors not assessed by the static battery of tests used.

The study also opens up a number of additional questions of central importance to the current investigation. It may still be asked whether it is the distinctive methodology of Reading Recovery that is effective in addressing reading failure, or whether it may not simply be that it is an intensive programme of sound individual instruction carried out competently at a frequency that is likely to lead to success. In the Iversen and Tunmer study, in addition to the expense of an intensive training programme for teachers in the use of the method, they opted in every group for teachers who were reading specialists with a master's degree in reading.

The study reported here recognises the superiority of individual as opposed to group instruction for poor readers. However, in seeking to establish methods of addressing reading failure that have economic viability for applying widely in disadvantaged populations, it has focused on using a variety of staff and volunteers to carry out a programme that is inexpensive in materials and that involves only a minimal level of training.

Aims

The aims of this study were:

- to identify the children at P7 stage throughout the authority with the most significant reading difficulties
- to carry out an effective intervention for acquiring key literacy skills based on individual support.

Hypotheses

The following hypotheses were proposed:

- 1 that the experimental pupils in the secondary study would achieve higher reading scores than the controls
- 2 that the children on the programme in the primaries would show large gain scores.

Programme

The programme selected for individual support was 'Toe By Toe' (Cowling & Cowling, 1993). Its sub-title describes it as 'a highly structured multi-sensory phonetic approach to literacy'. The creator of the system, Keda Cowling, developed it painstakingly over a 25-year period as an answer to the frustrations she experienced as a teacher in addressing the needs of children with dyslexia. This programme was selected on the basis of content, method and cost-effectiveness.

In terms of content, it was phonics-based, and contained all the necessary building blocks of sounding, blending and rules required for the development of sound competence in literacy. In terms of method, it was direct. It provided individual instruction in exactly those aspects of literacy that were going to be most beneficial. It also followed a routine whereby learning was reinforced and consolidated so that it would be thoroughly established. It contained all the ingredients required for fast success in each of the skills taught, and was therefore likely to be motivating to children who had previously experienced failure. This was an important consideration, especially as the programme was not designed to entertain – it contained no pictures, no colour and nothing to give any light relief from sheer hard work. All of the rewards arose from success. In terms of cost-effectiveness, it cost a fraction of a programme like Reading Recovery, as it did not depend on a lengthy and expensive training programme. Indeed, the absolute simplicity of the scheme commended it to any user without training. It did not even require a separate 'teacher's book' to support the manual, as everything was plain on each page of the programme itself. Nevertheless, the project was supported by a small training input. Cost-effectiveness was further supported by the fact that it did not require teachers to do the tuition. It could be done by anyone with the skills and ability to sit down regularly with the child and provide the tuition.

The programme itself works its way systematically through individual letter sounds, digraphs and blending exercises, mainly using nonsense words to encourage fast and accurate word-attack skills. There are also fluency exercises using whole sentences. A key feature is the teaching of 'polynons' and 'syllable division'. Polynons are described as words that are both polysyllabic and meaningless. Their only function is to serve the needs of word-building and syllable division. Syllable division is described as a multi-sensory activity in which students learn to divide words physically with a pencil stroke prior to reading polysyllabic words.

Toe By Toe does not have an index or even a contents page. These omissions are deliberate. An essential feature of the scheme is that it starts at the beginning, and is followed systematically right through to the end. It is not designed for picking and choosing selected exercises to boost a chosen skill. Rather, it is based on a commitment to the view that children with reading difficulties have often failed using other methods and have uneven development of skills. The programme aims to address this by ensuring that all necessary basic skills are taught in a strict and comprehensive sequence. Those who already have abilities in particular areas find that they quickly move through the exercises covering these areas, so time is not wasted by dwelling on material that has been previously learnt successfully.

Each double page of Toe By Toe follows the same format. On the left-hand page there is a box with simple instructions for the tutor or 'coach'. There is also a column headed 'other information', containing any background comments of interest or relevance to the lesson being covered. On the right-hand page there is a grid with the date at the top showing the sounds or words to be read during that lesson, together with spaces for achieving three consecutive ticks to demonstrate competence on each item. It is recommended that up to 20 minutes is spent using Toe By Toe each day. Sample pages from the manual are shown in Appendix 5.

Toe By Toe: results of other studies

Prior to this study a number of unpublished studies using Toe By Toe had provided an encouraging basis for its use as an effective intervention. These included a pilot study carried out by the learning support department in the secondary school selected for the current project. The pilot study involved 14 pupils at S1 stage, and 10 pupils in S2. The S1 pupils began with an average reading age of 7y 11m and following the programme this had risen to an average of 9y 11m, a gain of two years in reading age over a period of eight months. The range of gains was nine months to 3y 8m. The S2 pupils began with an average reading age of 8y 3m rising to 10y 0m, a gain of 1y 7m, with a range of five months (gained by a pupil who was often absent) to three years. These pupils started at various different times in the session and the period over which the gains were made varied from pupil to pupil.

The school's pilot study had two limitations that this study intended to address. First, it showed gain scores only for pupils who had taken the programme, as there were no controls. Second, gains were measured on a test with a ceiling of 10y 6m (the Salford Reading Test). Ten out of the 24 pupils in the above S1 and S2 samples scored above this ceiling, indicating that gains were likely to have been greater than those reported.

Results had also been gathered from several other trials. Twelve pupils in P6 and P7 classes in a West Dunbartonshire primary school had shown average gains using Toe By Toe of 1y 9m from reading age 9y 2m to 10y 11m over a period of six months. Similar results had been found in schools in England and Wales, with 10 pupils age 15 years rising from average reading age 8y 1m to 10y 10m, a gain of 2y 9m over a period of 10 months at Yny Sawdre Comprehensive School, and 11 pupils from Primary 7 through to S2 at St Richard Gwyn High School rising from 8y 2m to 9y 10m, a gain of 1y 8m over a period of seven months. In the last mentioned case, the gains may have been greater since several pupils started below the 7y 5m norms, and so their exact reading age was not recorded. This school also took the opportunity to assess the self-esteem of the Primary 7 pupils using the questionnaires devised by Lawrence (1996), on the basis that an increase in reading ability was likely to be associated with a complementary gain in feelings of self-worth. All of the six pupils assessed showed an increase in self-esteem scores, with two moving from the 'low' to the 'medium' category, one from the 'medium' to the 'high' category and one from 'low' to 'high'.

A summary of these results is shown in Table 15-1.

Table 15-1 Unpublished studies of reading age gains using Toe By Toe

School	Stage	N	Pre-test RA	Post-test RA	Interval	RA gain
Braidfield High School, West Dunbartonshire	S1	14	7y 11m	9y 11m	8m	2y 0m
Braidfield High School, West Dunbartonshire	S2	10	8y 3m	10y 0m	Various	1y 7m
Primary school, West Dunbartonshire	P6-P7	12	9y 2m	10y 11m	6m	1y 9m
Yny Sawdre Compre- hensive School, Wales	Age 15	10	8y 1m	10y 10m	10m	2y 9m
St Richard Gwyn High School, England	P7-S2	11	8y 2m	9y 10m	7m	1y 8m

At the same time as the second phase of the present study was being carried out, a study of the use of Toe By Toe in two Scottish education authorities was conducted by McConnell (2004). In surveying the views of teachers, parents and pupils on the effectiveness of the programme she concluded that for the vast majority of children it represented an effective intervention that was enjoyed by the pupils, supported by parents and recommended by teachers. Differences in the way the programme was being used across the two authorities were investigated, with one using it less frequently than daily, and for a shorter time each day than the recommended period. Her conclusion was that the programme works most effectively when used as recommended. This view is strongly endorsed in the present study on the basis of the consistent experience both of the author of the Toe By Toe programme and of the extensive experience across many settings of the secondary learning support teacher who carried out the teaching programme in the first phase of the study.

Method

This study was carried out in two phases. First, a quasi-experimental study was conducted in one West Dunbartonshire secondary school, with 24 pupils referred for learning support because of low reading levels. Of these pupils, 12 controls were assigned to the normal learning support programme while 12 experimentals were enrolled in the intensive individual support programme using Toe By Toe. The allocation of cases to the two conditions was not random, as it had to be subject to the normal timetabling and other constraints encountered in a large secondary school. However, the two samples were matched as closely as possible.

Normal learning support for the controls comprised two one-hour tutorial sessions, one concentrating on the development of basic punctuation and comprehension skills

and the other developing phonic skills using standard phonic workbooks. Time was also allocated to individualised spelling and paired reading programmes which took place within the mainstream English class, with support for learning staff assisting in a co-operative teaching capacity.

All experimental pupils received individual tuition for 20 minutes a day, and the programme lasted approximately three months. Pre-post assessments were conducted at the start and finish of a 12-month period.

Following further piloting of the programme in secondary and primary schools in the authority, the second phase of the study involved the identification of pupils at upper primary level (mainly Primary 7) who were experiencing significant reading difficulties. These children were initially identified by staff in the 35 primary schools in the authority, and following individual testing of 118 children, 104 were selected from 32 schools as meeting support criteria. The 14 who were excluded from the sample had a reading age above 9y 6m and these pupils were not viewed as having a reading problem. The final sample comprised 91 in P7, 12 in P6 and 1 in P5. Pre-testing took place in November-December 2002, with post-test in May 2003.

Approximately 120 individual support workers were trained in the use of the programme by the researchers. These were drawn from a wide range of personnel – teachers, classroom assistants and volunteers. Monitoring and support structures were put in place to ensure effective implementation.

Results

The 12 experimental pupils in the secondary study showed mean reading age gains of 2y 0m (from 8y 2m to 10y 2m) following the three-month Toe By Toe intervention, and with a 12-month interval between tests. The controls gained only four months (from 8y 5m to 8y 9m) during the same period ($t = 5.65$, $p < 0.001$, effect size 1.74), (independent two-sample t tests using the Microsoft Excel Version 8.0 data analysis tools). The test used was the Gapadol Reading Comprehension Test (McLeod & Anderson, 1972). Effect sizes were calculated using the standard deviations in the test manual. These results are shown in Figure 15-1.

The 104 children in the primary schools study were all individually tested on the Neale Analysis of Reading Ability, 2nd Revised British Edition (Form 2). Their average pre-test reading age was 8y 0m, this being about three years behind their chronological age. After a period of just under six months their post-test reading age had risen to 9y 2m, giving an average gain score of 1y 2m. The breakdown of gains is shown in Table 15-2.

Figure 15-1 Changes in reading ages: secondary school sample (N = 24)

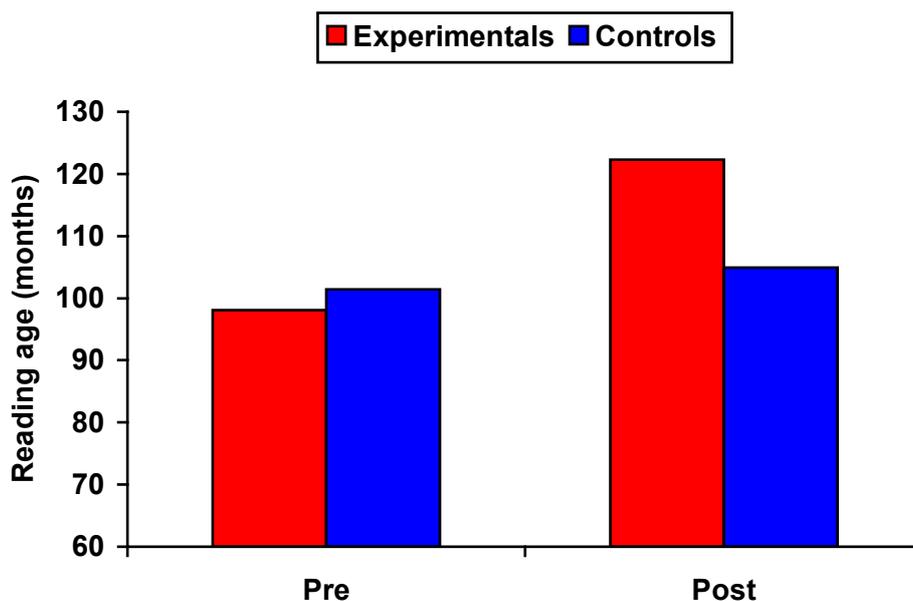


Table 15-2 Gains in reading age: primary school sample

<i>Gains in reading age</i>	<i>Number of pupils</i>
0-5 months	19
6-11 months	25
12-17 months	30
18-23 months	11
Over 2 years	16
Over 3 years	3

There were no differences of any note in the gains made by children in terms of lower or higher pre-test scores. When the sample was divided in half, those who started with lower reading ages had mean gains of 1y 1m, while those in the higher half gained 1y 3m. The lowest 25 children on pre-test – those who began with a reading age below 8y 0m – gained 1y 0m. Thus there were mean gains of a year or more for the whole sample, irrespective of starting point. At the bottom end, 18% of the sample (19 children) showed gains of less than six months, while at the top end 18% showed gains of two years or more.

Conclusions

The individual support study proved to be highly effective in addressing reading difficulties both in the secondary and in the primary school samples.

The average gain of two years in reading age for the secondary pupils was particularly encouraging. These gain scores were so large that they would have

demonstrated the effectiveness of the intervention even without the need for a control group. However, the use of a quasi-experimental design allowed informed comparisons to be made between experimentals and controls. The fact that, even with a good programme of traditional learning support, the gains of the controls were only four months in the course of a full year confirms the routine experience of learning support teachers and educational psychologists – namely, that pupils at this level of reading difficulty tend normally to make annual gains that are less than half of what might be expected. That is, in the course of a year they are making somewhat less than six months of improvement in their reading ages.

There are times when statistically significant changes do not make a real impact either on the perceptions of subject teachers or on the quality of life of the participants. This was not the case in the present study. Teachers of various subjects frequently commented that the pupils on the Toe By Toe programme were reading more competently than had ever been the case before with similar pupils. The position is best summed up in the words of one 14 year old pupil in the study, as cited elsewhere in this work:

When all this started I couldn't read. I was a failure. Now I have a cupboardful of books at home. My favourite authors are Roald Dahl and J.K. Rowlings. Now I am a success.

The gains achieved consistently by the secondary school group provided the basis for the extension of the study into the primary schools as an intervention with an established record of effectiveness. Here again the gains were sufficiently large that they showed the programme to be successful without the need for comparison with controls. These pupils would normally have made significantly less than six months of reading gains over a six-month period. The consistent experience of learning support staff, combined both with the results of the control group in the secondary study and with any assessment of normal rate of progress of these pupils over the years, would point to usual gains of between two and three months during any six-month period. The actual gains during the intervention of 1y 2m represent a population shift of 1.75 standard deviations on the Neale Analysis of Reading Ability. In terms of the above observations, the estimated 'expected' shift (that is, the normal mean gains without intervention) would have been about 0.35 standard deviations over the pre-post interval. Therefore, the effect size of the intervention may be estimated at about 1.4. This suggests that the programme was extremely powerful over this short period in increasing levels of reading achievement.

Two further observations may be made regarding the progress of the primary school sample. The first is that in comparison with the secondary sample their gains were 1y 2m as against 2y 0m. However, two main considerations are likely to be of importance here. First, the pre-post test interval was one year for the secondary pupils, while it was less than six months for the primary sample. It was expected that the primary pupils would continue to make reading progress over the succeeding months, thereby narrowing the observed gap in average reading gains. Second, the secondary pupils were all taught by one teacher, who was already an expert in the use of the programme. By comparison, the intervention with the primary pupils was

spread across 120 helpers, many of them volunteers, and most of them being inexperienced with the programme other than for a training session of half a day.

Comparisons of the weight and significance of gains in reading made by pupils of different ages and in different intervention conditions can be difficult to interpret. The concept of 'ratio gains' may be applied to the scores reported here, as defined by Topping and Lindsay (1992) – '...the gain in reading age made...on a reading test during a chronological time span, expressed as a ratio of that time span'. The mean ratio gain of 1y 2m in approximately six months for the primary sample is therefore about 2.3.

