

THE EFFECTS ON THE PHONOLOGICAL PROCESSING
SKILLS OF DISABLED READERS OF PARTICIPATING
IN DIRECT INSTRUCTION READING PROGRAMS

by

KERIN JOHN HEMPENSTALL
B.Sc., Dip.Ed., Dip.Soc.Studies,
Dip.Ed.Psych., MAPsS

A THESIS

submitted to

Royal Melbourne Institute of Technology

in fulfilment of the requirements

for the degree of

DOCTOR of PHILOSOPHY

Faculty of Applied Science

Department of Psychology and Intellectual Disability Studies
DECEMBER 1997

DECLARATION

I, the candidate, Kerin John Hempenstall, certify that:

- a) except where due acknowledgment has been made, the work is that of the candidate alone;
- b) the work has not been submitted previously, in whole or in part, to qualify for any other academic award;
- c) the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program.

Signature of the candidate

Kerin John Hempenstall

ACKNOWLEDGMENTS

I would like to express my gratitude to Professor Alan Hudson for his advice, support and practical assistance, and to Dr John Reece and Sophie Kaufmann for their statistical and computing assistance.

I wish to dedicate this dissertation to my parents whose belief in higher education and their continuous support provided valuable motivation, and to my family for enabling me to assign priority to the task.

PARTS OF THE THESIS ALREADY PUBLISHED:**Chapter Two published as:**

Hempenstall, K. (1997). The Whole Language-phonics controversy: An historical perspective. *Educational Psychology, 17*, 399-418.

Chapter Three published as:

Hempenstall, K. (1996). The Whole Language approach to reading: An empiricist critique. *Australian Journal of Learning Disabilities, 1*(3), 22-32.

Chapter Four published as:

Hempenstall, K. (1997). The role of phonemic awareness in beginning reading: A review. *Behaviour Change, 14*, 1-14.

Chapter Five published as:

Hempenstall, K. (1996). The gulf between educational research and policy: The example of Direct Instruction and Whole Language. *Behaviour Change, 13*, 33-46.

TABLE OF CONTENTS

TITLE PAGE	(i)
DECLARATION	(ii)
ACKNOWLEDGMENTS	(iii)
PARTS OF THESIS ALREADY PUBLISHED	(iv)
TABLE OF CONTENTS	(v)
LIST OF FIGURES	(xi)
LIST OF TABLES	(xiii)
ABSTRACT	(xv)

CHAPTER ONE: INTRODUCTION

<i>Problem</i>	1
<i>Instructional Methods</i>	1
<i>Causes of Reading Failure</i>	1
<i>Phonemic Awareness: The Research</i>	2
<i>Other Phonological Processes</i>	3
<i>Phonemic Awareness and Reading: The Relationship</i>	4
<i>Phonemic Awareness and Older Students</i>	5
<i>Models of Reading Development</i>	6
<i>Dyslexics: A Special Subgroup?</i>	7
<i>Instructional Issues</i>	10
<i>The Corrective Reading Program</i>	11
<i>Phonological Processes and Spelling</i>	11

CHAPTER TWO: HISTORY OF THE READING DEBATE

<i>A History of English Reading and Writing</i>	13
<i>Writing Systems: From Logograms to the Alphabet</i>	13
<i>Problems of Written English</i>	15
<i>The Problems of Literacy</i>	16
<i>The Teaching of Reading: The Emergence of Meaning-Centred Approaches</i>	17
<i>The History of Educational Research into Teaching Reading</i>	19
<i>The Great Debate</i>	22

<i>The USOE Study</i>	23
<i>Follow Through</i>	24
<i>Becoming a Nation of Readers</i>	25
<i>The Impact of Research on Practice</i>	27

CHAPTER THREE: THE WHOLE LANGUAGE APPROACH

<i>Whole Language: History</i>	28
<i>Assumptions of the Whole Language Model</i>	30
<i>Naturally Unfolding Development</i>	30
<i>Reading as a Natural Process</i>	30
<i>The Induction of the Alphabetic Principle</i>	31
<i>Can Whole Language and Phonics be Reconciled?</i>	33
<i>The Impact of Whole Language in Australia</i>	36
<i>Whole Language Philosophy in Practice</i>	37
<i>Semantic, Syntactic and Graphophonic Cues</i>	37
<i>Practices Recommended in Whole Language Programs</i>	39
<i>Assessment Techniques Used in Whole Language Classrooms</i>	40
<i>Providing Corrective Feedback</i>	41
<i>Whither Whole Language?</i>	43

CHAPTER FOUR: PHONOLOGICAL PROCESSING RESEARCH

<i>Phonemic Awareness: What Does It Mean?</i>	46
<i>Rhyme and Alliteration</i>	47
<i>Onsets & Rimes</i>	48
<i>Phoneme Awareness</i>	50
<i>Phonemic Awareness: Its Relationship to Reading Development</i>	51
<i>Correlational Studies</i>	51
<i>Training Studies</i>	52
<i>Linking Phonemic Awareness to Reading</i>	53
<i>Other Phonological Abilities</i>	55
<i>The Consequences of Phonemic Unawareness</i>	60
<i>Early Identification and Intervention</i>	61

<i>Research and Education Systems</i>	62
---	----

CHAPTER FIVE: THE DIRECT INSTRUCTION MODEL

<i>Characteristics of the Direct Instruction Model</i>	64
<i>Roots of the Direct Instruction Model</i>	66
<i>Evaluation of the Direct Instruction Model</i>	67
<i>Criticisms of Direct Instruction</i>	69
<i>Problems of Acceptance in Education</i>	70

CHAPTER SIX: RATIONALE AND RESEARCH QUESTIONS

<i>Rationale</i>	74
<i>Research Questions</i>	75

CHAPTER SEVEN: METHOD

<i>Participants</i>	84
<i>Testing Procedure</i>	85
<i>Teachers</i>	85
<i>Measures</i>	86
<i>Construct: Phonemic Awareness</i>	86
<i>Construct: Phonological Recoding in Lexical Access</i>	88
<i>Construct: Phonological Recoding in Working Memory</i>	89
<i>Construct: Decoding</i>	91
<i>Construct: Spelling Ability</i>	92
<i>The Corrective Reading Program</i>	93
<i>Selection</i>	93
<i>Program Design</i>	97
<i>Data Analysis</i>	98

CHAPTER EIGHT: RESULTS

Objectives	100
Descriptive Statistics	100
<i>Reading Disability Criterion</i>	101
Multivariate Analyses	102
Univariate Analyses	103

<i>Test Of Phonological Awareness (TOPA)</i>	104
<i>Word Attack</i>	105
<i>Picture Naming Test</i>	106
<i>Digit Span</i>	107
<i>Brigance Spelling</i>	108
How Widespread are the Effects?	109
<i>Effect Size Calculation by Sex</i>	110
<i>Effect Size Calculation by Age</i>	110
<i>Effect Size Calculation by School</i>	111
<i>Individual Differences in the Effects</i>	112
<i>Orana: A Special Setting</i>	113
<i>Results for Students Who Participated in Consecutive Programs</i>	114
Are There Differences Between the Program Levels - A, B?	114
<i>Level A and Level B Separately</i>	115
<i>Effect Size Calculation by Program</i>	118
<i>The Outcomes For the Level A Program</i>	118
<i>The Outcomes For the Level B Program</i>	119
<i>TOPA (Level A)</i>	120
<i>TOPA (Level B)</i>	121
<i>Word Attack (Level A)</i>	122
<i>Word Attack (Level B)</i>	123
<i>Picture Naming Test (Level A)</i>	124
<i>Picture Naming Test (Level B)</i>	125
<i>Digit Span (Level A)</i>	126
<i>Digit Span (Level B)</i>	127
<i>Brigance Spelling (Level A)</i>	128
<i>Brigance Spelling (Level B)</i>	129
Are the Effects of Educational Value? Examining	
Results in Relation to Test Norms	130
<i>TOPA (Level A & B)</i>	130
<i>Word Attack (Level A & B)</i>	131
<i>Picture Naming Test (Level A & B)</i>	132
<i>Digit Span (Level A & B)</i>	132

<i>Brigance Spelling (Level A & B)</i>	133
<i>Normed Graphs for Level A and Level B Separately</i>	134
<i>TOPA (Level A)</i>	134
<i>TOPA (Level B)</i>	135
<i>Word Attack (Level A)</i>	135
<i>Word Attack (Level B)</i>	136
<i>Picture Naming Test (Level A)</i>	137
<i>Picture Naming Test (Level B)</i>	137
<i>Digit Span (Level A)</i>	138
<i>Digit Span (Level B)</i>	139
<i>Brigance Spelling (Level A)</i>	139
<i>Brigance Spelling (Level B)</i>	140
What is the Relationship Between the Measured Phonological Variables?.....	141
<i>Correlations</i>	141
<i>Regression Analyses</i>	141
<i>Principal Components Analyses</i>	142
<i>Control Group Alone</i>	143
<i>Experimental Group Alone</i>	144
Is Success in the Corrective Reading Program	
Predicted by Any of the Pretest Scores?.....	145
<i>Regression Analyses</i>	145

CHAPTER NINE: ADDITIONAL STUDY

Introduction	149
Method	151
<i>The Participants</i>	151
<i>The Program</i>	151
Results	157
<i>TOPA</i>	159
<i>Word Attack</i>	160
<i>Picture Naming Test</i>	161
<i>Digit Span</i>	162
<i>Brigance Spelling</i>	163
<i>Effect Size Calculation for Program</i>	164
Are the Effects of Educational Value? Examining	
Results in Relation to Test Norms	164
<i>TOPA</i>	165
<i>Word Attack</i>	166
<i>Picture Naming Test</i>	167
<i>Digit Span</i>	168
<i>Brigance Spelling</i>	169
Results Summary	169

CHAPTER TEN: DISCUSSION

Summary of Results	170
Theoretical Considerations	172
<i>Phonological Representations</i>	172
<i>Similar Studies</i>	174
<i>Reading and Phonological Awareness: Reciprocal Relationship?</i>	175
<i>The Nature of Phonological Processes and the Program Effects</i>	177
<i>Word Attack</i>	178
<i>Picture Naming</i>	179
<i>Digit Span</i>	180
<i>Spelling</i>	181
<i>How Phonologically Disabled Were These Students?</i>	182

