Achieving the Vision

The Final Research Report of the

West Dunbartonshire Literacy Initiative

Tommy MacKay

November 2007

Ordered from: education.centralregistry@west-dunbarton.gov.uk

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

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 10^{th} 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 *t* 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 *t* tests with effect sizes calculated on the standard deviations of the pre-test cohorts Targeted sample: related two-sample *t* 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 *t* 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 *t* 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 *t* 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 10^{th} 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 10^{th} 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 (t = 8.038)
Primary 4	8y 1m	1026	8y 4m	1044	0.14	p<0.001 (t = 3.461)
Primary 7	10y 6m	1022	11y 0m	1091	0.20	p<0.001 (t = 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 10^{th} 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	$p<0.001$,	1.74
Controls	8y 5m	8y 8m	$t = 5.65$	

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

Gains in reading age	Number of pupils
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 were1y 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 %
Bottom 5%	*25*	*25*	*23*	*23*
Bottom 10%	*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 10^{th} 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 if 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 in 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: Coninuum.

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