Overall, the results achieved are clearly very good and point strongly to this programme as being an effective intervention that does not depend on experts or on intensive training. Indeed, follow-up investigations carried out since this part of the study was completed have indicated that many of the helpers were not even carrying out the programme correctly and needed a higher level of monitoring and support. Arrangements to undertake this were put in place, and it is expected that future interventions may show still higher gains.

The second observation is that the primary sample included 19 pupils whose progress over the six-month period was less than six months in terms of reading age gains. However, several of these children were in fact making progress at a reasonable rate given their level of difficulties. In many cases they had still not completed the programme at post-test and further progress was therefore expected. In other cases it was clear why less progress had been made. With a sample as large as this, dependent on so many helpers, there will always be some slippage in the intervention, such as helpers being off sick or pupils being absent from school. Every case of low progress was investigated and further help was arranged to ensure success for all pupils.

In conclusion, this study provided strong support for strand 7 of the main study, 'identification of and support for children who are failing'. A foundation has been laid for the eradication of illiteracy throughout the primary schools, with a rolling programme in place each year to identify every individual pupil with a reading difficulty. At the end of the brief intervention for primary school children, over one-third of them were no longer described by their teachers as having a 'reading problem'. Opportunities have also been extended throughout all of the secondary schools in the area to provide intensive help using the programme for pupils who have not achieved adequate levels of literacy. This is a crucial area of intervention for these pupils as they prepare to face a future beyond school. As Wells (1998) has observed:

While it's important to get the teaching of literacy and numeracy right in primary schools, early intervention will be too late for some older pupils. So opportunities for catching up in secondary school will need to be given high priority. If they aren't, many pupils won't be able to get much benefit from the wider curriculum. And some will leave school with basic skills that provide hardly any grounding for the world of work and later education and training.' (p. 1).

While it is clearly of crucial importance to identify children who are failing in literacy at the later stages of their primary schooling or in secondary school, and to provide them with effective interventions, methods must be found at an earlier stage of identifying those children who are likely to experience difficulties at a future date. This may be done by a combination of baseline assessment results and responding to concerns raised by teachers about children who are not making adequate progress in the early stages of literacy. Four-fifths of the children who were identified as reading failures at around Primary 7 level were to be found in the bottom quartile of baseline assessment scores for key literacy skills when they were in Primary 2. Almost half had scores falling in the bottom 10%. Only one child was identified as having significant difficulty in reading at Primary 7 who scored above the midpoint for reading scores in Primary 2.

Table 15-3 shows a breakdown of the distribution of scores for early reading skills for the 53 children out of the 104 in the primary sample who could be identified in the Primary 2 baseline assessments. Early reading skills (letter sounds, the alphabet, letter names, non-word reading and word reading) proved to be a much more robust predictor of later reading difficulties at this stage than phonological awareness (nursery rhymes, initial sounds, rhyme detection and rhyme production).

These results suggest that a suitable starting point for identification of future reading difficulties at Primary 2 stage would be the bottom 10% of early reading skills scores on the baseline assessments. This one measure would identify approximately half of the children who would be seen to have a reading difficulty in the later primary years.

Table 15-3 Prediction in Primary 2 of later reading difficulties (N = 53)

	Early reading skills		Phonological awareness	
	%	Cumulative %	%	Cumulative %
<i>Bottom 5%</i>	25	25	23	23
<i>Bottom 10%</i>	22	47	3	26
Quartile 1	79	79	51	51
Quartile 2	19	98	26	77
Quartile 3	2	100	17	94
Quartile 4	0	100	6	100

SUMMARY

This chapter describes a subsidiary study designed to develop a key strand in the multiple-component intervention – identification of and support for children who are failing. It examines the rationale for providing individual rather than group support. A quasi-experimental study at secondary school and a gains score study at primary

are described. The secondary study involved a comparison of 24 pupils, 12 experimentals who received an intervention programme based on individual support and 12 controls who received the normal learning support package. A commercially-available programme, *Toe By Toe*, that met the specifications for the study, was used. This involved individual, structured tuition in basic literacy skills for 20 minutes each day. The intervention lasted for three months, and pre-post assessments were conducted 12 months apart. Significant gains were made by the experimentals, with an average reading gain of two years compared with 4 months for controls. Following the secondary study, 104 pupils in upper primary school were identified on the basis of low reading scores. They were given the *Toe By Toe* intervention from volunteers and teachers following a brief training session. In less than six months the average gain scores for these pupils was 1y 2m. These results pointed to an economical and effective way of addressing reading failure in pupils at upper primary and secondary school level.

Chapter 16

The West Dunbartonshire Literacy Initiative in Action

Kathy Morrison

Project Leader and Head Teacher Early Intervention

The West Dunbartonshire Literacy Initiative is vigorously based on the application of research findings, which have demonstrated successfully that they can raise reading achievement. The standardised approach adopted in all schools and pre-five establishments has ensured success. Early intervention staff are deployed in every school and pre-five establishment in the authority – a total of 35 primary schools and 25 pre-five establishments (23 at the commencement of the project in 1997). Staff work with all children in their pre-school year and with pupils in Primaries 1, 2, 3 and 4 to raise attainment and achievement in literacy and numeracy.

The early intervention team work collaboratively within schools and pre-five centres. The main focus is literacy with a recommendation that of the allocated time in a school or nursery 75% is spent on literacy activities and the remaining 25% on numeracy activities. In reality most schools have deployed their early intervention teacher for literacy. There are no targeted or pilot schools: all schools and nurseries receive an allocated amount of teaching time from the early intervention team. This can vary from one day to three days dependent on the needs of schools. Allocation is based on levels of socio-economic disadvantage. Originally the early intervention team supported children in their pre-school year and pupils at Primary 1 and Primary 2. From Session 2002-03 the programme was rolled out to failing readers at Primary 3 and Primary 4. Extra staffing has allowed this to happen and all schools have received an extra half-day's allocation as a result. The focus at this stage is reading and children who are failing.

The role of the early intervention team is to add value to what schools do to raise attainment in literacy and numeracy for all children. Deployment and best use of the early intervention teacher is the responsibility of individual establishment managers. Needs of establishments vary. For example, in some schools the early intervention teacher can work with all children at a particular stage, while in others they support small groups and individuals. The role of the early intervention teacher in the pre-five sector is to develop the pre-school child's phonological awareness and to provide a wide range of experiences to develop children's concepts about print. Children are also actively encouraged in their attempts at early writing.

At Primaries 1 and 2 the teaching of phonics is well planned to ensure a good pace of learning. As a result of the pilot project set up to look at the teaching of phonics, nine experimental schools and nine control schools were selected (Chapters 9 and 11). The synthetic phonics programme *Jolly Phonics* was adapted for the pilot group. Success was proven through the literacy baseline assessment results. Children's ability in blending two and three sounds together and word reading improved

dramatically for the pilot group. Most other schools subsequently took on this approach. Only a few schools did not change to *Jolly Phonics*. However, they adjusted their programmes for phonics to increase the pace of learning of individual sounds. All schools and nurseries have been provided with resources to implement the early intervention strategies, including teachers' books, big books for whole class teaching and puppets. Nothing is left to chance, ensuring success for children and confidence in staff.

The project is based on the 10 key strands designed specifically for West Dunbartonshire (Chapter 8): the development of phonological awareness, a strong and structured phonic emphasis, extra classroom help in the early years, fostering a literacy environment within the home and the community as a support to school-based approaches, increased teacher awareness through focused assessment and target setting, increased time spent on key aspects of reading, identification of and support for children who are failing, accelerated and interactive learning strategies using multi-sensory approaches, home support for encouraging literacy, and developing positive attitudes and values towards literacy and building self-esteem

Resources to support literacy and numeracy are purchased regularly for schools and pre-five establishments - for example, big books, literacy programmes such as *Learning To Listen*, *Oxford Reading Tree Rhyme and Analogy* and *North Lanarkshire Writing*. All schools have adopted a structured approach to the teaching of phonics, in almost all cases adopting synthetic phonics methods. The pace of learning has changed, as has the approach to learning and teaching. Where schools taught one sound per week this has increased to three and children are using their sound knowledge immediately to put sounds together to make words, read words and write words. A multi-sensory approach has been adopted and has been very successful with the whole range of learners. *Jolly Phonics* programmes for P1 and P2 teachers are distributed for each teacher each session. Training and resources are provided for new schools. Staff development for all staff is crucial to the success of the programme. Expectations of all pupils are extremely high. The techniques of *Brain Gym*, accelerated learning and visual, auditory and kinaesthetic learning are integral to the teaching process. Teaching methodology is based on current research using the best available methods in literacy and numeracy. A highly-skilled team has ensured that school staff can benefit from modelling of new ideas, research into practice and use of new resources and techniques.

Support for pupils

The early intervention team have been trained to use a number of strategies to promote learning and to increase self-esteem and confidence. For example, all staff have been trained to use *Brain Gym* and *Mind Maps*. Through the use of these strategies the aim is to raise pupil's awareness of self and personal responsibility for learning, and to set achievable targets. Staff use a multi-sensory approach for teaching and learning through accelerated learning, thinking skills and the theories of multiple intelligence. Positive declarations are also used and have been shown to increase children's confidence and self-esteem. Staff also use praise and rewards to

encourage children, to promote self-esteem and confidence and in particular to celebrate success in reading and writing. For example, pupils have been given opportunities to share stories with the wider group and have been rewarded as ‘Star Writers’.

Assessment and monitoring progress

Attainment in literacy and numeracy is measured through the baseline assessment scheme designed for the authority. Literacy baseline assessment was introduced in November 1997. This has continued on an annual basis for each child in the pre-school year and for all pupils in Primaries 1 and 2. The assessment has four sections: Concepts of Print, Phonological Awareness, Early Reading Skills and Developmental Tasks. The content is based on research and how literacy skills develop. Between 3,000 and 4,000 children are assessed individually each year (Chapter 8). Content of the test is not taught; however, the skills are. For example, children are exposed to a wide range of texts, both fiction and non-fiction, and hearing and sharing books with adults on a daily basis. Phonological awareness is developed through hearing and reciting nursery rhymes, and specific programmes are used to teach skills in rhyme and phonological awareness (*Learning to Listen* and *Oxford Reading Tree Rhyme and Analogy*). This is based on research which has shown that the foundation skills for later reading success are children’s concepts about print and ability to hear and identify rhyme (Chapter 6).

The baseline assessment is administered by school staff and the early intervention team to individual pupils and has a very high ceiling and low floor. For example, most children in the nursery will know how to hold a book but they may not be able to read the word reading test. However Primary 2 pupils will be able to carry out most of the tasks and ultimately read some of the words in the word reading test. Children at all stages have made good progress and there have been significant gains made in each section of the assessment (Chapter 9). The baseline assessment results are an excellent indicator of how well children are doing in their literacy learning in West Dunbartonshire. The results of the assessments in literacy and numeracy are passed on by the pre-five sector to their associated primary schools each year as part of the arrangements for transition of information. This ensures that schools can track children’s progress from an early stage.

As well as this the Norman France Reading Test has been administered annually to pupils in Primaries 3, 4 and 7. This is a reading comprehension group test for each class. The results are used to track pupil progress and to ensure that the gains made at the early stages are maintained. The Norman France allows the tracking of individual pupils’ reading attainment through P3, P4 and P7. Schools pass on the P7 results to their associated secondary schools. Schools and the area network support team administer the test annually during May.

Numeracy baseline assessment, using the scheme designed for West Dunbartonshire, takes place during June each year, and as from 2003 is administered in pre-five establishments only. Until then Primaries 1 and 2 were also assessed. However, the rigour of the national test in mathematics has overtaken the need for numeracy

baseline in the primary stages. The results from the literacy and numeracy baseline assessment provide pre-five establishments and primary schools with information about each pupil's progress, enabling them to identify next steps, set targets and intervene early where problems are identified.

Failing readers in the later primary stages

During Session 2002-03 failing readers at P6 and P7 were targeted using the criteria of being at or below Level A in reading in the national tests. A principal teacher for learning support was seconded one day per week. A package of training materials, training and support for schools and tutors was made available throughout the session. A standardised approach was taken and all staff, the network team, the early intervention team, volunteers and parents who acted as tutors were trained in the use of *Toe By Toe* (Chapter 15). Although many schools had already been using *Toe By Toe*, they tended to dip in and out. This approach to training standardised the way in which *Toe By Toe* was delivered to pupils and therefore ensured success. Thirty-two primary schools took part in this initiative – 104 pupils. Children were also supported through special events to raise their confidence and boost their self-esteem. For example, all children took part in *My Amazing Brain*. Pupils were also given a certificate from the author of *Toe By Toe* to celebrate their success. Parents were supported through a series of meetings and training to enable them to support their own individual children. This was crucial in that children needed to feel that their efforts were valued and also support was necessary during the summer holiday period. The *Toe By Toe* initiative has achieved excellent results. Of the 104 pupils who took part in the reading intervention during 2002-03 most increased their reading age significantly.

Support has continued for these pupils with the transfer to secondary school. Nothing was left to chance and arrangements were made to ensure that these pupils would receive the same level of support at secondary as they did at primary. Secondary learning support staff were given the pre- and post-test reading results and have carried on the *Toe By Toe* programme with the identified children. Some children's reading ages increased to a level where they required a comprehension reading programme. *Stride Ahead* was used for these children. Training in *Stride Ahead* and support for the secondary schools has continued with both reading programmes to ensure a standard approach and ultimately success for the pupils involved.

Failing readers will continue to be identified. During Session 2003-04 *Stride Ahead* was also used in the primary sector. This is in response to schools' concerns about children who can read text but who cannot understand what they read. A greater number of children have been identified as failing readers – approximately 200. However, some of these children have reading ages above eight years, which is a suitable reading age for *Stride Ahead*. The same rigour has been applied to the monitoring of this intervention. The same process of pre- and post testing ensures that each pupil's progress is monitored to provide continuity and progression.

Staff development

The continued investment in staff development has been crucial to the success of the programme, not only for the early intervention team but also for staff from pre-five through to S1, and for special and support services staff. All staff have benefited from the best available training, based on the delivery of the 10 key strands. The most effective methods established through research findings have been incorporated and are being carried out by all staff at the early stages.

Schools and pre-five centres are supported through a planned programme of training and development, including *Interactive Maths*, *Brain Gym*, *Creating A Reading Environment*, *Learning To Listen*, *A Closer Look at Reading* and *Jolly Phonics*. Joint initiatives are extremely important and tremendous success has been achieved with *Total Communication*, a programme that aims to improve the communication skills of staff and provide them with a tool for assessing communication in young children, and *Hanen for Early Educators*. This is a joint initiative with speech and language therapy and psychological services.

Training is provided and aimed at all staff and not just a representative from a school or pre-five centre. This has included literacy and numeracy. Staff's knowledge of research, teaching methodology, cognition, thinking skills and accelerated learning is constantly revisited and updated. This has ensured that language and numeracy coordinators, head teachers, individual teachers and nursery nurses have the necessary skills and knowledge of current research to implement the initiative effectively. The early intervention team have also received training to develop their skills.

Training continues to be provided in *Brain Gym* for all pre-five staff and all P1 teachers. This helps staff understand how the brain works, and the importance of movement, hydration and motivation for learning. Staff in schools and nurseries have valued the *Brain Gym* techniques and have a clearer understanding of learning styles. Water is available in all schools and pre-five centres for pupils and *Brain Gym* is used as part of the daily programme for learning and teaching. *Hanen for Early Educators* is also part of a rolling programme for pre-five staff. This course has had considerable impact on practice and has provided staff with the knowledge and skills to work effectively with individual children, small groups and large groups. The course content covers the developmental stages of language and looks very closely at how to communicate effectively. For example, simple techniques are taught, such as the OWL which, is observe, wait and listen. This delay in waiting for an answer provides children with crucial thinking time.