Implications for Practice	183
<i>Program Effectiveness and Individual Differences</i>	183
<i>Treatment Resistance</i>	184
<i>What Are the Limits of Instructional Influence on Progress?</i>	185
<i>Learning Styles</i>	187
<i>Other Program Characteristics and Effectiveness</i>	188
<i>Program Fidelity</i>	190
<i>Teacher Training</i>	190
<i>Within Program Controls</i>	191
<i>Where To For These Students?</i>	192
Considering Results in Wider Contexts	193
<i>Empirically Validated Treatment</i>	193
<i>NICHHD Findings: An Example of Empirically Validated Treatment</i>	194
<i>Social Validity</i>	196
<i>Clinical Significance</i>	197
<i>Another Wider Issue: The Under-identification of</i> <i>Reading Difficulty in Females</i>	198
Methodological Considerations	199
<i>Research Design</i>	199
<i>Internal Validity</i>	199
<i>External Validity</i>	203
<i>Construct Validity</i>	204
<i>Statistical Conclusion Validity</i>	204
<i>Choice of Analyses</i>	205
<i>Null Hypothesis Testing</i>	207
<i>Effect Size Calculation</i>	208
Further Research	209
Concluding Comments	212
 REFERENCES	 213
 APPENDICES	 242
Appendix A. The Picture Naming Test.....	243

Appendix B. Parent and Teacher Questionnaires.....250
Appendix C. DISTAR Library Readers.....252

LIST OF FIGURES

1. Interaction (+SE) between experimental and control group at pre- and posttest for TOPA (A and B combined).....	105
2. Interaction (+ SE) between experimental and control group at pre- and posttest for Word Attack (A and B combined).....	106
3. Interaction (+ SE) between experimental and control group at pre- and posttest for Picture Naming Test (A and B combined).....	107
4. Interaction (+ SE) between experimental and control group at pre and posttest for Digit Span (A and B combined).....	108
5. Interaction (+ SE) between experimental and control group at pre and posttest for Brigance Spelling (A and B combined).....	109
6. Interaction (+SE) between experimental and control group at pre and posttest for TOPA (Level A).....	121
7. Interaction (+SE) between experimental and control group at pre and posttest for TOPA (Level B1).....	122
8. Interaction (+SE) between experimental and control group at pre and posttest for Word Attack (Level A).....	123
9. Interaction (+SE) between experimental and control group at pre and posttest for Word Attack (Level B1).....	124
10. Interaction (+SE) between experimental and control group at pre and posttest for Picture Naming Test (Level A).....	125
11. Interaction (+SE) between experimental and control group at pre and posttest for Picture Naming Test (Level B1).....	126
12. Interaction (+SE) between experimental and control group at pre and posttest for Digit Span (Level A).....	127
13. Interaction (+SE) between experimental and control group at pre and posttest for Digit Span (Level B1).....	128
14. Interaction (+SE) between experimental and control group at pre and posttest for Brigance Spelling (Level A).....	129
15. Interaction (+SE) between experimental and control group at pre and posttest for Brigance Spelling (Level B1).....	130

16. Mean scores for TOPA (A & B Combined).....	131
17. Mean scores for Word Attack (A & B Combined).....	131
18. Mean scores for Picture Naming Test (A & B Combined).....	132
19. Mean scores for Digit Span (A & B Combined).....	133
20. Mean scores for Brigance Spelling (A & B Combined).....	134
21. Mean scores for TOPA (Level A).....	134
22. Mean scores for TOPA (Level B).....	135
23. Mean scores for Word Attack (Level A).....	136
24. Mean scores for Word Attack (Level B).....	136
25. Mean scores for Picture Naming Test (Level A).....	137
26. Mean scores for Picture Naming Test (Level B).....	138
27. Mean scores for Digit Span (Level A).....	138
28. Mean scores for Digit Span (Level B).....	139
29. Mean scores for Brigance Spelling (Level A).....	140
30. Mean scores for Brigance Spelling (Level B).....	140
31. Interaction (+ SE) between experimental and control group at pre and posttest for TOPA.....	160
32. Interaction (+ SE) between experimental and control group at pre and posttest for Word Attack.....	161
33. Interaction (+ SE) between experimental and control group at pre and posttest for Picture Naming Test.....	162
34. Interaction (+ SE) between experimental and control group at pre and posttest for Digit Span.....	163
35. Interaction (+ SE) between experimental and control group at pre and posttest for Brigance Spelling.....	164
36. Mean TOPA scores for the <i>100 Lessons</i>	165
37. Mean Word Attack scores for the <i>100 Lessons</i>	166
38. Mean Picture Naming Test scores for the <i>100 Lessons</i>	167
39. Mean Digit Span scores for the <i>100 Lessons</i>	168
40. Mean Spelling scores for the <i>100 Lessons</i>	169

LIST OF TABLES

Table 1: Correlations Reported in Recent Studies.....	58
Table 2: Socio-Economic Indices.....	84
Table 3: Experimental vs Control Group: Mean Raw Scores.....	100
Table 4: Experimental vs Control group: Mean Power Transformed Scores.....	101
Table 5: Correlations between Pretest and Posttest Scores.....	101
Table 6: Tests of Normality And Homogeneity Of Variance.....	104
Table 7: Summary of Linear Regression Analyses for Age, School, and Sex Predicting Word Attack (n = 134).....	110
Table 8: Effect Size (<i>d</i>) Calculation by Sex.....	110
Table 9: Effect Size (<i>d</i>) Calculation by Age.....	111
Table 10: Effect Size (<i>d</i>) Calculation by School.....	112
Table 11: Students whose Raw Score Increased by 1 Standard Deviation or More.....	113
Table 12: Mean Scores for Orana vs. Other Schools (Experimental and Control Combined).....	113
Table 13: Effect Sizes (<i>d</i>) obtained by Orana vs. Other schools (Experimental Only).....	114
Table 14: Effect Sizes (<i>d</i>) for Level A and Level B.....	114
Table 15: Differences in Phonological Processes Between Levels.....	115
Table 16: Experimental vs Control groups: Mean Raw Scores for Level A.....	116
Table 17: Experimental vs Control group: Mean Power Transformed Scores for Level A.....	116
Table 18: Experimental vs Control groups: Mean Raw Scores for Level B.....	117
Table 19: Experimental vs Control group: Mean Power Transformed Scores for Level B.....	117
Table 20: Effect Sizes (<i>d</i>) for Level A and Level B.....	118
Table 21: Tests of Normality and Homogeneity of Variance: Level A.....	119
Table 22: Tests of Normality and Homogeneity of Variance: Level B.....	120
Table 23: Correlations Between Pretest Scores (<i>N</i> = 206).....	141
Table 24: Summary of Simultaneous Regression Analysis for Pretest Variables Predicting TOPA Pretest Score.....	141
Table 25: Summary of Simultaneous Regression Analysis for	

Pretest Variables Predicting Word Attack Pretest Score.....	142
Table 26: Summary of Simultaneous Regression Analysis for Pretest Variables Predicting Spelling Pretest Score.....	142
Table 27: Varimax Rotated Two Factor Solution for Total Sample at Pretest.....	142
Table 28: Varimax Rotated Three Factor Solution for Total Sample at Pretest.....	143
Table 29: Varimax Rotated Three Factor Solution for Control Group at Pretest.....	143
Table 30: Varimax Rotated Three Factor Solution for Control Group at Posttest.....	144
Table 31: Varimax Rotated Three Factor Solution for Experimental Group at Pretest.....	144
Table 32: Varimax Rotated Three Factor Solution for Experimental Group at Posttest.....	144
Table 33: Summary of Multivariate Multiple Regression for Pooled Pretest Scores Predicting Gains in the Experimental Group ($n = 134$).....	145
Table 34: Details of Multivariate Multiple Regression for Pooled Pretest Scores Predicting Gains in the Experimental Group ($n = 134$).....	146
Table 35: Summary of Simple Linear Regression for Each Pretest Score Separately Predicting Each Gain for Experimental Group ($n = 134$).....	147
Table 36: Summary of Stepwise Regression Analysis for Program and Pretest Scores Predicting Word Attack Posttest scores ($N = 206$).....	148
Table 37: Summary of Stepwise Regression Analysis for Program and Pretest Scores Predicting Word Attack Gains Scores ($N = 206$).....	148
Table 38: Distar Orthography.....	153
Table 39: Script for Blending Activity.....	155
Table 40: Experimental vs. Control Groups: Mean Raw Scores for <i>100 Lessons</i>	157
Table 41: Experimental vs. Control group: Mean Power Transformed Scores for <i>100 Lessons</i>	158
Table 42: Tests of Normality and Homogeneity of Variance for <i>100 Lessons</i>	159
Table 43: Effect Sizes (d) for <i>100 Lessons</i> Group.....	164

ABSTRACT

This thesis examines the effects of phonics-emphasis Direct Instruction reading programs on the phonological processes of students with teacher-identified reading problems in nine northern and western Melbourne primary schools. The students (131 males and 75 females, mean age 9.7 years, standard deviation 1.2 years) were assigned to the treatment condition or to wait-list comparison groups. Based on the results of a program placement test of rate and accuracy, students were assigned to one of two entry points into the *Corrective Reading* program (A, B1). The students in the intervention group received 60-65 lessons (in groups of five to ten students) from teachers at their schools, or, for some students, at a resource centre for surrounding schools. An additional study, with younger (mean age 8.8 years) less advanced readers involved a similar design and teaching approach. The program, *Teach Your Child to Read in 100 Easy Lessons*, was presented to thirteen students in two settings.

When compared with a similar cohort of wait-list students, the students in each program made statistically significant and educationally important gains in such phonologically-based processes as word attack, phonemic awareness, and spelling; and, statistically significant gains, of at least moderate effect size, in phonological recoding in lexical access and phonological recoding in working memory.

A further question involved the prediction (from pretest scores) of those students who would not make progress in word attack solely from the reading programs. In this thesis, only the presence or absence of the reading programs predicted improvement in word attack.