The early intervention team have also been given opportunities to take part in a wide range of courses to develop professionally as well as contributing to the work of the Literacy Initiative – for example, multiple intelligences, how young children learn, reading, writing, accelerated learning, thinking skills and numeracy. Staff supporting pupils from P6 to S1 with the reading programmes *Toe By Toe* and *Stride Ahead* have also received training in both programmes.

Training will remain a priority with programmes of staff development planned for each new session. New ideas will continue to be explored ensuring all teachers benefit from new research, methodology and approaches to the teaching of reading and writing. The rolling programmes in *Brain Gym*, *Jolly Phonics*, literacy, *Learning to Listen*, *Hanen*, *Total Communication*, *Toe By Toe*, *Stride Ahead*, literacy baseline assessment and numeracy will continue each session as part of a maintenance agenda. These training programmes are crucial to the success of the strategy and form the foundation for good practice. The head teacher will also continue to provide for individual groups of staff as required.

Ethos

The early intervention team are based at Dalreoch Primary School. The original team of 10 teachers and head teacher is now a head teacher, 17 teachers, a deputy head teacher and a principal teacher. The team have worked hard to develop a very good ethos and the Literacy Initiative is well known and respected throughout the authority. The atmosphere is friendly and welcoming. This is characteristic of the relationship of staff with their allocated schools and nurseries and with the children and staff they work with. Good communication is a priority.

Staff meet regularly through their planned collegiate time as well as for staff development at the end of June, at the beginning of August and after the literacy baseline in December. This does not include in-service days throughout the school year and other opportunities as and when they arise. This has made a positive impact on team building and the professionalism of the team.

Early intervention has very good links with the primary and pre-five sector as well as with the wider community. The project supports all schools and pre-five centres in their partnership with parents and provides parent information meetings and workshops as requested. Parent information meetings are requested frequently. Workshops have been developed in phonological awareness, phonics, sharing a book, reading, paired reading, emergent writing, *Toe By Toe*, *Stride Ahead*, numeracy, *Playalong Maths* and reading for information (aimed at boys).

Schools and pre-five centres receive regular information about early intervention and the wider initiative through reports, letters and invitations to staff development conferences as well as the distribution of resources to support literacy and numeracy learning. There is an authority policy for early intervention, and the steering group for the Literacy Initiative take an active role in the work of the early intervention project. The head teacher actively seeks the opinion of heads of establishments. This helps to formulate future policy, deployment of staff and staff development for authority schools and pre-five centres.

Partnership with parents

A crucial part of the project is to support the role of parents. Sixteen primary schools receive extra help for children who have little support at home, providing a link for family learning. Children who are absent for long periods, poor attendees and those

children who have little or no support at home also receive additional support. Pre-five centres are supported in their work with parents. A range of parent information packages and leaflets is available for all sectors of education as well as partner agencies in health and social work. Family literacy and numeracy packages have been developed to support home learning. For example, issues such as gender have been addressed. The package for paired reading has been developed to include a wide range of non-fiction books taking account of the interest of boys. Starting School Bags have been produced for every new child in Primary 1. This has been a major achievement for early intervention and for West Dunbartonshire. Through this bag the authority has been able to target all new parents and children and introduce them to the Literacy Initiative. Every child starting school has been given the opportunity of an equal start on the road to lifelong learning. The contents have been specifically designed for families to learn together. For example, there are parent information leaflets on every aspect of early literacy development – phonological awareness, sharing books and emergent writing, as well as a game, alphabet mat, a book, paper and a pencil. A survey of parents provided very positive feedback on the Starting School Bag. Annually approximately 1,200 children receive the bag when they start school. Schools have reported that the bag is a practical resource for parents to prepare their children for school and is useful in that all children each receive an identical bag.

In summary, the West Dunbartonshire Literacy Initiative has been highly successful to date. Very good publicity continues to be received for its achievements, with recognition at local, national and international levels.

SUMMARY

This chapter provides a practical account of the West Dunbartonshire Literacy Initiative in operational terms under the headings of support for pupils, assessment and monitoring progress, failing readers in the later primary stages, staff development, ethos and partnership with parents. It outlines how the project was organised and resourced, and the steps taken by key personnel to ensure that as a large-scale and complex logistical exercise it was effectively managed to ensure successful outcomes.

Chapter 17

Conclusions

In terms of the breadth of its vision, the scale of its implementation, the extent of its long-term gathering of follow-up data and the distinctiveness of its novel interventions, it is proposed that this study had made an original contribution to research in this field, and that it has significant implications for educational policy and practice. This concluding chapter examines these implications, as well as summarising the main findings of the study and providing a critique of its methodology.

The design, implementation and evaluation of a literacy intervention

The research initiative, of which the study reported here was the major part, was designed to last 10 years. The present work has reported on the progress of the project through its first phase, covering a period of six years from the initial preparation, planning and design. The aim as outlined at the outset was to design, to implement and to evaluate the effects of, a multiple-component intervention to raise achievement and address illiteracy in areas of socio-economic disadvantage, taking full account of the factors affecting educational change in the context of real world research.

It is concluded that the study has been successful in the achievement of this aim. Not only has a complex, large-scale intervention in literacy been designed from its first conception to the point of comprehensive implementation and evaluation with a whole population of children and young people, but it has also been rewarded with highly encouraging outcomes. These outcomes have led to significant raising of achievement for all groups of pupils from the most to the least able. In tackling the attainment levels of the latter, the intervention has laid a firm foundation for the achievement of the ambitious agenda it has set for the second phase – the eradication of illiteracy throughout the school population by the end of the 10th year of the programme.

This study was designed on a grand scale. The main study alone involved assessments of 43,370 children, of whom 22,986 were assessed individually, the remaining 20,384 being group assessments. It also involved 58 educational establishments, of which 23 were nurseries and 35 were primaries. Across these schools over 400 staff had to be trained, monitored and supported to ensure high fidelity of programme delivery.

Implementation involved key personnel working constantly and systematically through the years with every level of education management, from educational directorate, through quality insurance structures and primary and secondary head teachers and heads of pre-five establishments. It also required extensive contact with hundreds of workers at ground level – class teachers, classroom assistants, learning

support teachers, SEN auxiliaries and volunteers, together with collaboration with the psychological service and other services contributing to the overall programme. In addition, it involved planning and recommending the key support structures required to carry out the project effectively, as well as the organisational tasks involved in the appointment of early intervention teachers and other key workers.

To evaluate the main study a full baseline assessment scheme was designed and produced. This involved consideration of a complex national and local context regarding baseline assessment in schools (Scottish Office Education and Industry Department, 1998), to take account of which a survey was conducted across all 32 education authorities in Scotland (MacKay, 1999a). To support the baseline assessment, hundreds of teachers and other staff had to be trained in assessment methods, and this had to be repeated every year for new cohorts of workers. The subsidiary studies required setting up and carrying out training programmes for hundreds of additional staff, together with designing and carrying out further individual assessments for hundreds of additional children.

Carrying out all of the above requirements was a considerable logistical exercise. It necessitated the coordination of the work of staff at many different levels, across a number of services and departments and across a wide range of locations, as well as the establishment of training arrangements and support structures. Ensuring that the mechanisms and resources were in place at the right time on every occasion baseline assessment or other assessments were required also involved complex logistics. The resultant data represented an immense mountain of test papers that required to be scored and then processed, and it is estimated that for the main study around two tons of assessment papers were generated in the course of this exercise.

Key findings

The main study

This study was the centre-piece of the research, and the four subsidiary studies served the purpose of informing its development and supporting different strands in the multiple-component intervention. It was the study in which the vast preponderance of the available funding and other resources were invested. The findings relating to this study were therefore of the most crucial importance to the overall aims of the initiative.

A higher level of achievement in literacy was demonstrated year-on-year throughout the six-year intervention period when each cohort was compared with the population that served as its control, that is, the cohort at the same age level in the previous year. Enhanced results were obtained across every age group, pre-school, Primary 1 and Primary 2, across every baseline test and across every level of achievement. The proportion of pupils obtaining high scores rose significantly, while at the other end of the scale those with very low scores reduced dramatically in numbers. Indeed, by 2003 only a handful of pupils were obtaining scores that could be described as very low.

Two years had been selected to illustrate the effects of the intervention – the first year (1998) and the last year (2003) of the first phase of the initiative. While significant and extensive improvements in literacy had been achieved for the first of these years, with 34 significant increases out of 36 measured scores, and an average effect size of 0.45, not every test had shown an increase. In terms of effect sizes, five were very small (or in one case negative, -0.01). However, in the last year all increases were significant, the lowest effect size was 0.41 and the average effect size was 1.11. This suggested that as the programme became established in the delivery of the reading curriculum, and as the effects of training accumulated, overall effectiveness increased.

The group reading tests at Primary 3 and Primary 4, although very much less sensitive than the baseline assessments in their scoring, indicated that the gains made as a result of early intervention were being reflected in reading scores in the years following the intervention. The effect sizes at this stage were smaller. Nevertheless, the practical effect of these improvements was very apparent. Indeed, the class teachers in Primary 3 noted that the new reading levels of children entering their classes since the programme began were challenging the delivery of the normal P3 curriculum, which required to be re-appraised.

There was already evidence of the potential benefits of carrying out literacy interventions in areas of socio-economic disadvantage using a multiple-component strategy in which the components individually were drawn from research evidence. The Pilton Project had developed from the work of McMillan, Fox and Wood (1994) and had good follow-up evaluations prior to the commencement of this study (Lothian Regional Council, 1995). In addition, the preparatory study conducted by MacKay and Watson (1996) was of particular relevance, as it had introduced the early elements of the baseline assessment scheme developed for this study, had implemented some similar components of intervention and had shown an awareness of the importance of context as well as content in designing interventions. These studies had targeted areas of socio-economic disadvantage in Scotland. However, there were also crucial differences. Both were on a much smaller scale, and they did not involve the long-term considerations of accounting for the processes of educational and organisational change across whole Council areas.

The synthetic phonics study

Each of the three hypotheses for this subsidiary study was supported by the results: first, that after intervention experimentals would have higher scores than controls in the key areas addressed by the programme, namely, knowledge of letter sounds and blending skills, and that this would be reflected in word reading abilities; second, that they would score higher on tests of spelling; and third, that these gains would be reflected in the longer term in the mid-primary school years.

Assessments carried out at about the mid-point of the first year of the intervention showed superiority of experimentals over controls in the sounding, blending and word-reading skills assessed, the best result being for blending of nonsense words on the non-word reading test (effect size 0.75). Assessments at the end of the session on

the spelling test designed for the study showed higher scores for the experimentals. Although differences between the samples were not found in mean scores on the group reading tests in Primary 3, they were found the following year in Primary 4 (effect size 0.46). In addition, in Primary 4 the quartile distribution of scores showed an interesting pattern of results. The distribution of the group test scores of the controls did not differ significantly from distribution of baseline assessment scores for these children four years earlier while they were in nursery. However, the scores of the experimentals had shown a significant shift upwards, with fewer being in the lower quartiles.

Extensive qualitative feedback from teachers during the first year of the study supported the view that the synthetic phonics method was an effective means of raising achievement in their classes. As the first year was associated with all of the excitement that a new and very different approach can generate, structured teacher feedback was again obtained five years later through questionnaires. These again provided a high degree of support for the effectiveness of the approach. By this time the dissemination of the findings of the study among schools, and the enthusiasm of the teachers using synthetic phonics, had resulted in virtually every primary school in the authority opting for this method.

This study made a major contribution to a key strand of the overall intervention, namely, a strong and structured phonics emphasis. It supported the results obtained by Johnston and colleagues in using synthetic phonics in Scotland (Johnston et al., 1995; Johnston & Watson, 1997; Watson & Johnston, 1998), and elsewhere by Stuart (1998, 2004) and by Willows and her colleagues (Kwan & Willows, 1998; Stornelli & Willows, 1998; Sumbler & Willows, 1996). There were, however, some differences between these studies and the current one. While the other Scottish studies had included a focus on socio-economic disadvantage, this was not at the levels applicable to the sample used here, where disadvantage was very marked. The studies in the UK by Stuart focused on children for whom English was a second language, and the studies in Canada by Willows and colleagues also had high numbers in this category and reported limited socio-economic data. Importantly also, none of the other studies was conducted in a situation where literacy was already being addressed through a wide range of other strategies, including support in all control schools for effective and structured phonics teaching.

The attitudes study

The attitudes study indicated that over five years after a brief intervention to raise literacy scores by changing attitudes and values, the experimentals were still reading at a significantly higher level than the controls, even though they had received no differential treatment during these intervening years. This applied to reading accuracy, with an average reading age for experimentals of 10y 11m compared with 9y 3m for controls (effect size 0.82) and to reading comprehension, where the figures were 12y 3m for experimentals and 10y 6m for controls (effect size 0.70). No differences were found at that stage in responses to an attitude questionnaire.

Again, this study was of importance in planning the main long-term intervention with a view to intergenerational change. It reinforced the importance of an emphasis on attitudes and values in supporting the reading curriculum. It is recognised that the relationship between reading achievement and attitudes, and the strength of that relationship, is complex. However, this study and the earlier randomised control trial on which it was based (MacKay, 1995) supported the view that, irrespective of questions of which factor is antecedent (attitudes or achievement), attitudes and self-concept can predict later reading achievement.

The declaration study

The declaration study resulted in significantly higher scores on early literacy skills for the experimentals. For the nursery group, the experimentals showed significantly greater gains than the controls on three tests: nursery rhymes (size 0.51), rhyme production (effect size 0.98) and lower case letter sounds (effect size 1.51). For the primary group, the experimentals showed significantly greater gains than the controls on five tests: nursery rhymes (effect size 0.44), lower case letter sounds (effect size 0.21), letter names (effect size 0.64), non-word reading (effect size 0.60) and word reading (effect size 0.64). The experimentals also showed significant shifts towards more positive attitudes and expectations regarding reading. These findings were supported by structured feedback from both teachers and participating children.

As a study of children in their pre-school year and Primary 1 the declaration study was particularly relevant to informing a large-scale early intervention. The approach was novel and universally regarded as enjoyable by both children and teachers, and by making declarations in groups and at whole class level children were able to participate in a further activity involving social and interactive learning.

The individual support study

The individual support study was of central importance to the overall intervention as it reflected one of the two key aims of the initiative, that of addressing illiteracy in the pupils who were failing, as well as raising general reading achievement in the population. It supported one of the key strands, namely, identification of and support for children who are failing.

Both of the hypotheses for this study were supported by the findings: first, that the experimental pupils in the secondary study would achieve higher reading scores than the controls; second, that the children on the programme in the primaries would show large gain scores. The 12 experimental pupils in the secondary study showed mean reading age gains of 2y 0m (from 8y 2m to 10y 2m) following the three-month Toe By Toe intervention, and with a 12-month interval between tests, while the controls gained only four months (from 8y 5m to 8y 9m) during the same period (effect size 1.74). The average pre-test reading age of the 104 children in the primary schools study was 8y 0m, this being about three years behind their chronological age. After a period of just under six months their post-test reading age had risen to 9y 2m, giving an average gain score of 1y 2m. Using the concept of 'ratio gains' as defined by Topping and Lindsay (1992) – the gain in reading age made...on a reading test

during a chronological time span, expressed as a ratio of that time span – the mean ratio gain of 1y 2m in approximately six months for the primary sample was about 2.3. Since these pupils had had years of school experience in which the amount of progress they made was very much lower than the amount of time passing, the change in rate is a very encouraging one.

These findings are in accord with studies that point to the superiority of individual tuition over group tuition for failing readers (Bloom, 1984; Iversen & Tunmer, 1996). In addition to the expectations created by such studies that a carefully-planned individual support programme will lead to good outcomes, the results of this study were anticipated in the pilot work carried out in preparation for it, during which very high gain scores were reported for pupils receiving this specific intervention. This provided a firm basis for setting up the training and implementation arrangements by which the study was extended from secondary school throughout the upper classes in the primary sector.

Limitations

Any critique of this study must recognise a whole range of limitations. Some of these are perhaps inherent in a study of this kind and might almost be predicted by Robson's (1993) observation that 'real world research' tends to be done in 'complex, messy, poorly controlled "field" settings' (p. x). His assessment of real world research as mainly solving problems rather than just gaining knowledge, as predicting effects rather than finding causes, as looking for large effects rather than studying relationships between variables, as developing and testing interventions and services rather than theories and as using multiple methods rather than single methods, all resonates with the nature of this study.

It was observed at the outset of this work that as sample size increases from small numbers in a single school to whole populations in many diverse establishments, as co-workers increase from one or two dedicated research assistants to vast numbers of teachers and other staff, and as intervention methods increase from single to multiple-component strategies, so also is there a corresponding increase in the complexity, messiness and poor control that characterise these settings. It is acknowledged that these features were often very prominent in this investigation, and they may provide a context in which at least some of the limitations in the study may be considered. This section notes a range of limitations and weaknesses or possible weaknesses. Others of a more specific nature have already been noted within the studies to which they refer.

First, any study that claims to be driven by a commitment to values in science and the aspiration to 'do good', that is, to seek to apply psychology to human welfare, will almost certainly, among other possible criticisms, be vulnerable to the censure of those who have a different view of what is 'good'. A key aim of this study was to raise literacy levels among disadvantaged children – indeed, to change their own attitudes and values in the process, and in particular to ensure that all would attain 'functional literacy'. This philosophy might be (and indeed has been) challenged, and is open to Levine's (1986) criticism that this merely serves to increase conforming

behaviour in the participants and bring them yet more within bureaucratic communication and authority, with their literacy at just an adequate level to be 'functional' in increasing their usefulness and subservience to society rather than functional to themselves. The alternative view (for example, Freire, 1994, 2000) is that literacy provides empowerment, increases choices and improves quality of life. This is the view that is taken in this study.

A second general weakness is methodological. Throughout the study the researcher's ideal of double-blind randomised control trials was left far behind for designs which, even though they usually reached quasi-experimental level, were far from being blind as far as the people conducting the assessments were concerned, and furthermore, at times the assessor was also the researcher. This is frequently again a feature in real world research in schools, and often is a reflection of resources. For example, in the declaration study it was expected that the assessments would be undertaken by others, but when the time came the resources for this were not available. The researcher therefore reluctantly had to carry out the unplanned task of undertaking 120 individual baseline assessments.

Possible vested interests of the assessor in the assessment results were perhaps less of an issue in the vast samples that constituted the main study. Although most of the assessments were carried out by the class teachers and nursery staff themselves there were several factors that mitigated or monitored assessor effects. While undoubtedly the teachers would have had a wish for their own children to do well, the same wish was likely to have been present at pre-test baseline. The differences, however, between pre-test and post-test situations were generally large. Also, as the tests were conducted in November/December each year, the teachers were not assessing situations that reflected only their own work, but also that of the teacher or establishment responsible in the previous session.

Monitoring took place at a high level. Assessment papers were screened for any apparent anomalies that might cast doubt on the validity of the results, using many sources of enquiry such as patterns of test performance, knowledge of children's previous performance levels and the detailed knowledge of the levels of the children in each school held by the early intervention teachers. If such factors gave any cause for concern, and also at many times for other reasons, other staff such as members of the early intervention team undertook half or all of the assessments in a particular class. In addition, parallel forms of key baseline tests were designed and were used in a range of establishments without prior warning, not just where there seemed cause for concern but more generally. Throughout the six years of the first phase of the study only two schools caused concerns that they might not be using the tests appropriately. These situations were investigated thoroughly, and as a result one school was advised of its practice.

Third, despite the vastness of the sample used in the main study, a number of issues may be raised about sample size and characteristics. For example, in the attitudes study, although all but four of the original sample were traced in their secondary schools five and a half years after the original study, the resultant number (after removing one pupil because of significant issues of 'caseness') was only 19, and 11

of these were experimentals. Certainly, the original study, despite its small sample (dictated by the constraints of a single establishment) was a randomised control trial involving three groups, and the sheer size of the experimental v. control differences provided robust results. The differences at follow-up were also large, but the small numbers did constitute an obvious weakness.

Both sample size and sample characteristics were a limitation in the declaration study. The total number of participants was high at 565, but the core data available for the main analysis was 60, reduced to 54 at post-test because of changes of school or being absent at every opportunity for assessment. This meant that the sample was then reduced to 27 experimentals and 27 controls. This somewhat limited the range of analyses available. Of these 27, 18 were primary, leaving only 9 per group at nursery level (reflecting the smaller number of nurseries compared with primaries). The results pointed to real change in favour of the experimentals, supported by a wide range of other more qualitative data. However, the number of possible breakdowns that would have been of interest with this sample – by test, by socio-economic status of establishment, by high v. low achievers, by gender – was clearly very limited.

Sample characteristics proved to be an unexpected and indeed irksome feature in the declaration study. Analysis of pre-test scores in the group who were individually tested pointed to systematic differences in baseline performance in favour of the controls (that is, the controls began with higher scores). Without calibration from other sources, this might have appeared as an accidental or deliberate assessor effect (the assessor being the researcher, as noted above). That is, the assessor might have preferred the scores of the experimentals to be depressed to make later change easier (although it would be obvious that unwelcome pre-test differences in the groups would then be apparent). While randomised allocation to experimental or control conditions would have controlled for this, the practical situation did not support it, as the experimental schools had to be agreed with the education authority from the start. The possible assessor bias was calibrated when assessments provided by the schools showed the same pre-test differences. Despite a good systematic sampling procedure, it is likely that, other than totally random variation, the control schools influenced the ultimate sample by not wanting to have poor scorers as their representatives, while the experimental schools wanted more needy children to have the pre-post intervention assessments. In the event, the analysis took account of these differences and indicated that pupils across both groups who started with higher scores did not differ in their progress from those with lower starting scores.

Fourth, a number of issues may be raised in regard to the range of assessment methods used. The baseline assessment designed for the study had many strengths, and its usefulness was demonstrated by the fact that when education authorities were surveyed a number of them were using it in whole or in part. However, of the seven principles it was designed to meet – reliability, validity, utility, directness, reactivity, sensitivity and feasibility – an overwhelming one was feasibility. That is, it had to be possible for the education authority to fund its production and administration for thousands of individual assessments each year. For evaluation purposes it was also desirable to design a test that could be used across the full age range of the main

study, from the pre-school year to Primary 2. One of the outcomes of this was that some tests were limited in usefulness for the lower age range, having too high a floor (such as word reading), and others for the upper age range because they had too low a ceiling (such as concepts of print). This meant that floor and ceiling effects were often apparent. Nor was it possible to follow a simple expedient of combining scores into a grand total, an exercise with possible advantages but raising further questions about adding together some quite disparate tests in terms of weight and balance, as well as losing data specific to individual tests.

The Norman France Reading Tests presented an unavoidable gap in assessment measures between the stages up to P2 and the stages from P3 onwards. Their scores did not have the sensitivity or detail of the baseline tests, and they clearly had a high loading on attentional and cognitive factors, as well as being difficult to monitor effectively to ensure independent work when undertaken in groups.

Fifth, potential weaknesses in the assessment measures were reflected in the analyses carried out, and these have been discussed in relation to individual studies, and most particularly in relation to data skewed by floor and ceiling effects. The unsophisticated level of analysis reflected Robson's (1993) observation cited above regarding real world research often looking for large effects rather than studying the relationship among variables, and the vast body of data generated has not as yet been exploited in terms of the range and sophistication of the analyses that might illuminate aspects of the study more clearly. Certainly, many of the effects were indeed large, consistent and self-evident, but not all of them were, and these are the ones that might particularly benefit from further appraisal. Also, the advantages and disadvantages of using pre-intervention population cohorts as controls for post-intervention cohorts at the same stage rather than establishing 'true' control groups have been recognised. The issue must be raised as to whether the subsequent year groups might have had better scores without the intervention. This question is of great importance in relation to the main study, and is covered separately below.

Sixth, a further question that may be raised about the project is not so much whether the results rose, but whether the skills measured represented a balanced and meaningful range of literacy abilities. The rationale for selecting the items for the baseline assessment in terms of their usefulness and predictive validity is covered in detail in Chapter 7. At the same time it may be noted that the assessments were limited to the very mechanical skills of sounding, blending and word recognition that can be easily assessed in the early stages rather than the broader area of reading comprehension. However, apart from the inherent difficulties in assessing comprehension accurately in ways that separate it from broader cognitive abilities in these early years, it is recognised that skill in the mechanical measures assessed is critical to the development of higher order reading abilities (Chapter 6). When a child has not developed mechanical fluency, but has to spend time deciphering words and their phonic elements, comprehension is impeded. All who are familiar with administering tests covering both accuracy and comprehension, such as the Neale Analysis of Reading Ability, will be very familiar with the issues arising here. Mechanical skills were therefore the focus of baseline assessment for the early stages so that a solid foundation would be laid to develop higher order skills later.

The final limitation noted here again reflects Robson's observations about studies in the 'real world' looking for large effects rather than studying relationships between variables, developing and testing interventions and services rather than theories and using multiple methods rather than single methods. While this study has achieved its key aim of designing, implementing and evaluating this large-scale intervention, and demonstrating that it has raised achievement and reduced illiteracy, it leaves a number of unanswered questions regarding which variables had what effects. The whole sample had the whole intervention. The weight of each strand or component in the intervention may have differed between establishments, and the very objective of maximising 'ownership' of the project in each school tends to highlight differences in approach, but the effects of each component cannot be separately assessed. Might some components have been omitted without weakening the strategy? Might most of the variance have been accounted for by one or two components, such as more time spent on reading? What was the contribution of the context variables, involving such intangible concepts as 'vision' and 'profile'? For many reasons the simplest response to these questions is that they must represent the subject matter of another study. For the present study there was a different ambition that transcended these more detailed questions, namely, to lay the foundation for intergenerational change in the achievement levels of a disadvantaged population, and to seek to maximise this in every way possible. Many strands were interwoven to produce a successful and sustained outcome. The end result was that achievement levels rose, not just statistically, but in a meaningful way that can be crystallised in the experience of a rising generation of successful readers and their families.

Pre-intervention cohorts as post-intervention controls

The issue that affects all year-on-year cohort studies where the baseline results for the first year of a study represent the standard by which future cohorts at the same stage are evaluated is the question as to whether standards might not have been rising anyway from one year to the next. Since this question is of fundamental importance to the main study it is treated here in detail. Four comments may be made. First, the Assessment of Achievement Programme was established by the Scottish Office Education and Industry Department in 1981 to monitor the performance of pupils in Scottish schools in particular areas of the curriculum including English language, on a three-year cycle for each subject. The stages assessed were P4, P7 and S2. Although the results do not allow precise comparisons with the assessments reported here in terms either of age group or of content, the overall drift of findings does not suggest a tendency towards systematic rises in reading attainment from year to year. For example, for pupils in Primary 4, the period from 1984 through to 1995 showed at times a slight fall and at other times no change in reading attainment. Despite gains reported in relation to a number of early intervention programmes (Fraser, MacDougall, Pirrie & Croxford, 2001), reading standards at all stages showed a marked fall through the main part of the period reported in this study, from 1998 to 2001 (Scottish Office Education and Industry Department, 1996; Scottish Executive Education Department, 2003). From that perspective there would be no indications that a rise in results might have been expected at that period.

Second, no evidence is available to suggest that standards were rising in core educational attainments in general during these years in the area in which the main study was carried out. While the project was taking place a much less intensive intervention on raising attainment in numeracy was being designed, implemented and evaluated, including the design of a numeracy baseline assessment for the purpose. Only modest gains were reported in numeracy throughout that period, more or less proportional to the size of the intervention, which received much less resourcing than the literacy intervention. Therefore there is no indication that significant improvements in literacy were simply part of a general rise in educational standards over these years.

Third, the changes in the baseline assessment results did not reflect minor variation but showed large effect sizes. As reported for the main study, pupils scoring on the various baseline tests at the 50th percentile in 1997 before the intervention began would by the standards of the 2003 baseline have been described no longer as average but as low scorers. A score at the 50th percentile in 1997 for lower case letter sounds in Primary 1 would have been at the 4th percentile in 2003. In short, the results are consistent with the expectations of a highly successful intervention programme.

Fourth, at a less formal level, the consistent reports of class teachers, specialist teachers and many others associated with the programme overwhelmingly confirmed their view that the intervention had brought about manifest changes in the levels of attainment of the pupils who were participating at all levels of ability.

The process of educational change

The study has been guided throughout by the belief that ‘change is a process, not an event’ (Fullan, 2001, p. 52). Full recognition has been given to Fullan’s (2001) three phases of educational change – initiation, implementation and institutionalisation – together with a commitment to the view that years of systematic collaborative working are required to achieve a successful outcome of these phases.

The preparatory studies (Chapter 5) not only provided a good foundation for planning a large-scale intervention in literacy: they also contributed a crucial awareness of the inevitable and seemingly incessant factors that can undermine real world research programmes and the changes they seek to effect. An outline has been given earlier (Chapter 2) of the immense and diverse difficulties that arose in a short-term, small-scale study involving only two establishments. It was expected that such factors would be multiplied in a study involving hundreds of teachers and other workers, many thousands of children and scores of establishments over a period of years.

Indeed, this expectation was fulfilled. The difficulties were manifold, and the following are far from exhaustive. First, there were funding issues. The project required very considerable finances, including funds for the employment of a team of specialist teachers. The bulk of the resourcing was obtained from the Scottish Executive Education Department for the part of the project that related to early

intervention – essentially the main study, and overwhelmingly the most costly element. This was available initially for the first three years, and was then obtained for a further two years. At that stage the Council had to make major decisions about mainlining project staff, and also had to work in a different funding context. At times these funding concerns affected the longer-term planning process, and led also to the next difficulty as noted below.

The second major difficulty encountered was in the realm of personnel. Issues over funding and related areas meant that specialist staff were initially seconded or employed on temporary contracts for a fixed term of up to three years. This meant that in the third of these years staff had no guarantee of a future with the project, and understandably many felt they had to look either for permanent posts elsewhere or for a return to the post from which they had been seconded. For a period this resulted in many schools having very limited support. Crucially this affected the position of the person appointed as head teacher for early intervention, who returned to a mainstream head teaching post. She was central to programme implementation and thus a major discontinuity could not be avoided for a period. In addition, other contractual issues arose. For example, the teacher seconded for the implementation of the individual support study was largely recalled from her secondment because of staffing difficulties in her establishment, bringing about significant problems for a period in monitoring and supporting the study.

Third, towards the end of the first phase of the research a protracted period of industrial action affected the nursery nurses in the pre-school sector. This resulted in the implementation of the programme being disrupted to a greater or lesser extent in the 23 pre-five establishments. It affected training, the availability of staff to carry out baseline assessments and the overall readiness of staff to implement any initiative they felt was additional to their established duties.

Finally, an example of the issues that can be much more easily controlled in tidier research projects than in the messy arena of such a large-scale, multiple-component intervention is found in the synthetic phonics study, where the opportunity for continuing with a control sample was overtaken by the proselytising enthusiasm of the experimental schools. Efforts of researchers to stop controls from ‘doing it’ can prove in vain when a group of animated teachers persist in talking about and sharing their practices and materials.

The range of difficulties noted above could be expanded in almost every direction. The intervention had to continue and to succeed through virtually every major change or turmoil taking place in its midst – including a total restructuring of the educational directorate, together with significant changes in the Council. All of these challenges are fundamental to the process of achieving positive, long-term change in carrying out real world research on a large scale.