The studies in this thesis contribute to the long-standing debate on how best to ensure that children learn to read; to the understanding of the relationship between phonemic awareness and reading; to an understanding of the effects of the current system on at-risk children; and, how additional or alternative approaches more attuned to the findings of reading research may improve the effectiveness of the system.

CHAPTER ONE: INTRODUCTION

Problem

There is a significant problem with the attainment of literacy in Australian schools. The Australian Government House of Representatives Enquiry (1993) estimated that between 10-20% of students finish primary school with literacy problems. In Victoria, as many as 16% have been labelled reading disabled (Prior, Sanson, Smart, & Oberklaid, 1994; Richdale, Reece, & Lawson, 1996). Further concern has been expressed that, after their Year Three at school, students with reading problems have little prospect of adequate progress (Australian Government House of Representatives Enquiry, 1993). Providing additional foundation for that fear was a Victorian study (Hill, 1995) that noted little discernible progress in literacy for the lowest decile between Year Four and Year Ten. In the adult population, at least 10% cannot read the telephone book, follow a medical claim form, or write an application for a job (International Year of Literacy Brochure, 1990, cited in Cairney, Lowe, McKenzie, & Petrakis, 1993). Further, the economic costs of low levels of literacy in Australia have been estimated at \$6.5 billion annually (DEET, 1991). Although schools are now expected to achieve a more difficult objective - literacy for all, rather than literacy solely for a sub-class as in the past, there is growing concern that our society is far from achieving that objective.

Instructional Methods

Methods of teaching literacy have been the subject of long (often acrimonious) dispute. Much of this debate has centred on the degree to which children need to have an understanding of the structure of an alphabetic language in order to become skilled readers. Whereas, some have strongly supported this as of central importance in beginning reading instruction (phonics emphasis), others perceive the recognition of whole words as the more productive strategy, and hence have a whole word (or meaning) emphasis in their instructional approach. This history is described in some detail in Chapter 2.

In Australia, there has been wholesale adoption of one particular model of literacy development, known as whole language (Australian Government House of Representatives Enquiry, 1993). The whole language model is a particular example of the class of approaches that adopt a whole word (or meaning) emphasis. It is argued in Chapter 3 that the whole language model does not constitute a comprehensive approach to reading instruction, is not consistent with what is known about the reading process and how children learn to read, and

contains assumptions and practices that are demonstrably unhelpful, even obstructive, to progress for some students.

Causes of Reading Failure

Over the last ten to twenty years there has been increasing consensus on the cause(s) of reading success and failure. The area most cited involves phonological processes, and particularly, phonemic awareness. See Chapter 4 for a thorough analysis of that research, a summary of which follows. Reviews by Hurford, Darrow, Edwards, Howerton, Mote, Schauf, and Coffey, 1993; and Mann, 1993 have made it clear that the presence or absence of phonemic awareness predicts the future membership of good/bad reader categories, and discriminates good readers from poor readers. The avalanche of consistent findings led prominent researcher, Marilyn Jager Adams (1991) to enthuse “To my mind the discovery and documentation of the importance of phonemic awareness ... is the single most powerful advance in the science and pedagogy of reading this century” (p. 392).

Phonemic Awareness: The Research

Stanovich (1986) defined phonemic awareness as the “conscious access to the phonemic level of the speech stream, and some ability to manipulate cognitively representations at this level” (p. 362). Tasks used to assess shallow phonemic awareness tend to emphasise sensitivity to rhyme and alliteration; whereas, a more complex task might involve the manipulation or separation of sounds in a word. A further developmental advance involves a progressive reduction in the size of the unit comprehended - from whole word, to syllables, to intra-syllabic units, to individual phonemes.

The dramatically increased interest in this area is unsurprising given the finding that phonological abilities (of which phonemic awareness is a subset) are the most powerful predictors of reading success - better than more general cognitive abilities such as intelligence, vocabulary, and listening comprehension (Adams, 1990; Bradley & Bryant, 1983; Juel, 1988; Wagner & Torgesen, 1987; Yopp, 1988). There have been many correlational studies (see Wagner & Torgesen, 1987 for a review) that support this link; however, such studies cannot provide evidence of causality. Another problem for such correlational studies, argued Felton (1992), is their facility for predicting good reading outcomes, but inability to shed light on just which children will not make progress.

In addition to the correlational evidence, there have now been a number of longitudinal training studies showing that the relationship between phonemic awareness and reading progress is indeed causal. This latter finding is of great significance, for without it

phonemic awareness may be simply a consequence of reading development, or alternatively, merely a proxy for a third variable such as intelligence, or social class. The most famous of these studies, presented in Bradley and Bryant's seminal paper in 1983, was described by Coltheart (1983) as the "first clear evidence of the mental procedures important in the early stages of learning to read" (p. 421). The authors were interested in whether high levels of phonemic sensitivity were associated with later reading success, and low levels associated with reading difficulty over the following four years. They were able to demonstrate high correlations between the original sound categorisation scores and students' reading and spelling more than 3 years later. Selecting 65 of the students with low phonemic awareness scores, Bradley and Bryant randomly assigned them to either a training group, or a non-training group. The first group was taught (in 40 sessions over two years) to attend to the sound structure of words, while the second was taught to categorise words in terms of their meaning. The children received normal reading instruction in school, and at the end of the project were re-assessed. The training group had made significantly more progress in reading, an effect specific to reading, as the two groups were similar in a standardised mathematics test. Bradley (1990) retested the original experimental and control groups 5 years after the training was completed. Remarkably, the differences were still present in all four reading and spelling tests.

Subsequent intervention studies (Ball & Blachman, 1988, 1991; Byrne & Fielding-Barnsley, 1989, 1990, 1991, 1993, 1995; Cunningham, 1990; Hatcher, Hulme, & Ellis, 1994; Lundberg, Frost, & Petersen, 1988; Tangel & Blackman, 1992) obtained similar results, and those that employed follow-up have noted the endurance of the effects. In a number of these studies, the teaching of phonemic awareness has occurred in conjunction with letter-sound instruction, a process described by Hatcher et al. (1994) as a "phonological linkage" (p. 42). Children in dual-input programs demonstrate more improvement in reading and spelling than those exposed to a solely oral phonemic awareness program. Thus, it has been demonstrated that phonemic awareness is amenable to environmental manipulation to the benefit of students at-risk. Given the claim that phonemic insensitivity is at least partly an inherited problem (Flowers, 1993; Olson, Wise, Conners, Rack, & Fulker, 1989; Rack, Hulme, & Snowling, 1993), it is very important to discover the degree to which intervention can ameliorate such a deficit. If the deficit proved relatively impervious to intervention, then an argument could be mounted for an emphasis on a different mode of word identification, such as a purely visual strategy.

Students entering first grade with little phonemic awareness have less success in reading than peers who enter school with a conscious awareness of the sound structure of words and the ability to manipulate those sounds in words (Adams, 1990; Liberman & Shankweiler, 1985; Mann & Brady, 1988; Spector, 1995; Stanovich, 1985, 1986, 1988a, 1988b; Wagner, 1988). Presumably the reason for this advantage lies with the manner in which phonemic awareness provides a signpost to beginning readers that there is a logic to the reading process (Liberman, Shankweiler, & Liberman, 1989). The recognition of this logic and the capacity to make use of such recognition in beginning reading implies the attainment of the alphabetic concept.

Other Phonological Processes

Phonemic awareness is part of a larger construct in coding and retrieving verbal information known as phonological processing (Hurford, Darrow, Edwards, Howerton, Mote, Schauf, & Caffey, 1993; Smith, Simmons, & Kameenui, 1995; Vellutino & Scanlon, 1987; Wagner, 1986, 1988; Wagner & Torgesen, 1987). Results from this wider research area further indicate that deficits in processing the phonological features of language explain a significant proportion of beginning reading problems, and correlated difficulties in reading comprehension, background knowledge, memory, and vocabulary differences (Liberman & Shankweiler, 1985; Mann & Brady, 1988; Rack, Snowling, & Olson, 1992; Torgesen, Wagner, Simmons, & Laughon, 1990; Wagner & Torgesen, 1987). The research in this area to date is less far advanced but there are suggestions that deficits in naming speed, and short term auditory memory may further hinder some students, and may even make progress in a phonemic awareness training program difficult (Brady, Fowler, Stone, & Winbury, 1994; Torgesen, Wagner, & Rashotte, 1994). Wood and Felton (1994) consider naming speed to be the best predictor of the progress of reading impaired students in a reading intervention; and a number have suggested that naming speed may be a critical limiting component in learning to read successfully (Brady, 1991; O'Connor, Jenkins, Cole, & Mills, 1993; Wolf, 1991). Such a speculation suggests the possibility of using naming speed pretest scores to predict which students will make greater or lesser progress in a reading program. Such an hypothesis will be tested in this thesis.

In summary, this research indicates that difficulties with awareness, coding, and retrieval of the sounds in words are critical impediments to reading development. Whether these three phonological processes are independent, or elements of a more general process are as yet unresolved. Elbro, Nielsen, and Petersen (1994) argue that poor phonological

representations of words form the core deficit in disabled readers. In this view, lexical access and working memory are restricted not because of specific modular deficits in these processes, but rather because what is sought in the lexicon, or to be held in working memory, is lacking in readily distinguishing features. They noted the confusion of similar sounding words, and the less distinct word-naming in such readers. This view also finds support in a study by Eden, Stein, Wood, and Wood (1995). The phonological representation explanation allows for the possibility that improved phonemic awareness may lead to an assessed improvement in one or more of these other phonological processes. In fact, Rubin, Rottella, Schwartz, and Bernstein (1991) found that training Year 3 children in phonemic awareness had a significant beneficial effect on the picture naming speed of both the good and poor readers. In this thesis, pretest and posttest measures of phonological processes will provide further information about this possibility.

Phonemic Awareness and Reading: The Relationship

Establishing a causal relation between phonological awareness and reading acquisition does not preclude other directional relations. Some have argued that phonemic awareness is a consequence of learning to read rather than a causal factor (Morais, Alegria, & Content, 1987). The evidence that phonological awareness is developed by reading instruction and the act of reading arises from several sources:

(a) reviews of studies with skilled readers in non-alphabetic languages (Huang & Hanley, 1994),

(b) studies with adult illiterates in alphabetic languages (Lukatela, Carello, Shankweiler, & Liberman, 1995; Morais, 1991), and

(c) controlled studies (Bentin, Hammer, & Cahan, 1991; Bentin & Leshem, 1993).