Implications for policy and practice

There is a sense in which the key implications of this study for educational policy and practice do not need to be spelt out, as they are axiomatic. A large number of

education authorities – in Scotland, in the UK and in other parts of the world – have populations marked by high levels of socio-economic disadvantage. There is a consistent body of evidence on the impact of this factor on health, quality of life, educational achievement in general and literacy levels in particular. Such communities are characterised by educational underachievement and high levels of illiteracy. These problems have not been shown to respond either to generally available educational strategies or to many special initiatives. The multiple-component approach described here is not another ‘package’ to be purchased, but a process for enhancing the effective delivery of the reading curriculum. Its long-term effectiveness has been demonstrated, and while it requires resources at significant levels, it is argued here that it is a necessary and cost-effective investment.

Cost-benefit analysis

Visionary aims of raising educational achievement and eradicating illiteracy cannot be pursued without reference to the economic costs involved in implementing effective programmes. Without question, the intervention reported in this study was costly. The target annual budget for the whole initiative, adjusted to reflect current values, was around £300,000. Translating this into cost-benefit terms is complex in terms both of defining the population that benefited directly from the intervention and of assessing the wider aims and impact of the project.

In terms of the population that received direct benefits, the main study included every child in the pre-school year, Primary 1 and Primary 2. This was an average of 3,221 pupils per year through the six years of intervention reported. If the costs of the initiative are related only to this population the cost per pupil would be £93 per year. However, it may be argued that this was an initiative for the whole education system, although with a primary focus on the groups specified in the main study. It involved inputs at many levels throughout nursery, primary and secondary schools. This was reflected not only in the application of the intervention to pupils in the upper primary and secondary stages, but also in more general effects of the initiative, such as enhancing the curriculum from Primary 3 onwards and in affecting other siblings in families of targeted children. Viewed as a whole education authority initiative, the project represented about £13 per pupil per year for pupils attending schools or nurseries in the authority. As a proportion of the education department budget, this represented under 0.5% of education spending.

In terms of the wider aims and impact of the project, it is argued that this was an initiative with potential to have significant effects on quality of life and the economy throughout the whole population. If any of its stated goals were going to be achieved – of higher self-esteem, lower disruption in schools, better school ethos, better staff morale, economic savings in remedial support, lower crime, a more skilled workforce and a stronger economy – then in cost-benefit terms the expenses of running the project represented a modest investment indeed.

As a footnote on the cost-benefit analysis, it may be noted that the area in which the main study was conducted – the second poorest Council area in Scotland – despite the financial strictures it was facing, made a decision to fund the entire project after

the specific research funding for it had been discontinued. This involved mainlining the posts for the head teacher of the project, the early intervention team and the home-school support teachers.

One of the issues that has been considered in terms of the policy and practice implications of this study is whether the baseline assessments should continue after they were no longer required for the research, as they too have considerable cost implications, not so much in production as in administration. It is a testimony to the usefulness of the scheme that the education authority opted to continue with the baseline assessments with or without the research project. They were found to be of great value in informing teachers of the progress of every individual pupil, in identifying children at risk and setting targets. They supported an essential strand in the multiple-component intervention – raising teacher awareness through focused assessment.

Among the individual components of the intervention, the synthetic phonics study has highlighted the benefits of a strong and structured phonics emphasis. The study indicated the superiority of the synthetic over the analytic or traditional approach, and the clearest policy recommendation would be for schools to adopt this approach. Although this recommendation could be with confidence of good outcomes, caution would still be associated with this area. If there is an aspect in which the synthetic phonics study, and all of the existing evaluations, have left a continuing question it is whether the synthetic approach is ultimately superior because of its distinctive synthetic methods, or whether it has not yet been sufficiently systematically compared with better analytic phonics teaching. Such teaching would involve not only a faster pace of teaching letter sounds but also a clearer focus on elements of blending instruction that were absent, for example, from the studies by Johnston and others cited above.

In addition to the need for a multiple-component intervention in literacy for the early years, the provision of intensive individual support for failing readers in the later stages of primary and into secondary school is essential. Again, individual tuition is costly – but it is not ultimately so costly as illiteracy. The methods used in this study have proved to be successful, but without the high level of costs associated with other programmes of demonstrated effectiveness. Clay's Reading Recovery programme (Clay, 1979b, 1991, 1993a) requires extensive training of teachers to ensure successful delivery, and Iversen and Tunmer's (1993) modified programme used teachers with a higher degree in reading. The individual support study reported here was based on a single brief training session for staff, many of whom were volunteers.

Finally, it is essential for education authorities to establish literacy initiatives that are long-term and not short-term, and that take full account of the processes of educational change and the loss of impetus that characterises many projects after the initial enthusiasm had waned. Although early intervention projects for literacy were established throughout Scotland in the late 1990s, many of these were small in scale, limited in scope and not enduring. Informal surveys by the author have highlighted several ambitious projects that have now terminated, some having simply faded out

with the passage of time, with personnel changes or with the transfer of central funding to individual schools.

In summary, for areas of socio-economic disadvantage this study supports the establishment of long-term literacy initiatives with multiple components, with intensive individual programmes for failing readers, with detailed assessment and evaluation measures and with high levels of training, monitoring and staff support.

Vision, profile, ownership, commitment and declaration

Throughout this report reference has been made to the recognition given to key context variables: that is, to the idea that major educational change must be supported not only by good programme content but by a much wider context marked by vision, profile, ownership, commitment and declaration. These were built into the project at every opportunity. They were formally articulated and discussed at the meetings of the steering group that met regularly to monitor and plan the progress of the project. They were highlighted as being of fundamental importance at major conferences held to keep the project on track and to celebrate its results from year to year. They were taught to staff and volunteers at every level during virtually every training session. The essential message was: this programme has 10 content variables – the ‘10 strands’ – and it has five context variables.

The result was that the project has been defined and recognised as being ‘visionary’. The language associated with it at all times has therefore been visionary language. The original research proposal on which it was based began with the words, ‘This is a vision for transforming reading standards for all children in all schools throughout the education authority’. It was presented in the highest profile at every opportunity. On over 30 occasions it was seen in newspaper headlines or in magazine articles, and on several occasions it was featured on radio and television. The message to all who participated was that they were involved in something very important. Vision and profile promoted ownership and commitment. The project belonged wholly to everyone – it belonged to the Council, the directorate, to the early intervention team, to the head teachers, to the class teachers, assistants and volunteers, to the parents, to the children and also to the researcher. Commitment began at Councillor level. Few conferences or media reports lacked the visible presence of the Council leader or the chair of the education committee. In turn commitment was expected at every level throughout the authority.

The final context variable was declaration. It was practised informally from the start. There were great expectations – and they were declared boldly. This project was to be a world leader. It was to raise attainment, it was to wipe out illiteracy – and as a result it was to change lives. The aim in tackling low achievement and illiteracy was to tackle everything known to be associated with it. The results anticipated from intergenerational change were higher self-esteem, lower disruption in schools, better school ethos, better staff morale, economic savings in remedial support, lower crime, a more skilled workforce and a stronger economy.

If the success of this research has built the foundation on which these high ambitions will be accomplished, then it will have promoted the framework of values in science that it has endorsed – the values of promoting health, caring and compassion, self-determination and participation, human diversity and social justice.

Epilogue

‘For all the money, time, energy and ingenuity we have spent on reading research, we are still at the stage of saying that children learn to read when there is something they want to read and an adult who takes the time and trouble to help them’ (Meek, 1983, p. 1).

In the 20 years or so since Meek published her qualitative, longitudinal studies of adolescents learning to read a great deal of knowledge about the reading process and the basis of effective teaching has been systematically accumulated (Chapter 6). Nevertheless, there remains essential truth in the statement that the achievement of competence in reading is based on these two fundamental requirements – the motivation of the learner to learn, and the commitment of the teacher to teach.

It is for the purpose of elucidating this dual foundation of reading achievement that this research has been conducted. It has sought to investigate the circumstances in which the learner will be best encouraged to have the motivation to learn, especially in a socio-cultural context marked by educational underachievement and lack of engagement with formal learning processes. It has also sought to consider the circumstances in which teachers will not only have a high commitment to teach in settings often marked by failure and discouragement, but also how they will be best equipped with the curricular content and methodologies most suited to successful outcomes, together with strategies to address the needs of those whose progress is impaired.

In pursuit of its aims, this study sought explicitly to address the ambitious agenda of applying psychology on a large scale to endemic social and educational problems with a view to laying a foundation for major, intergenerational change in disadvantaged populations. To achieve such a vision it was necessary not only to design effective interventions in terms of basic content and method, but also to manage the processes of large-scale educational change and to sustain and develop these processes over a long period.

There would have been an appealing simplicity in tackling an alternative research agenda – one that avoided the many pitfalls of large samples, of vast individual assessment programmes that at various times looked as if they might be unmanageable and of the messiness of multiple-component interventions with their associated problems of separating the effects of different variables and ensuring fidelity in delivery. The extensive world of literacy research is replete with more straightforward research choices in discrete and manageable areas, and West Dunbartonshire Council could have discharged its responsibilities by conducting a simple early intervention study on a less grand scale, and with a less formal commitment to rigorous research methodology. Instead, the arena of real world research was embraced in one of its untidiest settings; the agenda of critical psychology was espoused in the selection of a research programme designed to address crucial areas of human need; the thorny issues surrounding the concept of values in science had to be considered; and ways had to be found of incorporating

challenging – at times almost mercurial – concepts such as ‘vision’ and ‘declaration’ into the overall equation.

If, in meeting its aim to design, implement and evaluate an intervention for raising achievement and eradicating illiteracy, this study has contributed to the well-being of some of the most vulnerable children and young people in society, then perhaps it will have taken forward to a small degree the vision embraced in the Foreword:

‘Psychology has the potential to help bring about a significantly better world, in keeping with its ethical mandate to promote human welfare. Yet too often we settle for too little.’ (Prilleltensky & Fox, 1997, p. 4).

It is with the commitment to values in science, and to psychology as a force for positive change in society, especially amongst its most needy members, that this epilogue concludes and this work ends.

References

- Achenbach, T. (1982). *Developmental Psychopathology*. New York: Wiley.
- Acklaw, J. (1990). Change and the management of educational psychology services in England and Wales. *School Psychology International*, 11, 3-7.
- Adams, M. (1990). *Beginning to Read: Thinking and Learning about Print*. Cambridge, Mass.: Bradford.
- Albee, G. & Gullotta, T. (Eds) (1997). *Primary Prevention Works*. Sage, London.
- Allington, R. (1990). How policy and regulation influence instruction for at-risk learners: why poor readers rarely comprehend well. In B. Jones & L. Idol (Eds), *Dimensions of Thinking and Cognitive Instruction*. Hillsdale, NJ: Erlbaum.
- American Institutes of Research (1999). *An Educator's Guide to School-Wide Reform*. Washington, DC: American Institutes of Research.
- American Psychological Association (2001). *Publication Manual of the American Psychological Association* (5th edition). Washington, DC: Author.
- Ames, C. & Ames, R. (1989). *Research in Motivation in Education*, Vol. 3. San Diego: Academic Press.
- Atkinson, J. & Raynor, O. (1974). *Motivation and Achievement*. Washington: Winston.
- Axford, S., Blundell, L., Brown, C. & Macphee, H. (1994). Playground behaviour and activities: a learning context for personal and social development. In The Scottish Office Education Department, *Psychological Perspectives on Curriculum Development with Particular Reference to the 5-14 Programme*. Edinburgh: SOED.
- Ball, E. & Blachman, B. (1988). Phoneme segmentation training: effect on reading readiness. *Annals of Dyslexia*, 38, 208-225.
- Ballard, P. (1918). Norms of performance in simple oral arithmetic. *Journal of Experimental Pedagogy*, 4, 236-238.
- Bandura, A. (1986). *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1994). Self-efficacy. In V. Ramachaudran (Ed.), *Encyclopedia of Human Behaviour*, Vol. 4 (pp. 71-81). New York: Academic Press.
- Barber, T. (1976). *Pitfalls in Human Research*. Oxford: Pergamon.
- Barnes, J. & Lucas, H. (1975). Positive discrimination in education: individuals, groups and institutions. In Barnes, J. (Ed.) *Educational Priority: Curriculum Innovation in London EPAs*. London: HMSO.
- Barnett, W. & Escobar, C. (1990). Economic costs and benefits of early intervention. In S. Meisels & J. Shonkoff (Eds), *Handbook of Early Childhood Intervention*. Cambridge: Cambridge University Press.
- Bernstein, B. (1970). Education cannot compensate for society. *New Society*, 387, 344-347.
- Birman, B., Orland, M., Jung, R., Anson, R., Garcia, G., Moore, M., Funkhouser, J., Morrison, D., Turnbull, B. & Reiser, E. (1987). *The Current Operation of the Chapter 1 Program: Final Report from the National Assessment of Chapter 1*. Washington, DC: Government Printing Office.

- Blachman, B. (1984). What we have learned from longitudinal studies of phonological processing and reading, and some unanswered questions: a response to Torgesen, Wagner and Rashotte. *Journal of Learning Disabilities*, 27, 287-291.
- Blachman, B. (1989). Phonological awareness and word recognition: assessment and intervention. In A. Kamhi & H. Catts (Eds), *Reading Disabilities: A Developmental Language Perspective* (pp 133-158). Austin, Tex: PRO-ED.
- Blatchford, P., Burke, J., Farquhar, C., Plewis, I. & Tizard, B. (1987). Associations between pre-school reading related skills and later reading achievement. *British Education Research Journal*, 13, 1, 15-24.
- Blatchford, P. & Cline, T. (1992). Baseline assessment for school entrants. *Research Papers in Education*, 7, 247-269.
- Blatchford, P., Creeser, R. & Mooney, A. (1990). Playground games and playtime: the children's view. *Educational Research*, 32,3, 163-174.
- Blaxter, M. (1990). *Health and Lifestyles*. London: Routledge.
- Bloom, B. (1984). The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 4-16.
- Bloom, M. & Fischer, J. (1982). *Evaluating Practice: Guidelines for the Accountable Professional*. Englewood Cliffs, NJ: Prentice-Hall.
- Bond, G. & Dykstra, R. (1967). The co-operative research programme in first-grade reading instruction. *Reading Research Quarterly*, 2, 5-142.
- Botting, B. (1997). Mortality in childhood. In F. Drever & F. Whitehead (Eds), *Health Inequalities: Decennial Supplement*, Series DS No. 15. London: The Stationery Office.
- Boyle, J. & MacKay, T. (1990). Pupils with learning difficulties: What do schools expect of their educational psychologists? A Scottish perspective, *Learn*, 12, 24-32.
- Boyle, J. & MacKay, T. (submitted). Changing school expectations of educational psychologists: a 10-year follow-up study of continuity and change.
- Bradley, L. & Bryant, P. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 221, 746-747.
- Bradshaw, J. (2002). Comparisons of child poverty and deprivation internationally. In J. Bradshaw (Ed.), *The Well-Being of Children in the UK*, (pp. 17-26). London: Save the Children.
- Bramley, G. (1997). Poverty and local public services. In D. Gordon & C. Pantazis, *Breadline Britain in the 1990s*. Aldershot: Ashgate.
- Briggs, C. & Elkind, D. (1973). Cognitive development in early readers. *Developmental Psychology*, 9, 279-280.
- Briggs, S., MacKay, T. & Miller, S. (1995). The Edinbarnet playground project: changing aggressive behaviour through structured intervention. *Educational Psychology in Practice*, 11, 37-44.
- Brill, K. (1962). *Children, Not Cases*. Convocation Lecture. London: National Children's Home.
- Brophy, J. & Evertson, C. (1978). Context variables in teaching. *Educational Psychologist*, 12, 310-316.
- Brown, U., Scott, G., Mooney, G. & Duncan, B. (2002). *Poverty in Scotland 2002*. Glasgow: CPAG/Scottish Poverty Information Unit.
- Bryant, P. & Bradley, L. (1985). *Children's Reading Problems*. Oxford: Blackwell.
- Bryant, P., Bradley, L., Maclean, M. & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language*, 16, 407-428.