Some (Bentin & Leshem, 1993; Perfetti, Beck, Bell, & Hughes, 1987) have argued that exposure to reading instruction is the catalyst for the development of phonemic awareness. More properly this assertion should include the caveat that it is *successful* instruction (rather than mere exposure) which may trigger phonemic awareness, as unsuccessful readers typically demonstrate continued deficits in this area. Increasingly, there is acceptance that the relationship between phonemic awareness and reading development is a reciprocal one, in which shallow forms of phonemic awareness enhance progress in early reading, and this progress stimulates the development of deeper phonemic awareness, that is, at the phoneme level (Adams, 1990; Bentin, & Leshem, 1993; Stanovich, 1985; Vellutino & Scanlon, 1987; Wagner & Torgesen, 1987). On the other hand, Bruck (1992) found that

phonemic awareness deficits in dyslexic readers are very persistent even if reading ability improves, presumably through reliance on orthographic and other strategies (Hulme & Snowling, 1992). Hence, it may be that the reciprocal relationship between reading and phonemic awareness does not hold for all students. Some students, perhaps those usually labelled dyslexic, may be highly resistant to developing phonemic awareness despite reading instruction. Indeed, Berninger and Abbott (1994) view resistance to validated treatment interventions as the distinguishing feature of learning disabilities such as dyslexia.

It is now apparent that one can enhance phonemic awareness skills through the implementation of a dedicated phonemic awareness program. However, it is not clear whether this represents the only means of achieving the objective. Perhaps a reading program that draws attention to the relationship between written word parts (including graphemes) and oral word parts (including phonemes) may promote the growth of phonemic awareness without the application of a dedicated phonemic awareness program. Standing against this speculation is the argument that unless a student has an understanding of the structure of oral language *first*, then focussing on such structural issues in the written form may be ineffective (Juel, 1993; Lindamood, 1994; Simner, 1995; Torgesen, Morgan, & Davis, 1992). Yet it may be possible that both the alphabetic principle and phonemic awareness can be evoked through a phonics-emphasis reading program, carefully taught (at least for those students with some lower threshold level of phonemic awareness). The question is probably best addressed as an empirical one, as there are still a number of unresolved theoretical issues relating to phonemic awareness. A major focus of this thesis is the extent to which such a program does increase phonemic awareness - both as an aural/oral skill, and as it is applied to the task of reading.

Phonemic Awareness and Older Students

It is as yet unclear what implications the phonemic awareness research has for older children who struggle with reading. It may be that there is an upper threshold level of phonemic awareness (O'Connor, Notary-Syverson, & Vadasy, 1996) beyond which there is no advantage for reading development in attempting its enhancement. Indeed, it is possible that, for older children, phonemic awareness is no longer the appropriate focus, as students may be more in need of orthographic rather than phonemic strategies. Share (1995) argues, however, that without the induction of the alphabetic principle, skilled reading (implying the use of a generative strategy capable of decoding novel words) will not occur. His view is supported by the finding that dyslexic adult readers (even those with strong orthographic capacities) still demonstrate phonemic awareness deficits, and struggle to decode novel words

(Bruck, 1992; Hulme & Snowling, 1992; Pratt & Brady, 1988; Siegel, 1993; Solman & Stanovich, 1992). This thesis involves the participation of students in mid and upper primary school to help shed light upon this issue.

In order to understand why older readers can still benefit from instruction aimed at developing decoding strategies that will become less and less used as their level of reading skill improves, it is necessary to examine models of reading development.

Models of reading development

A number of researchers have developed models of reading development based on stages (Chall, 1979; Ehri, 1993, 1994; Frith, 1985). Although variations occur among writers, there is increasing general acceptance among empirical researchers that the sequence of development of the word identification system moves from logographic to alphabetic to orthographic. In the first stage, the beginning reader learns to recognise a visual pattern by its shape (a letter landscape). The shape is recognised wholistically, and significant alterations to the letter structure may be made without altering the child's response (e.g., McDonalds, Pepsi-Zepsi, etc.).

At this stage, the child has not learned to analyse the written word structure, and would not need to if our written language were logographic. It is, however, alphabetic, and contains far too many words to be recognised by the visual pattern of peaks and troughs, whirls and intersections that comprise our written language.

The movement to the alphabetic stage is probably driven by the gradual awareness of speech segmentation which the child induces or is taught (Adams, 1990). This phoneme awareness may more readily be invoked in children whose earlier experiences have included a focus on the structure of the spoken word, albeit in larger units such as rhymes, syllables, onset and rimes. Some children do not develop this awareness unaided (Chall, 1989) and without assistance may remain at this early stage (Spear-Swerling & Sternberg, 1994), reliant on memory of the letter landscapes, or contextual guessing strategies. Such readers are doomed as the demands of a rapidly increasing visual vocabulary become overwhelming in middle to upper primary school, that which Share and Stanovich (1995) term "an orthographic avalanche" (p. 17).

In the alphabetic stage, simple letter pattern-to-sound conversion provides a means of decoding (albeit, laboriously) unknown words. Initially this may involve use of only partial letter-sound cues (Spear-Swerling & Sternberg, 1994) until, with the arrival of alphabetic insight (Byrne, 1991), this strategy becomes reliable, at least with regular words, and

continues to provide some clues for irregular words (Goulandris & Snowling, 1995). In irregular words, it is vowels that provide the quality of irregularity, but consonants remain regular for the most part, and it is the consonants that are most important in word recognition (Share & Stanovich, 1995). Hence, this phonological recoding strategy enables cues for decoding most words along the regular-irregular continuum.

Share (1995) sees this alphabetic period as crucial, and he developed a self-teaching hypothesis in which “... each successful decoding encounter with an unfamiliar word provides an opportunity to acquire the word specific orthographic information that is the foundation of skilled word recognition and spelling” (Share & Stanovich, 1995, p. 18). This gradual “lexicalization” (p. 18) occurs through repeated opportunities to use letter-sound correspondences for decoding. The strategy is used with less frequency as the range of familiar word patterns increases, through a “self-teaching” (Share, 1995, p. 155) mechanism. The phonological recoding strategy remains useful for decoding unfamiliar words - and of course, our language has many low frequency words. Eighty percent of English words have a frequency of less than one in a million (Carroll, Davies, & Richman, 1971, cited in Share & Stanovich (1995). Thus, the phonological recoding mechanism has a usefulness beyond its initial ability to provide the opportunities for the formation of orthographic representations. Share and Stanovich (1995) assert that orthographic strategies are developed through multiple examples of success in decoding phonologically. If one accepts this view, then orthographic strategies should not be taught directly, and the instructional emphasis for older students must still be placed on ensuring letter-sound correspondences, blending and segmenting, and practice. It may also be that only through such laborious serial letter-by-letter decoding can precise letter-order become entrenched in the orthographic representation that forms the basis for accurate spelling (Adams, 1990; Jorm & Share, 1983; Williams, 1991). However, since many different words share similar spelling patterns, practice on any one word may simultaneously enhance the recognition of other similar words. It is this facility, known as decoding-by-analogy, that helps explain the capacity of readers to develop a large reading vocabulary so quickly.

Dyslexics: A Special Subgroup?

There has been concern expressed in the literature that dyslexics may be irretrievably insensitive to phonemes (Bruck, 1990, 1992; Pratt & Brady, 1988; Torgesen, Wagner, & Rashotte, 1994), and thus unable to derive benefit from sounds-based programs. However, a study by Alexander, Anderson, Heilman, Voeller, and Torgesen (1991) with a group of (93-

154 months) dyslexic students noted significant improvement in phonemic awareness and phonological recoding following instruction in the Auditory Discrimination in Depth Program (Lindamood & Lindamood, 1969). In another study (Lovett, Border, De Luca, Lacerenza, Benson, & Brackstone, 1994), dyslexic children (average age 114 months) were able to demonstrate significant gains in phonological processing (in both speech and print) using a variant of direct instruction procedures. Although no attempt has been made to identify dyslexics in this present thesis, a figure of 4% of the population is sometimes used (DSM-IV, 1994), although the methods of identification vary significantly. Given the assumption that the schools involved in this current study are representative of the population, one in which an estimated 16% (Prior et al., 1995) have significant reading difficulty, then perhaps one in four students in the present studies might have been identified as dyslexic. It is of interest then to consider whether any identifiable (from pretest scores) subgroups of the treatment cohort failed to make progress in the reading program, or whether the program effect was sufficiently strong to overcome individual deficits. Foorman and Francis (1994) noted that when blending and segmenting are suitably incorporated into a code-emphasis program, individual differences in these skills (which are usually predictive of reading success or failure) disappeared. That is, instruction had led to the disappearance of individual differences in this skill. This thesis may add to the research on this question.

Share (1995) points out that struggling readers tend to rely more on non-phonological strategies such as context, or whole word recognition, or partial visual cues. These strategies are non-generative, and do not assist skilled reading to eventuate. The replacement of these strategies with a core of phonological recoding skills is not an easy task - all the more difficult as the student grows older, and ineffective strategies become more deeply entrenched. Share (1995) is adamant that “... there can be no case of competent reading in the absence of functional decoding” (p. 173). There is then theoretical evidence that decoding strategies are of primary importance to all non-facile readers, and empirical evidence that they can also be taught to older struggling readers, using programs that make explicit the connections between sounds and letters.

The relationship between phonics and phonemic awareness is often misunderstood. Phonemic awareness is an aural/oral skill that (at least in part) can exist without contact with print. At an advanced level, it involves the capacity to dissect the spoken word, and manipulate the resultant sound segments. Until contact with writing however, there is no communicative value in developing such a skill, and many children do not routinely pay

attention to these meaningless segments of speech, and hence do not develop this capacity. Other children become fascinated with rhymes and alliteration, Pig Latin, Spoonerisms, and enjoy inventing words - constructing them from speech segments. A lack of phonemic awareness alone cannot be classified as a primary language deficit, as it is unnecessary for oral communication, and only becomes evident if one is confronted with the reading task.