- Burton, C. & King, L. (2004). The health benefits of writing about intensely positive experiences. *Journal of Research in Personality*, 38, 150-163.
- Campaign for Learning (1998). *Attitudes to Learning 98. MORI State of the Nation Survey: Summary Report*. London: Author.
- Carroll, H. (1972). Remedial teaching of reading: an evaluation. *Remedial Education*, 7, 10-15.
- Chall, J. (1967). *Learning to Read: The Great Debate*. New York: McGraw-Hill.
- Chall, J. (1983). *Learning to Read: The Great Debate* (2nd edition). New York: McGraw-Hill.
- Chall, J. (1989). Learning to read: the great debate twenty years later. A response to 'Debunking the great phonics myth'. *Phi Delta Kappan*, 71, 521-538.
- Chapman, J. & Tunmer, W. (1991). Recovering Reading Recovery. *Australia and New Zealand Journal of Developmental Disabilities*, 17, 59-71.
- Clark, M. (1976). *Young Fluent Readers*. London: Heinemann.
- Clay, M. (1979a). *Stones – The Concepts about Print Test*. Exeter, NH: Heinemann.
- Clay, M. (1979b). *The Early Detection of Reading Difficulties* (3rd edition). Portsmouth, NH: Heinemann.
- Clay, M. (1991). *Becoming Literate*. Portsmouth, NH: Heinemann.
- Clay, M. (1993a). *Reading Recovery: A Guidebook for Teachers in Training*. Auckland, NZ: Heinemann.
- Clay, M. (1993b). *An Observation Survey of Early Literacy Achievement*. Auckland, NZ: Heinemann.
- Clerehugh, J., Hart, K., Pither, R., Rider, K. & Turner, K. (1994). *Early Years Easy Screen (EYES)*. Windsor, Berks: NFER-Nelson.
- Cowling, K. & Cowling, H. (1993). *Toe By Toe: A Highly Structured Multi-Sensory Reading Manual for Teachers and Parents*. Baildon, West Yorks: Author.
- Cox, T. (1983). The educational attitudes and views about school of a sample of disadvantaged fifteen-year-olds. *Educational Studies*, 9, 69-79.
- Cox, T. (2000). *Combating Educational Disadvantage: Meeting the Needs of Vulnerable Children*. London: Falmer.
- Cox, T. & Jones, G. (1983). *Disadvantaged Eleven Year Olds*. Oxford: Pergamon Press.
- Crawford, M., Kydd, L. & Riches, C. (1997). *Leadership and Teams in Educational Management*. Buckingham. Open University Press.
- Daniels, A. (1990). Social skills training for primary age children. *Educational Psychology in Practice*, 6, 3, 159-162.
- Davey Smith, G. & Egger, M. (1992). Socioeconomic differences in mortality in Britain and the United States. *American Journal of Public Health*, 82, 1079-81.
- Davie, R., Butler, H. & Goldstein, H. (1972). *From Birth to Seven: The Second Report of the National Child Development Study (1958 Cohort)*. London: Longman.
- Dawes, R. (1994). *House of Cards: Psychology and Psychotherapy Built on Myth*. Free Press, New York.
- de Shazer, S. (1985). *Keys to Solution in Brief Therapy*. New York: Norton.
- Deforges, M. & Lindsay, G. (1995). *Infant Index*. London: Hodder & Stoughton.
- Department of Education & Science (1991). *The Teaching and Learning of Reading in Primary Schools, Autumn 1990: A Report by HMI*. London: DES.
- Department of Social Security (1996). *Households Below Average Income: A Statistical Analysis, 1979-1993/94*. London: HMSO.

- Douglas, J. (1964). *The Home and the School*. London: MacGibbon & Kee.
- Dweck, C. (1989). Motivation. In R. Glaser & A. Lesgold (Eds), *The Handbook of Psychology and Education*, Vol. 1 (pp. 198-239). Hillsdale, NJ: Erlbaum.
- Dweck, C. (1991). Self-theories and goals: their role in motivation, personality and development.
- Dweck, C., Hong, Y. & Chiu, C. (1993). Implicit theories and individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin*, 19, 644-656.
- Ehri, L. (1991). Learning to read and spell words. In L. Rieben & C. Perfetti (Eds), *Learning to Read: Basic Research and its Implications* (pp. 57-73). Hillsdale, NJ: Erlbaum.
- Elliott, J., Arthurs, J. & Williams, R. (2000). Volunteer support in the primary classroom: the long-term impact of one initiative upon children's reading performance. *British Journal of Educational Research*, 26, 2, 227-244.
- Elmore, R. (1995). Getting to scale with good educational practice. *Harvard Educational Review*, 66, 1, 1-26.
- Equal Opportunities Commission (1996). *Briefings on Women and Men in Britain*. London: Author.
- Eron, L., Huesmann, L., Brice, B., Fisher, P. & Mermetstein, R. (1983). Age trends in the development of aggression, sex typing and related television habits. *Developmental Psychology*, 19, 71-77.
- Essen, J. & Wedge, P. (1982). *Continuities in Childhood Disadvantage*. London: Heinemann.
- European Union (2001). Draft report on social inclusion: results of the European Community Household Panel Survey 1998. Mimeo.
- Evans, M. & Wright, A. (1987). The Surrey school psychological service: an evaluation through teacher perceptions. *Educational Psychology in Practice*, 3, 12-20.
- Feitelson, D. (1988). *Facts and Fads in Beginning Reading: A Cross Language Perspective*. Norwood, NJ: Ablex Publishing Corporation.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- Fogelman, K. (Ed.) (1983). *Growing Up in Great Britain: Collected Papers from the National Child Development Study*. London: Macmillan.
- Foot, H., Howe, C., Anderson, A., Tolmie, A. & Warden, D. (Eds) (1994). *Group and Interactive Learning*. Southampton: Computational Mechanics Publications.
- Fox, D. & Prilleltensky, I. (1997). *Critical Psychology: An Introduction*. London: Sage.
- France, N. (1978). *The Primary Reading Test*. Windsor, Berks: NFER.
- France, N. (1981). *The Primary Reading Test (Levels 1 & 2): Teacher's Guide (Revised Edition)*. Windsor, Berks: NFER.
- Fraser, H., MacDougall, A., Pirrie, A. & Croxford, L. (2001). *National Evaluation of the Early Intervention Programme*. Final report to the Scottish Executive Education Department. University of Edinburgh.
- Freire, P. (1994). *The Pedagogy of Hope: Re-living Pedagogy of the Oppressed*. New York: Continuum.
- Freire, P. (2000). *Pedagogy of the Oppressed*. 30th Anniversary Edition. New York: Continuum.

- Fullan, M. (1991). *The New Meaning of Educational Change*. New York: Teachers College Press.
- Fullan, M. (1993). *Change Forces: Probing the Depths of Educational Reform*. London: Falmer.
- Fullan, M. (1999). *Change Forces: The Sequel*. Philadelphia, PA: Falmer.
- Fullan, M. (2001). *The New Meaning of Educational Change*. Third edition. London: Routledge Falmer.
- Gaa, J. (1973). Effects of individual goal-setting conferences on achievement, attitude, and goal-setting behaviour. *Journal of Experimental Education*, 42, 22-28.
- Garner, C. (1989). *Does Deprivation Damage?* Edinburgh: University of Edinburgh Centre for Educational Sociology.
- General Registrar's Office (1993). *The 1991 Census Report for Scotland*. Edinburgh: HMSO.
- Gillham, B. (Ed.) (1978). *Reconstructing Educational Psychology*. London: Croom Helm.
- Glass, G., McGaw, B. & Smith, M. (1981). *Meta-Analysis in Social Research*. Beverley Hills, CA: Sage.
- Good, T. & Grouws, D. (1975). Teacher rapport: some stability data. *Journal of Educational Psychology*, 67, 179-182.
- Goodman, Y. (1986). Children coming to know literacy. In W. Teale & E. Sulzby (Eds), *Emergent Literacy: Writing and Reading*, (pp. 1-14). Norwood, NJ: Ablex Publishing Corporation.
- Goss, T. (1996). *The Last Word on Power: Executive Re-Invention for Leaders who Must Make the Impossible Happen*. London: Piatkus.
- Gough, P. & Hillinger, M. (1980). Learning to read: an unnatural act. *Bulletin of the Orton Society*, 20, 179-196.
- Gough, P. & Juel, C. (1991). The first stages of word recognition. In L. Rieben & C. Perfetti (Eds), *Learning to Read: Basic Research and its Implications* (pp. 47-56). Hillsdale, NJ: Lawrence Erlbaum.
- Gregory, E. (2000). Recognising differences: reinterpreting family involvement in early literacy. In, T. Cox (Ed.), *Combating Educational Disadvantage: Meeting the Needs of Vulnerable Children* (pp. 103-120). London: Falmer.
- Greig, A. (2001). The educational psychologist as practitioner researcher: reality or dream? *Educational and Child Psychology*, 18, 4, 75-88.
- Gurney, P. (1987). Self esteem in the classroom. *School Psychology International*, 8, 21-29.
- Hannon, P. (1985). *Literacy, Home and School: Research and Practice in Teaching Literacy with Parents*. London: Falmer.
- Harding, S. (1985). Values and the nature of psychological well-being. In M. Abrams, D. Gerard & M. Timms (Eds), *Values and Social Change in Britain*. Basingstoke: Macmillan.
- Hargreaves, A. (1994). *Changing Teachers, Changing Times: Teachers' Work and Culture in the Postmodern Age*. London: Cassell.
- Haste, H., Hogan, A. & Zachariou, Y. (2001). Back (again) to the future. *The Psychologist*, 14, 1, 30-33.
- Hatch, T. (2000). *What Happens When Multiple Improvement Initiatives Collide?* Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching.

- Hawton, K., Salkovskis, P., Kirk, J. & Clark, D. (Eds) (1989). *Cognitive Behaviour Therapy for Psychiatric Problems: A Practical Guide*. Oxford: Oxford University Press.
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.
- Hersen, M. & Barlow, D. (1976). *Single Case Experimental Designs: Strategies for Studying Behaviour Change*. New York: Pergamon.
- Hewison, J. (1988). The long-term effectiveness of parental involvement: a follow-up to the Haringey Reading Project. *British Journal of Educational Psychology*, 58, 184-190.
- Hewison, J. & Tizard, J. (1980). Parental involvement and reading attainment. *British Journal of Educational Psychology*, 50, 209-215.
- Hollingshead, A. & Redlich, F. (1958). *Social Class and Mental Illness*. New York: Wiley.
- Hopkins, D. & Stern, D. (1996). Quality teachers, quality schools, *Teaching and Teacher Education*, 12, 5, 501-17.
- Hughes, M. & Greenhough, P. (1994). Is it better to work with a peer, an adult or both? A systematic comparison with children aged 5, 7 and 11 years. In H. Foot, C. Howe, A. Anderson, A. Tolmie & D. Warden (Eds), *Group and Interactive Learning*, (pp. 53-59). Southampton: Computational Mechanics Publications.
- Humes, W. (2003). Policy making in Scottish education. In T. Bryce & W. Humes (Eds), *Scottish Education: Second Edition – Post-Devolution* (pp. 74-85). Edinburgh: Edinburgh University Press.
- Isaac, A. (1992). Mental practice: does it work in the field? *The Sport Psychologist*, 6, 192-198.
- Issroff, K., Jones, A. & Scanlon, E. (1994). Case studies of children cooperating with computers: a time based analysis. In H. Foot, C. Howe, A. Anderson, A. Tolmie & D. Warden (Eds), *Group and Interactive Learning*, (pp. 73-79). Southampton: Computational Mechanics Publications.
- Iverson, S. & Tunmer, W. (1993). Phonological processing skills and the Reading Recovery programme. *Journal of Educational Psychology*, 85, 1, 112-126.
- Jarvis, S., Towner, E. & Walsh, S. (1995). Accidents. In B. Botting (Ed.), *The Health of Our Children: Decennial Supplement* (Series DS No. 11). London: OPCS.
- Jensen, A., Malcolm, L., Phelps, F. & Stoker, R. (2002). Changing patterns of thinking: individuals and organisations. *Educational Psychology in Practice*, 18, 1, 35-45.
- Johnston, R. & Watson, J. (1997). Developing reading, spelling and phonemic awareness skills in primary school children. *Reading*, 31, 37-40.
- Johnston, R., Connelly, V. & Watson, J. (1995). The effects of phonics teaching on reading development. In P. Owen & P. Pumfrey (Eds), *Children Learning to Read: International Concerns*, Vol. 1. London: Falmer.
- Jones, E., Kannouse, D., Kelley, H., Nisbett, R., Valins, S. & Weiner, B. (Eds) (1972). *Attribution: Perceiving the Causes of Behaviour*. Morristown, NJ: General Learning Press.
- Keller, J. (1983). Motivational design of instruction. In C. Riegeluth (Ed.), *Instructional Design Theories and Models*. Hillsdale, NJ: Erlbaum.
- Kershner, J. (1990). Self-concept and IQ as predictors of remedial success in children with learning disabilities. *Journal of Learning Disabilities*, 23, 368-374.

- Keys, W. & Fernandes, C. (1993). *What Do Students Think About School? Research into Factors Associated with Positive and Negative Attitudes Towards School and Education*. Slough: NFER.
- Knapp, M., Scott, S. & Davies, J. (1999). The cost of antisocial behaviour in younger children. *Clinical Child Psychology and Psychiatry*, 4, 457-473.
- Kramer, M. (1992). Barriers to the primary prevention of mental, neurological and psychosocial disorders of children: a global perspective. In G. Albee, L Bond & T. Cook Monsey (Eds), *Improving Children's Lives: Global Perspectives on Prevention*. London: Sage.
- Kumar, V. (1993). *Poverty and Inequality in the UK: The Effects on Children*. London: National Children's Bureau.
- Kwan, A. & Willows, D. (1998). Impact of early phonics instruction on children learning English as a second language. Paper presented at the National Reading Conference, Austin, Texas, December 1998.
- Lawrence, D. (1971). The effects of counselling on retarded readers. *Educational Research*, 13, 119-124.
- Lawrence, D. (1973). *Improved Reading Through Counselling*. London: Ward Lock.
- Lawrence, D. (1996). *Self-Esteem in the Classroom*. London: Paul Chapman.
- Levine, K. (1986). *The Social Context of Literacy*. London: Routledge & Kegan Paul.
- Lindsay, G. (1980). The Infant Rating Scale. *British Journal of Educational Psychology*, 50, 97-104.
- Lindsay, G. (1998). Baseline assessment: a positive or malign initiative? In B. Norwich & G. Lindsay (Eds) *Baseline Assessment: Benefits and Pitfalls* (pp. 8-40). Tamworth, Staffs: NASEN.
- Lindsay, G. & Deforges, M. (1998). *Baseline Assessment*. London: David Fulton.
- Lindsay, G. & Deforges, M. (1999). The use of the Infant Index/Baseline-PLUS as a baseline assessment measure of literacy. *Journal of Research in Reading*, 22, 1, 55-66.
- Lloyd, S. (1992). *The Phonics Handbook*. Chigwell, Essex: Jolly Learning.
- Locke, E., & Latham, G. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.
- Lokke, C., Irvine, G., M'gadzah, H. & Frederickson, N. (1997). The resurrection of psychometrics: fact or fiction? *Educational Psychology in Practice*, 12, 4, 222-233.
- Lothian Regional Council (1995). *Deprivation, Early Intervention and the Prevention of Reading Difficulties: An Evaluation of Pupils' Progress in an Intervention Project*. August 1995. Edinburgh: Author.
- MacDonald, G. & Cornwall, A. (1995). The relationship between phonological awareness and reading and spelling achievement eleven years later. *Journal of Learning Disabilities*, 28, 8, 523-527.
- MacKay, T. (1971). *A Progressive Addition Test*. Unpublished dissertation, University of Glasgow Library.
- MacKay, T. (1982). Physical handicap and special education. *SALGEP Quarterly*, 1, 4, 3-10.
- MacKay, T. (1983). Remedial education: its present position and future direction. *SALGEP Quarterly*, 2, 4, 20-23.
- MacKay, T. (1987). Planning research in child guidance. *SALGEP Quarterly*, 6, 1, 3-11.