When print is encountered, the capacity to perform the phonemic operations described above becomes very important. In order to develop the alphabetic principle (that units of print map on to units of sound), students must already have (or soon develop) phonemic awareness. It is the alphabetic principle that allows students to move beyond the early logographic stage of reading in which each word is a unique, indivisible shape to be recognised visually. Memory constraints make that a strategy of limited usefulness as it does not assist students to decipher words not before seen and memorised. It is the understanding of the alphabetic principle that allows students to decipher such novel words. Using the alphabetic principle as the cipher represents what Perfetti (1991) calls a productive process, in contrast to the highly limited memorisation process. Share (1995) sees the phonological recoding process as critical to the development of skilled reading, and describes it as being "... a self-teaching mechanism, enabling the learner to acquire the detailed orthographic representations necessary for rapid, autonomous, visual word recognition" (p. 152).

Many students enter school with little phonemic awareness (Adams, 1990), and exposure to any one of a variety of forms of reading tuition may be sufficient to stimulate such awareness for them, thus making the alphabetic principle more readily accessible. However, in an unacceptably high number of students this process does not occur. The aim of phonics teaching in a code-emphasis program is to make explicit to students the alphabetic principle. In a whole language classroom, in which phonics is viewed as one (subsidiary) strategy among others, to be used when the prediction-confirmation strategy breaks down, there is likely to be considerably less emphasis on student mastery of this principle. Teachers may point out word parts to students in the context of authentic literature as the situation arises, but the limitations of such incidental phonics may impact most heavily on at-risk students (Simner, 1995). It seems that all phonics are not equal. It is possible to teach phonics carefully, and with parsimony; it is possible to do so ineffectively and excessively; and it is possible to do it in name only. The major problem for at-risk students, argued by Byrne (1996) involves the risk for such learners of failing to be explicit and unambiguous.

It might be prudent to tell children directly about the alphabetic principle since it appears unwise to rely on their discovery of it themselves. The apparent relative success of programs that do that (Bradley & Bryant, 1983; Byrne & Fielding-Barnsley, 1991, 1993, 1995) support the wisdom of direct instruction. (p. 424)

Similar sentiments have been expressed by a number of researchers in recent years (Adams & Bruck, 1993; Baker, Kameenui, Simmons, & Stahl, 1994; Bateman, 1991; Blachman, 1991; Felton & Pepper, 1995; Foorman, 1995; Foorman, Francis, Beeler, Winikates, & Fletcher, 1997; Moats, 1994b; Simmons, Gunn, Smith, & Kameenui, 1995; Singh, Deitz, & Singh, 1992; Spector, 1995; Tunmer & Hoover, 1993; Weir, 1990). Consensus remains to be achieved regarding the details of the strategies best able to ensure the understanding of the alphabetic principle; however, the cited authors are of the belief that (for some learners at least) direct instructional approaches are more likely to be successful.

The most common reading problem among students with reading difficulty is at the level of word recognition and decoding (Indrisano & Chall, 1995; Share & Stanovich, 1995). Hence, programs that emphasise such skills are (unsurprisingly) more effective than those that focus on meaning, as in the whole language approach (Stahl & Miller, 1989, 1995; Vellutino, 1991). A well designed and presented phonics-emphasis program may have the effect of boosting at-risk students' phonemic awareness because of its emphasis on word structure, and also their capacity to decode novel words - a marker of the presence of the alphabetic principle. There have been studies that have examined this question, including that by Williams (1980) in which she supplemented a school-based reading program for reading disabled students with a blending and segmenting procedure. The program was successful in increasing phonemic awareness and decoding skills in such children across an age range of 7 to 12 years. These skills transferred to the decoding of unlearned words, such as pseudowords. Wallach and Wallach (1979, as cited in Williams, 1991) obtained similar results in a tutoring program based on the same principles. Williams (1991) decries the minimal impact of such research on instructional activities, arguing the need for clearly delineated and adequately designed blending and segmentation training in reading programs.

The *Corrective Reading* program (Engelmann, Hanner, & Johnson, 1988; Engelmann, Johnson, Carnine, Meyer, Becker, & Eisele, 1988), based on principles similar to Williams (1980), has been evaluated many times, with consistently good results, especially with at-risk students, although the outcome measures have usually emphasised broad reading assessment

measures, rather than focussing on phonological processes. This thesis examines whether the use of the *Corrective Reading* program enables improved outcomes in these areas.

Instructional Issues

The content of reading instruction forms one element of the environmental contribution to the development of reading capacity. Another area concerns the manner in which any given content is delivered to students. Research on learning disabled, intellectually disabled, and disadvantaged children has demonstrated that not all students respond equally to instruction. Forms of instruction that are adequate for some students may not be for others (Adams, 1991; Tunmer & Hoover, 1993; Yates, 1988). Hence, an area that is receiving increasing attention is that of the quality of reading instruction (Adams, 1990; Felton, 1993) students receive. There is a strong argument that reading is not a natural process (as speech appears to be), although whole language advocates argue for its equivalence (Lieberman & Lieberman, 1990). The unacceptably high rate of illiteracy supports the reading-as-unnatural view, and schools can not afford to assume that phonemic awareness will develop in all children solely through exposure to literature (Adams, 1990; Cantwell & Rubin, 1992) that is the major conduit to such awareness provided in a whole language classroom (Iverson & Tunmer, 1993; Read, 1991). The question arises as to the best way to assure phonemic awareness development occurs, especially in at-risk students.

Fortunately, there is a strong literature on effective teaching (reviewed in Chapter 5) that provides an appropriate vehicle for delivering to students the content now known to be central to reading success. A number of recent studies have employed such a model known as direct instruction in successfully teaching phonemic awareness skills (Cunningham, 1990; Felton, 1993). A range of researchers have reached a similar conclusion about the need to emphasise direct instruction teaching principles in providing initial and remedial reading instruction to at-risk students (Baker, Kameenui, Simmons, & Stahl, 1994; Bateman, 1991; Blachman, 1991; Felton & Pepper, 1995; Foorman, 1995; Moats, 1994a; Perfetti, 1992; Spedding & Chan, 1993; Stanovich, 1994; Tunmer & Hoover, 1993; Weir, 1990; Wood & Felton, 1994). Direct instruction is an approach to teaching that is quite different to the child-centred, whole language model. It involves high levels of student time-on-task, goals that are made explicit to students, sufficient time allowed for instruction, extensive content coverage, careful monitoring of progress, and attention to lesson pacing, many low level questions that ensure a high proportion of correct responses, and feedback that is prompt and academically oriented. The major features of such explicit instruction are: (a) teaching in small steps, (b)

providing guidance during initial practice, (c) having students practise after each step, and (d) ensuring a high level of success (Rosenshine, 1986).

The Corrective Reading Program

A direct instruction remedial reading program with strong empirical support is known as *Corrective Reading: Decoding* strand (Engelmann, Hanner, & Johnson, 1988; Engelmann, Johnson, Carnine, Meyer, Becker, & Eisele, 1988). Numerous studies (Branwhite, 1983; Campbell, 1983; Clunies-Ross, 1990; Gregory, Hackney, & Gregory, 1982; Holdsworth, 1984; Kassandra & McQuaid, 1987; McLean & Moore, 1985; Maggs & Murdoch, 1979; Noon & Maggs, 1980; Polloway & Epstein, 1986; Polloway, Epstein, Polloway, Patton, & Ball, 1986; Sommers, 1995; Thorne, 1978) attest to its effectiveness in improving at-risk readers' performance on a range of standardised assessments. The author had noted (Hempenstall, 1988) that, in evaluations performed in schools of the *Corrective Reading* program over a number of years, students consistently made substantial gains in the Word Attack subtest of the Woodcock Reading Mastery Tests (Woodcock, 1973). The reading of pseudo-words, such as is assessed by the Word Attack subtest, is regarded by many as the best means of ascertaining the extent to which students can use their phonemic awareness in deciphering words never before seen - words unavailable to contextual or memorisation strategies (Elbro, Nielsen, & Petersen, 1994; Hoover & Gough, 1990; Scanlon & Tanzman, 1994; Share & Stanovich, 1995; Siegel, 1993; Stanovich, 1988a; Vellutino, Wood & Felton, 1994). It was these findings that led to an interest in whether the *Corrective Reading* program promotes phonemic awareness in both oral and written language in primary aged remedial reading students. It is hypothesised that a well designed and presented phonics-emphasis program should have the dual effect of boosting at-risk students' phonemic awareness because of its emphasis on word structure, and the students' capacity to decode novel words (a marker of the presence of the alphabetic principle). This thesis examines whether the use of the *Corrective Reading* program enables these outcomes. The pretest and posttest results will be compared with those of a group of similar readers who are on a waiting list to participate in the same program.

Phonological Processes and Spelling

As discussed earlier, other phonological processes may also play a part in reading processes; hence, naming speed (Hempenstall, 1995a) as a measure of phonological recoding in lexical access (Torgesen et al., 1994), and Digit Span (Wechsler, 1991) as a measure of phonological recoding in working memory (Catts, 1996) were assessed prior, and subsequent,

to the intervention program. There is less known about the role of these other phonological processes, including how amenable they are to direct or indirect intervention. Several studies have noted improvement in lexical access following phonemic awareness intervention (Beck, Perfetti, & McKeown, 1982; McGregor & Leonard, 1995, cited in Catts, 1996), though no studies thus far have supported the value of directly teaching naming skills.

Lindamood (1994) noted that children who have difficulty in appreciating the sound structure of words tend to be poor spellers. Ball and Blachman (1991) found that, for young children, improved phonemic awareness led to improved spelling. One explanation for this offered by Davidson and Jenkins (1994), and Treiman (1985) is that spelling, at least in part, is indicative of young children's ability to classify speech sounds. Burt and Butterworth (1996) have argued that phonological ability plays an even greater role in spelling than it does in reading, whereas, Stage and Wagner (1992) asserted that older students make less use of phonological processes in spelling than do young students, instead relying more on orthographic representations. It may be that this latter assertion refers only to older, *skilled* readers, and hence is really an assertion about stage rather than age. Thus, it is speculated in this thesis that participation in the *Corrective Reading* program will improve phonemic awareness and spelling (although spelling is not taught directly).