- MacKay, T. (1989). Special education: the post-Warnock role for the educational psychologist. *BPS Scottish Division of Educational and Child Psychology, Newsletter 1*, 1-8.
- MacKay, T. (1995a). Reading failure in an area of multiple social disadvantage: response of a psychological service to a school's priorities. In The Scottish Office Education Department, *Matching Service Delivery to Client Needs: Quality Assurance in Psychological Services* (pp. 210-233). Edinburgh: The Scottish Office Education Department.
- MacKay, T. (1995b). Scottish educational legislation. In The British Psychological Society *Professional Psychology Handbook* (pp. 110-118). Leicester: BPS Books.
- MacKay, T. (1997). Psychological service delivery to primary schools: do head teachers want research? *Educational Psychology in Practice, 13*, 165-169.
- MacKay, T. (1999a). Baseline assessment in Scotland. *Journal of Research in Reading, 22*, 1, 81-88.
- MacKay, T. (1999b). Can endemic reading failure in socially disadvantaged children be successfully tackled? *Educational and Child Psychology, 16*, 1, 22-29.
- MacKay, T. (1999c). Education and the disadvantaged: is there any justice? *The Psychologist, 12*, 7, 344-349.
- MacKay, T. (1999d). *Quality Assurance in Education Authority Psychological Services: Self-Evaluation Using Performance Indicators*. Edinburgh: Scottish Office Education Department.
- MacKay, T. (2000a). Commentaries on Prilleltensky and Nelson, 'Promoting child and family wellness: priorities for psychological and social intervention'. *Journal of Community and Applied Social Psychology, 10*, 113-116.
- MacKay, T. (2000b). Educational psychology and the future of special educational needs legislation. *Educational and Child Psychology, 17*, 2, 27-35.
- MacKay, T. (2000c). The Record of Needs: does its distribution across Scotland reflect real need? *Education in the North, 7*, 29-34.
- MacKay, T. (2001a). Life for the underprivileged: Can psychology make a difference? *Proceedings of the British Psychological Society, 9*, 1, 22.
- MacKay, T. (2001b). The future belongs to psychology. *The Psychologist, 14*, 9, 466-469.
- MacKay, T. (2002a). Adult illiteracy: short-term and long-term solutions. Keynote address to the Society for Industrial and Organisational Psychology of SA, 5th Annual Industrial Psychology Conference, Pretoria, 13 June. <http://www.siopsa.org.za/2002Presentations>
- MacKay, T. (2002b). Raising attainment through an integrated learning system: evaluation of SuccessMaker in the Dumbarton Academy cluster. Unpublished report: West Dunbartonshire Council.
- MacKay, T. (2002c). The future of educational psychology. *Educational Psychology in Practice, 18*, 3, 245-253.
- MacKay, T. (2003a). Equal opportunities and social justice: the abuse of power in public services. *Proceedings of the British Psychological Society, 11*, 1, 106.
- MacKay, T. (2003b). Psychological services and their impact. In T. Bryce & W. Humes, *Scottish Education: Second Edition – Post-Devolution* (pp. 842-852). Edinburgh: Edinburgh University Press.
- MacKay, T. (2003c). Social inclusion: priorities and paradoxes for psychology. *Proceedings of the British Psychological Society, 11*, 1, 3.

- MacKay, T. (in preparation a). A social psychological perspective on international illiteracy: the example of South Africa.
- MacKay, T. (in preparation b). Children's educational achievement: do denominational schools make a difference?
- MacKay, T. & Boyle, J. (1994). Meeting the needs of pupils with learning difficulties: what do primary and secondary schools expect of their educational psychologists? *Educational Research*, 36, 221-230.
- MacKay, T. & Boyle, J. (submitted). Schools' perceptions of the value of educational psychologists: the key determinants.
- MacKay, T. & Briggs, S. (1994). 'Fair Play': the primary school playground. *Educational Psychology in Scotland*, 2, 25-31.
- MacKay, T. & McDonald, A. (in preparation). Tackling reading failure among secondary school pupils.
- MacKay, T. & Vassie, C. (1998). Is Scottish educational psychology changing? The results of a 16-year time-sampling of the work of psychologists. *BPS Scottish Division of Educational Psychology: Abstracts of Annual Conference for Educational Psychologists in Scotland, 1998*.
- MacKay, T. & Watson, K. (1996). Literacy, social disadvantage and early intervention: a strategy to enhance reading in primary school. *Final report to Scottish Office Education and Industry Department*.
- MacKay, T. & Watson, K. (1999). Literacy, social disadvantage and early intervention: enhancing reading achievement in primary school. *Educational and Child Psychology*, 16, 1, 30-36.
- Mackenbach, J. (1994). Socio-economic inequalities in The Netherlands: impact of a five year research programme. *British Medical Journal*, 309, 1487-1491.
- Mackenbach, J., Caspar, W. & Joost, B. (1996). Differences in the misreporting of chronic conditions, by level of education: the effect on inequalities in prevalence rates. *American Journal of Public Health*, 86, 706-711.
- MacLean, M., Bryant, P. & Bradley, L. (1987). Rhymes, nursery rhymes and reading in early childhood. *Merrill-Palmer Quarterly*, 33, 255-281.
- Maddux, J., & Stanley, M. (Eds.) (1986). Special issue on self-efficacy theory. *Journal of Social and Clinical Psychology*, 4 (Whole No.3).
- Malone, T. (1981). Towards a theory of intrinsically motivating instruction. *Cognitive Science*, 4, 333-369.
- Marris, P. (1975). *Loss and Change*. New York: Anchor Press.
- Martin, K. & Hall, C. (1995). Using mental imagery to enhance intrinsic motivation. *Journal of Sport and Exercise Psychology*, 17, 1, 54-69.
- Mason, J. (1981). *Pre-reading: A Developmental Perspective*. Technical Report No. 198, Urbana, Ill: University of Illinois, Centre for the Study of Reading.
- Mason, J. (1985). Cognitive monitoring and early reading: a proposed model. In D. Forrest-Pressley, G. MacKinnon & T. Waller (Eds). *Mecacognition, Cognition and Human Performance* (pp. 77-101). Orlando: Academic Press.
- Maxwell, W. (1994). Special educational needs and social disadvantage in Aberdeen City school catchment zones. *Educational Research*, 36, 1, 25-37.
- Mayo, E. (1933). *The Human Problems of an Industrial Civilisation*. New York: MacMillan.
- McCallum, I. (1993). *Testing 7-year Olds – Performance and Context: Projecting School Roles and Assessing Performance*. London: London Research Centre.

- McClelland, D. (1985). *Human Motivation*. Glenview, IL: Scott, Foresman.
- McClelland, D., Atkinson, J., Clark, R. & Lowell, E. (1953). *The Achievement Motive*. New York: Appleton-Century.
- McConnell, D. (2004). An evaluation of the Toe By Toe programme in two Scottish local authorities. University of Glasgow Theses.
- McLeod, J. & Anderson, J. (1972). *The Gapadol Reading Comprehension Test*. Port Melbourne, Australia: Heinemann.
- McLoone, P. & Boddy, F. (1994). Deprivation and mortality in Scotland, 1981 and 1991. *British Medical Journal*, 309, 1465-1470.
- McMillan, G., Fox, D. & Wood, K. (1994). Deprivation, early intervention and the prevention of reading difficulties. In SOED/Regional Psychological Services. *Professional Development Initiatives 1992-93.: Psychological Perspectives on Curriculum Development with Particular Reference to the 5-14 Programme*. Edinburgh: SOED/Regional Psychological Services.
- McQuillan, J. & Tse, L. (1995). Child language brokering in linguistic minority communities. *Language and Education*, 9, 3, 195-215.
- Meek, M. (1983). *Achieving Literacy: Longitudinal Studies of Adolescents Learning to Read*. London: Routledge & Kegan Paul.
- Meltzer, H., Gatward, R., Goodman, R. & Ford, T. (2000). *Mental Health of Children and Adolescents in Great Britain*. London: The Stationery Office.
- Morgan, M. (1987). Self-monitoring and goal setting in private study. *Contemporary Educational Psychology*, 12, 1-6.
- Mortimore, J. & Blackstone, T. (1982). *Disadvantage and Education*. Aldershot: Gower.
- Mortimore, P. & Whitty, G. (2000). Can school improvement overcome the effects of disadvantage? In T. Cox (Ed.), *Combating Educational Disadvantage: Meeting the Needs of Vulnerable Children* (pp. 156-176). London: Falmer.
- Mullis, I., Martin, M., Gonzalez, E., & Kennedy, A. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools*. Chestnut Hill, MA: Boston College.
- Murphy, S. (1990). Models of imagery in sport psychology: a review. *Journal of Mental Imagery*, 14, 3 & 4, 153-172.
- Murray, J. (1963). *The Power of Your Subconscious Mind*. Englewood Cliffs, NJ: Prentice-Hall.
- Muter, V., Hulme, C., Snowling, M. & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, 71, 3-27.
- National Commission on Education (1996). *Success Against the Odds: Effective Schools in Disadvantaged Areas*. London: Routledge.
- Neale, M. (1989). *The Neale Analysis of Reading Ability: Revised British Edition*. Windsor, Berks: NFER-Nelson.
- NHS Scotland (2004). Constituency health and well-being profiles, 2004. ISBN 1-84485-076-5. Glasgow: Author.
- Nicolson, R., Fawcett, Moss, H., Nicolson, M. & Reason, R. (1999). Early reading intervention can be effective and cost-effective. *British Journal of Educational Psychology*, 69, 47-62.
- O'Hagen, F. & Swanson, W. (1983). Teachers' views regarding the role of the educational psychologist in schools. *Research in Education*, 29, 29-45.

- Office for National Statistics (2000). Infant mortality by social class, 1981, 1991, 1996 and 1997. *Social Trends Dataset*. London: ONS.
- Office for National Statistics (2001a). *Mortality Statistics: Childhood, Infant and Perinatal. Review of the Registrar General on Deaths in England and Wales, 1999*. Series DH3 No. 32. London: The Stationery Office.
- Office for National Statistics (2001b). *Health Statistics Quarterly* (Autumn). London: The Stationery Office.
- Office of Population Censuses and Surveys (1988). *Occupational Mortality: Childhood Supplement 179-180, 1982-83*. London: HMSO.
- Office of Population Censuses and Surveys (1991). *The 1991 Census*. London: OPCS.
- Olofsson, A. & Lundberg, I. (1985). Evaluation of long-term effects on phonemic awareness training in kindergarten: illustrations of some methodological problems in evaluation research. *Scandinavian Journal of Psychology*, 26, 21-34.
- Olson, R., Wise, B., Ring, J. & Johnston, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: effects on the post-training development of word recognition. *Scientific Studies of Reading*, 1, 235-253.
- Organisation for Economic Cooperation and Development (2000). *Literacy in the Information Age: Final Report of the International Adult Literacy Survey*. Paris: Author.
- Organisation for Economic Cooperation and Development (2004). *Learning for Tomorrow's World: First results from PISA 2003*. Paris: Author.
- Orlick, T., Zitzelsberger, L., Li-Wei, Z. & Qi-wei, M. (1992). The effect of mental-imagery training on performance enhancement with 7-10 year old children. *The Sport Psychologist*, 6, 230-241.
- Patterson, L. (1992). Social class in Scottish education. In S. Brown (Ed.), *Class, Race and Gender in Schools*. Glasgow: Scottish Council for Research in Education.
- Pettifor, J. (1996). Ethics: virtue and politics in the science and practice of psychology. *Canadian Psychology*, 37, 1-12
- Pinnell, G., DeFord, D. & Lyons, C. (1988). *Reading Recovery: Early Intervention for At-Risk First Graders*. Virginia: Virginia Educational Research Service.
- Plowden Committee (1967). *Children and Their Primary Schools: A Report of the Central Advisory Council for Education*. London: HMSO.
- Prescott-Carke, P. & Primatesta, P. (Eds). (1999). *Health Survey for England, 1997: The Health of Young People, '95-'97*. London: The Stationery Office.
- Prilleltensky, I. (1994). *The Morals and Politics of Psychology: Psychological Discourse and the Status Quo*. Albany, NY: State University of New York Press.
- Prilleltensky, I. & Fox, D. (1997). Introducing critical psychology: values, assumptions and the status quo. In D. Fox & I. Prilleltensky (Eds), *Critical Psychology: An Introduction* (pp. 3-20). London: Sage.
- Prilleltensky, I. & Nelson, G. (1997). Community psychology: reclaiming social justice. In D. Fox & I. Prilleltensky (Eds), *Critical Psychology: An Introduction* (pp. 166-184). London: Sage.
- Prilleltensky, I. & Nelson, G. (2000). 'Promoting child and family wellness: priorities for psychological and social intervention'. *Journal of Community and Applied Social Psychology*, 10, 2, 85-105.

- Pumfrey, P. (1979). Improving reading through counselling – royal road or cul-de-sac? In Thackrey, D. (Ed.) *Growth in Reading*. London: Ward Lock.
- Rappaport, J. (2000). Commentaries on Prilleltensky and Nelson, ‘‘Promoting child and family wellness: priorities for psychological and social intervention’’. *Journal of Community and Applied Social Psychology*, 10, 107-110.
- Rappaport, J. & Stewart, E. (1997). A critical look at critical psychology: Elaborating the questions. In D. Fox & I. Prilleltensky (Eds), *Critical Psychology: An Introduction* (pp. 301-317). London: Sage.
- Rawson, M. (1995). *Dyslexia Over the Life Span: A 55 Year Longitudinal Study*. Cambridge, MA: Educator’s Publishing Service.
- Reynolds, D. (1985). School effectiveness research: a review of the literature. *School Organisation and Management Abstracts*, 1, 1, 5-14.
- Riley, J. (1996). *The Teaching of Reading: The Development of Literacy in the Early Years of School*. London: Paul Chapman.
- Robins, L. (1978). Sturdy childhood predictors of adult antisocial behaviour: replications from longitudinal studies. *Psychological Medicine*, 8, 611-622.
- Robson, C. (1993). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford: Blackwell.
- Roethlisberger, F. & Dickson, W. (1939). *Management and the Worker*. Cambridge, Mass: Harvard University Press.
- Rogoff, B. (1990). *Apprenticeship in Thinking: Cognitive Development in Social Contexts*. Oxford: Oxford University Press.
- Roff, J. & Wirt, R. (1984). Childhood aggression and social adjustments as antecedents of delinquency. *Journal of Abnormal Child Psychology*, 12, 111-126.
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.
- Rosenthal, R. & Fode, K. (1963). The effects of experimenter bias of the performance of the albino rat. *Behavioral Science*, 8, 183-189.
- Rosenthal, R. & Jacobsen, L. (1968). *Pygmalion in the Classroom*. New York: Holt.
- Rosnow, R. & Rosenthal, R. (1997). *People Studying People: Artifacts and Ethics in Behavioural Research*. New York: Freeman.
- Ross, M. (1998). Ounces of prevention and pounds of cure: issues in the promotion of mental health. *The Bulletin of the Scottish Branch of the British Psychological Society*, 26, 10-17.
- Roure, R., Collet, C., Deschaumes-Molinario, C., Dittmar, A., Raqa, H., Delhomme, G. & Vernet-Maury, E. (1998). Autonomic nervous system responses correlate with mental rehearsal in volleyball training. *European Journal of Applied Physiology*, 78, 99-108.
- Rutter, M., Maughan, B., Mortimore, P. & Ousten, J. (1979). *Fifteen Thousand Hours: Secondary Schools and their Effects on Children*. London: Open Books.
- Rutter, M., Tizard, J. & Whitmore, K. (1970). *Education, Health and Behaviour*. London: Longman.
- Sacker, A., Schoon, I. & Bartley, M. (2001). Sources of bias in special needs provision in mainstream primary schools: evidence from two British cohort studies. *European Journal of Special Needs Education*, 16, 3, 259-276.
- Sacker, A., Schoon, I. & Bartley, M. (2002). Social inequality in educational achievement and psychosocial adjustment throughout childhood: magnitude and mechanisms. *Social Science and Medicine*, 55, 863-880.