This thesis will address a number of questions that have both theoretical and practical implications. The research questions outlined below refer to a sample of readers referred to as disabled readers. Whilst this is a term used by a number of authors to describe students whose reading development is unsatisfactory, there have been a range of criteria employed to discriminate this group from normally developing readers. Some researchers (Prior, Sanson, Smart, & Oberklaid, 1995) selected students below one standard deviation on a standardised reading test. Others included those students below the 25th percentile (Lovett, Border, De Luca, Lacerenza, Benson, & Brackstone, 1994; Stanovich & Siegel, 1994). Felton (1992) adopted the stricter criterion of the 16th percentile, whilst Vellutino, Scanlon, Sipay, Small, Pratt, Chen, and Denckla (1996) employed the 15th percentile. Some authors reported standard score thresholds - Newby, Recht, and Caldwell (1993) used standard scores below 85 on a word attack test, whilst Lyon and Moats (1997) selected 80 as the standard score upper limit. Employing age-equivalence norms, Lovett and Steinbach (1997) decided upon 1.5 year delay as their standard. In this study, the criterion adopted for the designation reading disabled was any student below the 25th percentile on the Word Attack subtest of the *Woodcock Tests of Reading Mastery* (1987).

CHAPTER TWO: HISTORY OF THE READING DEBATE

A History of English Reading and Writing

There has been considerable recent controversy over the competing emphases to beginning reading known as Whole Language and phonics. In order to provide a broad perspective on the debate, this chapter examines the history of disputes about reading, particularly as they apply to at-risk students. It commences with a brief discussion of the advantages and difficulties of our English alphabetic system, and the literacy problems associated with it. Identification of the major attempts to deal with the complexity of our writing system is followed by a history of the research into the most efficacious means of enhancing reading development. An examination of early research, such as The Great Debate, The USOE Study, Follow Through, and Becoming a Nation of Readers helps illuminate the current debate by indicating which issues are novel, and which are those from the past as yet unresolved.

The current controversy surrounding the extent of literacy failure is not a new phenomenon. Public interest in the issue is certainly at a high level currently, with the introduction of state and nationwide testing, and the possibility of the introduction of minimum standards of acceptable school performance. However, literacy, and the role of schools in promoting it, has had a fiery history in the educational community for almost two hundred years. Unfortunately, there is not a consensus within the education community on the existence, definition, or extent of a literacy problem, and on appropriate methods of solving the problem. This lack of unity leads to a fragmentation of efforts at resolution, precluding the focussed approach necessary to address effectively the systemic dilemma of illiteracy. A major continuing dispute involves the relevance of phonic strategies in beginning reading, and as an overarching theme, the role of educational research in influencing educational policy and practice. An examination of the history of such reading disputes may be useful as it places the current debate in a broader context, and indicates how some contemporary issues are similar, or analogous, to those of earlier periods.

As far as we know, spoken communication has existed for as long as our species has developed relationships. How languages began is unknown - perhaps initially from imitation of the sounds heard in the natural environment, followed by invention of other sounds to encompass the many additional requirements of an intelligent species. There are now at least

3000 different languages spoken in the world, yet the sounds that humans can produce are quite limited, and thus most languages require less than 50 distinct phonemes (Davis, 1988).

Writing Systems: From Logograms to the Alphabet.

Attempts throughout history and across cultures to communicate in a visual format have varied in the style that the messages take, and to a greater or lesser extent these visual systems have met the written communication needs of the time. Apart from the requirement of communicating the author's intent, the system should be capable of expressing the full range of ideas, emotions, and actions for that culture. Furthermore, any system, if it is to be available to the general population (not only for an elite), should be easy to write/draw, and have a manageable number of symbols. It is not surprising then that, over time and over different cultures, a variety of different systems have evolved. The earliest attempts involved pictures (e.g., cave drawings), and at their simplest they were quite effective when the writer had at least rudimentary skills. Complex ideas however were more difficult to draw skilfully and portray unambiguously; in fact, many ideas cannot be portrayed by drawing, for example, democracy. Agreed-upon symbols evolved to overcome this problem, at least within geographical regions, but did not have the universal comprehensibility of, for example, a drawing of a horse. Such symbols (called logograms) are slow to reproduce and, as each is unique, require impressive memory capacity. The Chinese have at least 40,000 logograms (Rayner & Pollatsek, 1989), far more than even a Chinese scholar could manage.

There is some evidence that the skills demanded of readers differ depending on the writing system in use. Huang and Handley (1995) noted that learning to read Chinese required less phonological awareness than did an alphabetic system such as English (and correspondingly greater visual skills). Such findings have important implications for the most appropriate instructional emphasis in initial reading. It is suggestive of the need to ensure the development of phonological awareness in students embarking upon beginning reading in an alphabetic system. In fact, there is now an irresistible body of evidence pinpointing phonological skills as powerfully causal in reading development. An examination of this evidence can be found in Chapter 4, and other recent reviews may be found in Ball, 1993; and in Smith, Simmons, and Kameenui, 1995.

About 4000 years ago the interest in word "look-alikes" shifted to word "sound-alikes". Thus, rather than symbols representing words-in-picture, they could represent the sounds-in-words, initially through the use of syllables. This had obvious advantages in economy because the same syllables appear in many words, and because all words are readily

decomposable into syllables. The Korean written language system is partly a syllabary, containing several thousand syllables, and English contains about five thousand (Adams, 1990). An emphasis on syllables, however, means that each reader and writer must learn and memorise a large number of syllables that have no pictorial meaning, a requirement restricting written communication to an elite. The Korean written language (a 15th Century invention) also incorporated an alphabetic system, in order to avoid this limitation.

As the evolution of written systems continued, the requirement of general accessibility of the written system led to the association of one symbol, or letter, with each basic speech-sound or phoneme. Just as syllabaries reduced the memory load required by logograms, so the use of letters made possible the reproduction of any word in far fewer symbols than were required by syllabaries. This made the task of learning to read and write far more accessible to the general population, and thus the alphabet was recognised as one of the more significant of human inventions.

Problems of Written English

Written communication, which was developed mainly as a means of making inventories of ordnance and cargo, and was available to only a select few, became a means of timeless communication - allowing communication from the prosaic to the profound in content, and with the expectation that almost any person could master its techniques. Unfortunately this expectation has proved rather difficult to fulfil. A problem for an incompletely alphabetic system like English is the lack of a one-to-one correspondence between letters (graphemes) and sounds (phonemes) (Adams, 1990; Rayner & Pollatsek, 1989). This is especially evident with vowels - we have more than a dozen sounds represented by only five letters.

Our oral language has changed markedly from Old English (which was quite regular). Old English was Germanic in origin (Francis, 1965), but new words and sounds have entered our language, mainly from Latin, Greek, and French. These new sounds and sound-combinations have to be encompassed within a print system that is unchanging - thus leading to the irregularities that are the bane of young (and not so young) readers and writers. The rules for letter-sound correspondence do not always provide the means to accurate decoding. It is interesting to speculate whether the opposition to phonic instruction would have been so great had the English language been more regular.

These irregularities have led to several unsuccessful attempts at reforming the alphabet over the past millennium - the most famous being George Bernard Shaw's attempt in

the Nineteenth Century, and the introduction of ITA (Initial Teaching Alphabet) in the Twentieth Century. These attempts are based on the principle that one-to-one correspondence between letter and sound will make phonics instruction more effective, and learning to read easier. Downing (1979, cited in Adams, 1990) reported on a large scale British study that found this to be true for ITA - students were more readily able to develop understanding of the alphabetic principle when taught using the ITA method of reading instruction. The counter-argument to such reform is that, while the ITA's phonemically regularised alphabet may aid beginning readers, it would be at the expense of skilled readers who are able to gain important and meaningful information from the traditional orthography. Under ITA, for example, homophones would have the same spellings, making comprehension more difficult; and intra-word conditional redundancies (an element in skilled word recognition) would be unavailable to the reader. Garner (1962, cited in Gibson & Levin, 1978) argued that letters are more constrained (and thus more predictable) than words. These conditional rules about clusters of consonants, and the allowable number of vowels in a sequence reduce uncertainty, thereby facilitating word recognition. It is the automatic simultaneous activation of intra-word units that distinguishes skilled readers (Roth & Beck, 1987). For example, in a word that begins T(consonant)(vowel) there is a very strong probability that the consonant is "h". Thus, the effect on existing readers of such a reformed orthography would be to decrease reading speed and comprehension for the majority, at least in the short term. Whether such reform would be advantageous overall is irrelevant given that the disadvantages would fall on adults (the decision-makers) who would be required to relearn the reading process.

An alternative strategy is to teach beginning readers the new orthography, and then teach the traditional form as they master the principles of reading. There is some doubt (Crowder & Wagner, 1992; Groff, 1990) whether such reform would be worth the trouble, as longitudinal studies (Rayner & Pollatsek, 1989) that have compared students in traditional and reformed systems found no significant between-group differences by the time that the transfer back to traditional orthography was completed. Perhaps the most enduring outcome of the bold attempt at reform will be the recognition that an early emphasis on learning the alphabetic principle is most efficacious in beginning reading instruction (Chall, 1967).

We are left with an English system of 26 letters and about 45 phonemes that can be spelled in at least 350 ways (Pollack & Pickarz, 1963). The permutations this allows makes both learning to read and teaching beginning readers formidable tasks. Confusion arises from

words that look alike but are sounded differently (*tough, bough, cough, dough*), and words that look different (*mail, male*) but sound alike.

Letter confusions are also common in beginning readers (Rayner & Pollatsek, 1989). It is probably the first time that they have encountered an object whose name changes when it is rotated. A chair is still a chair when it is rotated but a "b" becomes a "d" when rotated about the vertical axis, and a "p" when rotated about the horizontal axis. Letters may be upper-case or lower-case, and in differing script forms. To complicate the issue further, the sounds of individual phonemes are not precisely maintained when they are used in words, they are influenced by the letters around them. Thus, the phoneme /d/ has a different sound when followed by /oo/ than when followed by /i/. The converse is also true - there are not 3 distinct sounds in "cat". Only by a learned process of conscious analysis can one detect these individual phonemes (Stanovich, 1993), and it is a skill that eludes about 30% of first grade readers (Adams, 1990).