- Sammons, P. (1995). Gender, ethnic and socio-economic differences in attainment and progress: a longitudinal analysis of student achievement over nine years. *British Educational Research Journal*, 23, 4, 490-511.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key Characteristics of Effective Schools: A Review of School Effectiveness Research*. London: Office for Standards in Education.
- Saracho, O. (1991). Teacher expectations of student performance: a review of the research. *Early Childhood Development and Care*, 76, 27-41.
- Sarason, S. (1982). *The Culture of the School and the Problem of Change*. Second Edition. Boston, Mass: Allyn & Bacon.
- Sarason, S. (1990). *The Predictable Failure of Educational Reform*. San Francisco: Jossey-Bass.
- School Curriculum and Assessment Authority (1997). *Looking at Children's Learning*. London: SCAA.
- Schunk, D. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, 33, 2, 359-382.
- Schweinhart, L. & Weikart, D. (1997). *Lasting Differences: The High/Scope Preschool Curriculum Comparison Through Age 23*. Ypsilanti, MI: High/Scope Press.
- Schweinhart, L., Barnes, H. & Weikart, D. (1993). *Significant Benefits: The High/Scope Perry Preschool Study Through Age 27*. Ypsilanti, MI: High/Scope Press.
- Schweinhart, L., Weikart, D. & Larner, M. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly*, 1, 1, 15-45.
- Scottish Executive (1999). *Social Justice: A Scotland Where Everyone Matters*. The Scottish Executive Website Social Inclusion Section. <http://www.scotland.gov.uk/inclusion>.
- Scottish Executive (2002). *Review of Provision of Educational Psychology Services in Scotland* (The Currie Report). Edinburgh: Scottish Executive.
- Scottish Executive Education Department (2003). *Report of the Sixth Survey of English Language (2001)*. Assessment of Achievement Programme. Edinburgh: SEED.
- Scottish Office (1999). *Towards a Healthier Scotland: A White Paper on Health*. Edinburgh: The Stationery Office.
- Scottish Office Department of Health (1998). *Research Strategy for the National Health Service in Scotland*. Edinburgh: The Stationery Office No. J19705 7/98.
- Scottish Office Education and Industry Department (1996). *AAP English 1995: A Summary Report of the Fourth Survey*. Assessment of Achievement Programme. Edinburgh: SOEID.
- Scottish Office Education and Industry Department (1998). *Baseline Assessment in Scotland: Pilot Procedures*. Edinburgh: SOEID.
- Scottish Office Education Department (1992). *Curriculum and Assessment in Scotland, National Guidelines: English language 5-14*. Edinburgh: SOED.
- Scottish Office Education Department (1993). *Curriculum and Assessment in Scotland, National Guidelines: the structure and balance of the curriculum 5-14*. Edinburgh: SOED.

- Shaughnessy, M. & Self, E. (1997). Mentoring emotionally sensitive individuals. *Mentor*. Issue 1. Journal of the Mentoring and Leadership Resource Network. http://www.mentors.net/03journal/j1_emotional.html.
- Shuttleworth, I. (1995). The relationship between social deprivation, as measured by individual free school meal eligibility, and educational attainment at GCSE in Northern Ireland: a preliminary investigation. *British Educational Research Journal*, 21, 4, 487-504.
- Singleton, C., Thomas, K. & Horne, J. (1998). *CoPS Baseline Assessment System*. Beverley, East Yorkshire: Lucid Research.
- Slavin, R. & Madden, N. (1998). *Disseminating Success for All*. Baltimore: Johns Hopkins University.
- Slavin, R. & Madden, N. (2000). *One Million Children*. London: Paul Chapman.
- Sloane, H., Endo, G., Hawkes, T. & Jenson, W. (1990). Decreasing children's fighting through self-instructional parent training materials. *School Psychology International*, 11, 17-29.
- Smith, P., Boulton, M. & Cowie, H. (1990). *Ethnic Relations and Sociometric Status in Middle School: An Intervention Study*. Report to ESRC, Swindon.
- Smith, T. & Noble, M. (1995). *Education Divides: Poverty and Schooling in the 1990s*. London: Child Poverty Action Group.
- Stallings, F. (1976). How instructional processes relate to child outcomes in a national study of Follow Through. *Journal of Teacher Education*, 27, 1, 43-47.
- Stallings, F., Robbins, P., Presbrey, L. & Scott, J. (1986). Effects of instruction based on the Madeline Hunter model of students' achievement: findings from a Follow Through project. *The Elementary School Journal*, 86, 571-587.
- Stanovich, K. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-406.
- Stanovich, K. & Stanovich, P. (1995). How research might inform the debate about early reading acquisition. *Journal of Research in Reading*, 18, 2, 87-105.
- Stanovich, K., Cunningham, A. & Feeman, D. (1984). Relation between early reading acquisition and word decoding with and without context: a longitudinal study of first-grade children. *Journal of Educational Psychology*, 76, 668-677.
- Stigler, J. & Hiebert, J. (1999). *The Teaching Gap*. New York: The Free Press.
- Stobie, I. (1996). Plus ça change, plus c'est la même chose. *School Psychology International*, 17, 5-16.
- Stobie, I. (2002a). Processes of 'change' and 'continuity' in educational psychology: Part I. *Educational Psychology in Practice*, 18, 3, 203-212.
- Stobie, I. (2002b). Processes of 'change' and 'continuity' in educational psychology: Part II. *Educational Psychology in Practice*, 18, 3, 213-237.
- Stornelli, D. & Willows, D. (1998). Effect of more and earlier phonics instruction on kindergarten literacy outcomes. Paper presented at the National Reading Conference, Austin, Texas, December 1998.
- Strand, S. (1999). Ethnic group, sex and economic disadvantage associations with pupils' educational progress from baseline to the end of Key Stage 1. *British Educational Research Journal*, 25, 2, 179-202.
- Street, B. (1993a). The new literacy studies, guest editorial. *Journal of Research in Reading*, 16, 81-97.
- Street, B.(Ed.) (1993b). *Cross-cultural Approaches to Literacy*. Cambridge: Cambridge University Press.

- Stuart, M. (1999). Getting ready for reading: early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, 69, 587-605.
- Stuart, M. (2004). Getting ready for reading: a follow-up study of inner city second language learners at the end of Key Stage 1. *British Journal of Educational Psychology*, 74, 1, 15-36.
- Sumbler, K. & Willows, D. (1996). Phonological awareness and alphabetic coding instruction within balanced senior kindergartens. Paper presented as part of Symposium, *Systematic Phonics Within a Balanced Literacy Programme*. National Reading Conference, Charleston, SC.
- Sylva, K. (2000). Early childhood education to ensure a 'fair start' for all. In T. Cox, *Combating Educational Disadvantage: Meeting the Needs of Vulnerable Children* (pp 121-135). London: Falmer.
- Teale, W. (1986). Emergent literacy: reading and writing development in early childhood. In J. Readence & R. Baldwin (Eds), *Research in Literacy: Merging Perspectives* (pp. 45-74). Rochester, NY: National Reading Conference.
- The Basic Skills Agency (2001). *Adults' Basic Skills*. London: Author.
- The Jolly Phonics Case Studies (undated). Chigwell, Essex: Jolly Learning.
- The Scottish Office HMI (1997). *A curriculum framework for children in their preschool year*. Edinburgh: SOEID.
- Thomas, J. (1980). *The Self in Education*. Windsor: NFER
- Thomson, J. & Whelan, K. (1997). A community approach to road safety education using practical training methods: the Drumchapel project. *Road Safety Research Report No. 3*. London: Department of Transport.
- Thomson, L. (1996). Reading failure and social deprivation. *Educational Psychology in Practice*, 12, 3, 166-174.
- Tizard, B., Blatchford, P., Burke, J., Farquhar, C. & Plewis, I. (1988). *Young Children at School in the Inner City*. Hove: Lawrence Erlbaum.
- Tizard, J., Schofield, W. & Hewison, J. (1982). Symposium: reading-collaboration between teachers and parents in assisting children's reading. *British Journal of Educational Psychology*, 52, 1, 1-15.
- Tolmie, A. & Howe, C. (1994). Computer-directed group activity and the development of children's hypothesis testing skills. In H. Foot, C. Howe, A. Anderson, A. Tolmie & D. Warden (Eds), *Group and Interactive Learning*, (pp. 139-144). Southampton: Computational Mechanics Publications.
- Toomey, D. (1993). Parents hearing their children read: a review. *Educational Research*, 35, 223-236.
- Topping, K. (1977). An evaluation of the long-term effects of remedial teaching. *Remedial Education*, 12, 2, 84-6.
- Topping, K. (1978). Consumer confusion and professional conflict in educational psychology. *Bulletin of the British Psychological Society*, 31, 265-267.
- Topping, K. (1988). *The Peer Tutoring Handbook: Promoting Co-operative Learning*. London: Croom Helm.
- Topping, K. (1995). *Paired Reading, Spelling and Writing: The Handbook for Teachers and Parents*. London: Cassell.
- Topping, K. (1996). The effectiveness of family literacy. In S. Wolfendale & K. Topping (Eds), *Family Involvement in Literacy: Effective Partnerships in Education*. London: Cassell.

- Topping, K. (2001). *Thinking, Reading, Writing: A Practical Guide to Paired Learning with Peers, Parents and Volunteers*. London: Continuum.
- Topping, K. & Lindsay, G. (1991). Parental involvement in reading: the influence of socioeconomic status and supportive home visiting. *Children and Society*, 5, 306-316.
- Topping, K. & Lindsay, G. (1992). Paired reading: a review of the literature. *Research Papers in Education*, 7, 1-50.
- Topping, K. & Wolfendale, S. (Eds) (1985). *Parental Involvement in Children's Reading*. Beckenham: Croom Helm.
- Torgesen, J., Wagner, R. & Rashotte, C. (1994). Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities*, 27, 5, 276-286.
- Towner, E., Jarvis, S., Walsh, S. & Aynsley-Green, A. (1994). Measuring exposure to risk in schoolchildren aged 11-14. *British Medical Journal*, 308, 449-452.
- Tudor Hart, J. (1971). The inverse care law. *Lancet*, 1, 7696, 405-412.
- Tunmer, W., Herriman, M. & Nesdale, A. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly*, 23, 134-158.
- Turner, M. & Burkard, T. (1996). *Reading Fever: Why Phonics Must Come First*. London: Centre for Policy Studies.
- United Nations (2001). *Human Development Report 2001*. New York: United Nations Development Programme.
- Vassie, C. & Watson, K. (1990). 'First catch your hare – effecting a process of change in service delivery'. In *New demands and new responses: a comparative evaluation of service delivery*. Edinburgh: Scottish Education Department/Regional Psychological Services.
- Vellutino, F. & Scanlon, D. (1987). Phonological coding, phonological awareness and reading ability: evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly*, 33: 321-363.
- Vygotsky, L. (1978). *Mind in Society*. Cambridge, Mass: Harvard University Press.
- Wagner, R. & Torgesen, J. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192-212.
- Walker, A. & Walker, C. (1997). *Britain Divided: The Growth of Social Exclusion in the 1980s and 1990s*. London: Child Poverty Action Group.
- Watson, J. & Johnston, R. (1998). Accelerating reading attainment: the effectiveness of synthetic phonics. *Interchange No. 57*. Edinburgh: The Scottish Office Education and Industry Department.
- Wattenberg, W. & Clifford, C. (1964). Relation of self concepts to beginning achievement in reading. *Child Development*, 35, 461-467.
- Webb, E., Campbell, D., Schwartz, R., Sechrest, L. & Grove, J. (1981). *Non-reactive Measures in the Social Sciences* (2nd edition). Boston, Mass: Houghton Mifflin.
- Weinberger, J. (1996). *Literacy Goes to School: the parents' role in young children's literacy development*. London: Paul Chapman.
- Weiner, B. (1974). *Achievement Motivation and Attribution Theory*. Morristown, NJ: General Learning Press.
- Weiner, B. (1986). *An Attributional Theory of Motivation and Emotion*. New York: Springer-Verlag.
- Weiner, B. (1990a). History of motivational research in education. *Journal of Educational Psychology*, 82, 4, 616-622.

- Wells, A. (1998). *What Works in Secondary Schools? Catching up with Basic Skills*. London: Basic Skills Agency.
- Wells, G. (1987). *The Meaning Makers*. London: Hodder & Stoughton.
- West Dunbartonshire Childcare Partnership (1999). *Childcare audit: a summary*. West Dunbartonshire Council.
- Whitehead, M. & Drever, F. (1999). Narrowing social inequalities in health? Analysis of trends in mortality among babies of lone mothers. *British Medical Journal*, 318, 908.
- Whitmarsh, A. & Summerfield, C. (1996). *Social Focus on Ethnic Minorities*. Office for National Statistics. London: HMSO.
- Wilkinson, E. & Napuk, A. (1997). *Baseline assessments: a review of literature*. Edinburgh: Scottish Office Education and Industry Department.
- Wilkinson, R. (1992). National mortality rates: the impact of inequality? *American Journal of Public Health*, 82, 1082-84.
- Williams, J. (1986). The role of phonemic analysis in reading. In J. Torgesen & B. Wong (Eds), *Learning Disabilities: Some New Perspectives*. New York: Academic Press.
- Wimmer, H., Landerl, K. & Schneider, W. (1991). The role of alliteration and rhyme in learning to read a regular orthography. Unpublished manuscript. University of Salzburg, Institut für Psychologie. Salzburg, Austria. (Cited in Muter et al. 1998).
- Wolfendale, S. & Topping, K. (Eds.) (1996). *Family Involvement in Literacy: Effective Partnerships in Education*. London: Cassell.
- Wragg, E., Wragg, C., Haynes & Chamberlin, R. (1998). *Improving Literacy in the Primary School*. London: Routledge.
- Wright, A. (1992). Evaluation of the first British Reading Recovery programme. *British Educational Research Journal*, 18, 351-368.
- Zimmerman, B. (1989). A social cognitive view of self-regulated learning. *Journal of Educational Psychology*, 81, 329-339.

Appendices

Appendix 1 The Structure of Scottish Education

Appendix 2 Baseline Assessment

Appendix 3 Synthetic Phonics Study – Spelling Test

Appendix 4 Jolly Phonics Questionnaire

Appendix 5 Toe By Toe – Sample Pages

Appendix 6 Impact – The Response of the Media

Appendix 1

The Structure of Scottish Education

Statutory school age in Scotland is 5-16 years. Most children spend two years in pre-school education at nursery schools or other pre-five establishments from age 3, entering primary school at age 5 and secondary school at age 12. The system may be illustrated as follows:

Nursery school

First nursery school year	–	Age 3
The ‘pre-school year’	–	Age 4

Primary school

Primary 1 (P1)	–	Age 5
Primary 2 (P2)	–	Age 6
Primary 3 (P3)	–	Age 7
Primary 4 (P4)	–	Age 8
Primary 5 (P5)	–	Age 9
Primary 6 (P6)	–	Age 10
Primary 7 (P7)	–	Age 11

Secondary school

Secondary 1 (S1)	–	Age 12
Secondary 2 (S2)	–	Age 13
Secondary 3 (S3)	–	Age 14
Secondary 4 (S4)	–	Age 15
Secondary 5 (S5)	–	Age 16
Secondary 6 (S6)	–	Age 17

Appendix 2

Baseline Assessment

(Section 4 – Developmental Tasks – was not utilised in the study reported here)

Appendix 3

Synthetic Phonics Study – Spelling Test

Appendix 4

Jolly Phonics Questionnaire

Appendix 5

Toe By Toe – Sample Pages

Appendix 6

Impact – The Response of the Media