Despite the difficulties imposed by our system of writing, the majority of children do learn how to read with at least reasonable proficiency. This is true over time, and across different languages and systems of teaching; however, we are beginning to understand that not all systems lead to equivalent outcomes. The role of teaching, then, is to provide the opportunity, the encouragement, the environment, and the instruction appropriate to beginning readers' needs so that mastery occurs. The approaches adopted by educators have been many and varied, but a major focus has been the degree to which strategies involving intra-word analysis are necessary in the development of reading.

The Problems of Literacy

That literacy is highly valued in a democratic technological society is readily apparent. From enquiries and policies at various levels of government, from media interest, from employer-expressed concerns, through to parental involvement - it is evident that the goal of literacy-for-all is of considerable importance in our society. Recognising what constitutes literacy is rather more difficult, however. Literacy is a set of skills rather than a unitary concept, and people vary in the number of skills, and the degree of their mastery of those skills (Stedman & Kaestle, 1987). For example, the authors distinguish between reading achievement and functional literacy. They define reading achievement as those skills taught and assessed in schools - from learning to read words already in their lexicon through to complex critical and interpretive skills. The overlapping dimension, functional literacy, implies the ability to comprehend written communication outside school - in work, recreation

and general societal participation. While the two dimensions are clearly interrelated, Stedman and Kaestle (1987) estimate that 20 percent of the US population have serious deficits in their functional literacy.

Among the lowest fifth in functional literacy skills are many who are unable to read product labels and have to depend upon brand name logos for selection of items in a grocery store. Many are unable to determine whether they are getting the correct change. Many cannot read recipes very well and cannot read the directions on frozen food packages. (p. 34).

The authors further highlight as problematic for these adults: traffic signs, street names, transport schedules, children's homework, school reports, and emergency phone numbers. Kirsch and Guthrie (1984) suggest that the demands of the workplace for literacy skills is increasing, and jobs without a strong literacy requirement are becoming rare. Similarly in *A Nation at Risk* (National Commission on Excellence in Education, 1983) it was argued that literacy demands are outstripping supply.

In Australia, the Report of the House of Representatives Standing Committee on Employment, Education and Training "The Literacy Challenge" (1993), estimated that a similar number (up to 20%) of children complete primary school with residual literacy problems. An obvious question is whether schools, as they are currently structured, are capable of meeting the increasing literacy requirements of our society. In addition to the increased requirements for literacy, schools are pressured to include many tasks in their curricula which formerly were considered family responsibilities, for example, health issues including sex education, drugs, smoking, fitness, skin cancer, youth suicide. A focus on these questions concerning tasks, resources and methods in education, rather than acrimonious debate on whether standards are declining, encourages a more forward-looking, constructive approach to achieving improved student literacy.

The Teaching of Reading: The Emergence of Meaning-Centred Approaches

The first teachers of reading in English were priests in the Seventh Century. Children were taught the alphabet, syllables, and the Primer, or Prayerbook (Davis, 1973). Most reading was religious, and the ability to read was restricted to relatively few. With the invention of the printing press in the Sixteenth Century, the written word became much more prevalent, although the Bible was the only book available in most homes. Thus, reading was first promoted by religious authorities as a means to one end (salvation), and only later was considered important by governments, as a means to a quite different, secular end - an educated, democratic society.

The phonic technique of teaching component skills and then combining those skills was the norm until the mid-Nineteenth Century (Adams, 1990). It followed a sequence of teaching upper-case and lower-case letter names, two-letter and three-letter combinations, mono-syllabic words, multi-syllabic words, phrases, sentences, and finally, stories. Phonics is an approach to teaching reading that aims to sensitise children to the relationships of the spelling patterns of our written language to the sound patterns of our oral language. It is not a single method, however, as decisions need to be made regarding the timing of its introduction, the method of delivery, whether explicitly taught, or simply implied, taught in isolation, or solely in the context of literature, how many, and which, rules are appropriate.

It was not until 1828 that Samuel Worcester produced a primer that borrowed the European idea of teaching children to recognise whole words without sounding them out.

It is not very important, perhaps, that a child should know the letters before it (sic) begins to read. It may learn first to read words by seeing them, hearing them pronounced, and having their meanings illustrated; and afterward it may learn to analyse them or name the letters of which they are composed (Crowder & Wagner, 1992, p. 204).

Support for this view came from James Cattell in 1885 with his assertion that whole word reading was more economical (Davis, 1988); and later, from the Gestaltists who considered that the overall shape of the word (rather than the summation of the sound-parts) should provide the pre-eminent clue for young readers. An assumption behind this approach was that beginning readers should be taught to read in the way skilled readers were thought to do. Given the belief that skilled readers associated meaning directly onto the whole-word image, it follows that there would be time saved by showing beginners how this was achieved. The alternative view was that reading should be viewed as a developmental process in which the early stages of developing the alphabetic principle are necessary for later skilled-reading, even though those early skills may be rarely needed at the later stages.

A further assumption of what became known as the whole-word approach was that the knowledge of letter-sounds would naturally follow once whole-word recognition was established (Smith, 1978). It was not until some time later that doubt began to be expressed about the effects on some children of this whole-word initial emphasis. Unfortunately for such children, the consequence of the primacy of the whole-word method was an inability to decode unfamiliar words (Tunmer & Hoover, 1993). The major reason for the length of time that elapsed before empirical judgements could be made about the relative merits of the

contrasting teaching emphases relates to the dearth of investigators engaged in such research until comparatively recently.

The History of Educational Research into Teaching Reading

It was not until the first two decades of this century that educational research began in earnest. The development of formal reading tests, and the recognition that education was a fertile ground for research led to many investigations into such topics as remedial approaches, individual differences in development, test development, silent reading vs. oral reading, and reading-readiness. Although this research was in its infancy, early findings (often unsubstantiated by independent research) were quickly adopted by book publishers keen to exploit the new markets that mass education provided. A number of texts based on whole-word teaching were published, and the method became very influential throughout the 1930's and 1940's. It appeared to offer a curriculum sensitive to the developmental needs of students, and one that would be both more attractive to teachers than phonics drill, and a more interesting approach for the rapidly increasing numbers of students engaged in the important journey towards literacy in the modern democratic society.

The whole-word model, as Chall (1967) described it, involved introducing words through their meaning. Words should be recognised by sight, using the cue of their shape and length. A fall-back strategy relied on deducing meaning from other clues such as pictures, or from the context. Phonic strategies were considered potentially harmful, and used as a last resort - but, even then, usually only to provide partial cues, such as obtained by attention to a word's first or last letters. Systematic teaching of phonic strategies was antithetical to the wholistic nature of such meaning-oriented approaches. Because teaching should not take as the unit of instruction anything other than meaningful text, any phonic skills developed by students would be self-induced and idiosyncratic.

The approach was taken even further when the whole-sentence, and then the whole-story became the units of study. In the sentence method, the child looked at the sentence being read by the teacher - this was followed by a focus on particular words in that sentence. In the whole-story method, the story was read to the child by the teacher before sentences and words were addressed. This approach was designed to make the meaning of print, rather than the mechanics pre-eminent, and was thought to be more interesting for the child, thus enhancing learning. Unfortunately, as the unit of analysis enlarged the more necessary it became for students to rely on memory. Some books began to use controlled vocabulary in the early reading stages, but the problem of decoding unfamiliar words was merely postponed, and the anticipated self-directed recognition of word similarities (providing a generative strategy) was too frequently unforthcoming. The end result was that, for many at-

risk children, progress came to an abrupt halt around Year Three or Four when an overwhelming number of unfamiliar (at least in their written form) words are rapidly introduced. Carnine (1982) estimated that the number of words a child needs to recognise in Year 2 was between three and four hundred, and in Years 3 and 4 between three and four thousand. Share (1995) estimates that the average 5th Year student encounters about ten thousand new words.

Strategies that relied upon memory-for-shapes of words, picture-clues, or context-clues become unproductive (Spear-Swerling & Sternberg, 1994). Depending largely on their visual recognition store of word shapes, students too often did not develop any generative strategy for the decoding of novel words. It is true that many children do develop a working understanding of the alphabetic principle despite the absence of explicit instruction; however, those students who did not have the *Ahah!* experience tended to be left floundering without the structure necessary to progress.

Prior to the whole-word dominant period, it was oral-reading that was most commonly taught and tested; however, with the increased emphasis on reading for meaning, silent-reading began to increase in popularity. Unfortunately, the cost of abandoning oral reading was the loss of information available to the teacher about progress and problem areas. This change allowed reading errors to be practised to the point of being firmly established. In addition, oral reading assists readers to become more familiar with those words whose spellings do not match their pronunciations (Adams, 1990); it assists students to become aware that written language provides the same opportunities for communication as does its oral form; and, in beginning readers, it leads to higher word recognition and comprehension scores (Carnine & Silbert, 1979). For older, more skilled students the primary mode of reading is appropriately silent.

The seemingly obvious solution involves a suitable balance so that both oral and silent reading opportunities are regularly scheduled at the appropriate reading stages. However, to some theorists, oral reading does not provide an authentic experience, because meaning may be compromised. "The basic mode of reading is silent. Silent reading does not place constraints on the reader" (Barmby, Bonham, Lawry, & Nissner., 1986, p. 35). Even today, some schools schedule daily silent reading under a variety of acronyms, and consider unnecessary the provision of opportunities for corrective feedback through regular oral reading. The presumption that practice makes perfect - that increasingly skilful reading will occur as long as the child engages in reading regularly - is misplaced with at-risk readers in

particular. In the absence of feedback, practice is likely to make errors permanent (Fields & Kempe, 1992), and this is especially true for at-risk students (Kameenui & Simmons, 1990).

The debate over code-emphasis vs meaning-emphasis has always been vitriolic. During the 1840's, the Secretary of the Massachusetts Board of Education in reporting on code-based (phonics) teaching colourfully described:

... the odor and fungousness of spelling-book paper" from which "a soporific effluvium seems to emanate ... steeping (the child's) faculties in lethargy". By contrast, meaning-oriented lessons ... will be like an excursion to the fields of Elysium compared with the old method of plunging children, day by day, for months together in the cold waters of oblivion, and compelling them to say falsely, that they love the chill and torpor of immersion (Adams, 1990, p. 22).

In the 1950's, the first real challenge to the whole word approach was provoked by Rudolph Flesch, who, in an emotionally charged attack, wrote:

It seems to me a plain fact that the word method consists essentially of treating children as if they were dogs. It is not a method of teaching at all; it is clearly a method of animal training. It's the most inhuman, mean, stupid way of foisting something on a child's mind (Flesch, 1955, p. 126).

Flesch's arguments were fuelled by aligning them with a perceived threat to democracy posed by an alleged decline in reading standards in the working class. He hinted at conspiracies to disempower sections of the community by deliberately using methods of teaching that were ineffective. "The American dream is, essentially, equal opportunity through free education for all. This dream is beginning to vanish in a country where public schools are falling down on the job" (Flesch, 1955, p. 132). Flesch's call for a return to phonics teaching had an enormous impact - the book was a best seller, and perhaps for the first time, parents began to express a desire to be involved in educational decision-making.

Community interest has continued from that time up to the present, and some similarities can be seen in the current phonics vs Whole Language debate but this was the first real taste of public accountability in education, and it had a significant impact on researchers, publishers and politicians. Courses on reading became more important in teacher training, research intensified, and government enquiries into literacy became regular events. Publishers began producing a wider variety of reading programs, from code-emphasis to meaning-emphasis, and various combinations of features from opposing schools for those wishing to take an eclectic stance on the issue.

Flesch's arguments had moved beyond the pedagogical to the political. The linking of phonics with democracy was a most effective strategy, though not one that endeared him to educational historians. His actions, however, were based on his assessment of which approach was more effective in teaching children to read. The current writings of a number of the leaders of the Whole Language movement (Edelsky, 1990; Goodman, 1989) display a similar interest in politics, though politics are the stated primary concern of these writers rather than as a means towards ensure good teaching practice (as Flesch had employed it). Questions of teacher or instructional effectiveness are less important to such advocates than are the objectives of personal liberation for students, and for society. "Whole Language ... has human emancipation as its goal" (Shannon, 1994, p. 99). Given the conceptual disparity between these major objectives it is unsurprising that genuine dialogue between Whole Language advocates and those seeking instructional sophistication is yet to be achieved.

In Flesch's time, there was still little systematic evidence about the relative effectiveness of the two major emphases across the broad population, and much debate centred on philosophical issues. Thouless might have had just such an issue in mind when he formulated his Law of Certainty. It can be summarised by the observation that when there is cause for doubt about a particular belief, or conflicts between approaches that are not readily resolved, one may reasonably expect that most people would adopt a position of caution. In reality, such uncertainty seems to polarise people's views strongly so that more are prompted to hold extreme views of support or condemnation than to hold a moderate position. Thus, supporters may clutch even more strongly to a belief about which there is doubt, while detractors focus strongly on the apparent negative aspects of the belief, disregarding any positives. This profound observation may partly explain why educational policy making continues to be subject to such extreme pendulum swings. Such a swing appears to be developing at present, as the dominant model, Whole Language, which is a development arising out of the meaning-centred, or whole word approach, comes under attack for its apparent ineffectiveness when applied to at-risk students. Barbara Bateman (1991) argued passionately that the whole-word emphasis, evident in Whole Language classrooms, lacks the explicitness of instruction in the alphabetic principle that is the key to mastering reading for at-risk children.

The abysmal overall record of this meaning-emphasis (whole-word) instruction is now so well known it need not be elaborated on here. It is sufficient to observe that of the millions

of children failed, and being failed by meaning-emphasis programs, a very large portion are slow learners. (p. 7)

The Great Debate

During the 1960's, Jean Chall (1967) was an important figure because she accepted the herculean task of developing a scientific study to test the effectiveness of various approaches to reading. The outcome of her work "Learning to read: The great debate" was published in 1967, and her conclusions were, and remain, controversial. Having analysed twenty basal level reading programs across 300 classrooms in three countries, and having studied the literature (such as it was) on effectiveness comparisons of phonics and whole-word approaches, she concluded that systematic teaching of phonics tended to produce better word recognition, spelling, vocabulary and comprehension in all children, not only those from the at-risk groups (such as students of lesser intelligence, or those from lower socio-economic backgrounds). Chall's detractors (e.g., Carbo, 1988) have disputed her conclusions, arguing that much of the research she reviewed had a variety of methodological flaws involving non-standardised tests, non-random selection, and inadequate program descriptions. Despite the criticisms, Chall's contribution was influential in affecting the weight of opinion regarding phonics (at least among researchers and some empirically-minded educators), and in stimulating a great deal of subsequent research.

Some of this later research continued to be criticised as flawed but, in general, greater rigour began to be a feature of the design of educational studies. The results of Chall's research were, however, less successful in altering the products of the publishers of beginning reading texts, and in influencing education bodies to promote practices of proven effectiveness in their domain. The failure of research-based knowledge to have an impact upon educational decision-makers continues to be lamented to this day (Carnine, 1995; Hempenstall, 1996; Stone, 1996)

The USOE Study

In the USA, the strength of public interest ensured that concern and research funding from governments was forthcoming. Large scale projects followed throughout the late 1960's and 1970's. The US Office of Education Co-operative Research Program in First Grade Reading Instruction was designed to overcome the criticisms of Chall's work, and to extend the research questions. Which approaches to beginning reading work best? Does reading-readiness affect program effectiveness? What characteristics of communities, schools, teachers and students are correlated with better outcomes? (Bond & Dykstra, 1967). Twenty

seven separate projects involving hundreds of classrooms were established - each informing on an element of the research questions in carefully designed studies coordinated by Bond and Dykstra. The findings were similar to those of Chall regarding the importance of phonics teaching, but also noted that a balance between phonics and meaning-emphasis was most productive.

Effective programs were found to be effective even for students with differing degrees of readiness. This latter finding was important because a whole industry of reading-readiness training was springing up. It was based on the assumption that children should not be taught reading until they had mastered a variety of visuo-spatial, language and motor skills. There were problems in ascertaining the core fundamental skills, accurately assessing them, teaching them effectively, and demonstrating an impact on reading progress (Arter & Jenkins, 1979). What was not apparent at that time was that learning to read was the most effective way to master many of those skills - hence valuable instructional time was better spent on the target task. "If the goal is for children to learn a particular skill (like reading), it is more efficient to teach it directly than to expect it to transfer from other learning" (Singer & Balow, 1981, p. 107).

In an analogous sense, the reading-readiness debate that gave primacy to the students' developmental stage in the ascription of when and what to teach, is being mirrored today in the interest among some developmentalists in so-called "learning styles" (Carbo, 1992; Dunn, Beaudry, & Klavas, 1989). In this view, there are important differences among children in their processing skills that require the tailoring of instruction to take account of those differences. If we accept the proposition that learning styles are genuine and important variables in learning, and further that they can be validly and reliably measured, then matching the instruction to the individual preference should produce superior learning. The approach has considerable intuitive appeal, and is the subject of an increasing amount of research. As regards reading however, there is little evidence that such matching enhances the process of learning to read (Snider, 1992; Stahl & Kuhn, 1995).

It was significant that, in the Bond and Dykstra (1967) study, the meaning-oriented approach (out of which evolved "language-experience" and "Whole Language") did as well as basal (without phonic-emphasis) programs with high-readiness students, but less well with low-readiness students. The adverse finding has been echoed over the past twenty years as modern "Whole Language" approaches are frequently criticised because of their apparent

ineffectiveness with at-risk students (Bateman, 1991; Gersten & Dimino, 1993; Stahl, 1990; Stahl & Kuhn, 1995; Vellutino, 1991; Yates, 1988).

Bond and Dykstra's findings were concise regarding the characteristics of communities, classrooms, teachers and students that were predictive of successful reading instruction. The major student predictor was not intelligence but knowledge of letters (predating the now acknowledged predictive power of phonological skills). The other finding, which perhaps played a part in the rise of the "effective-teaching" movement (Rosenshine & Stevens, 1984), was the importance of the method of delivery (in addition to the issue of content). The conclusion that teacher variables have a significant influence on student success was very important at a time when teacher differences were considered by many to be of little significance. "The implication is that to improve reading achievement we must improve both programs and classroom delivery. Each seems to contribute separately and significantly to children's progress" (Adams, 1990, p. 43).

In the following large scale study an model known as Direct Instruction successfully combined an explicit phonics emphasis with a teaching style emphasising explicit, systematic instruction of the type described in the "effective teaching" research.

Follow Through

This major study was federally funded in the USA in the late 1960's, arising because of a concern about the poor educational outcomes for disadvantaged students. Entitled Follow Through, it was aimed at the primary school stage, and was designed to determine which methods of teaching would be most effective for disadvantaged students throughout their primary school career. It followed an early-intervention project called Head Start that had as its goal the overcoming of educational disadvantage prior to school entry (i.e., at the pre-school level). The results of Head Start interventions unfortunately were not durable, and failed to achieve its ambitious objectives.

The impact of the unfulfilled promise of Head Start was felt by Follow Through. Though initially intended as a massive intervention, it was reduced in scope to that of a study to assess how best to maintain and build on Head Start's fragile gains. It remained, however, a huge study - involving 75,000 children in 180 communities over the first three years of their school life. It continues to be the largest educational experiment ever undertaken, extending from 1967 to 1995, at a cost of almost a billion dollars. There were nine major competing sponsors covering a broad range of educational philosophies. They included child-directed learning, individualised instruction, language experience, learning styles, self-esteem

development, cognitive emphasis, parent based teaching, direct instruction, and behavioural teaching. The models can be reduced to three distinct themes - those emphasising either basic academic outcomes, cognitive development, or affective development. The targeted basic skills included reading, language, spelling, writing, and maths. The models that emphasised the systematic teaching of basic skills (Direct Instruction, and Behaviour Analysis) performed best. In reading, the Direct Instruction model, which also has a strong phonic emphasis, had the most impressive results in both academic and affective areas.

There were criticisms that variability in implementation across sites made judgements of model superiority dubious, and that overall effects were too small to be pleased about (House, Glass, McLean, & Walker, 1978). Nevertheless, when the data was re-analysed by several groups (House et al., 1978; Bereiter & Kurland, 1981; Meyer, Gersten, & Gutkin, 1983), the Direct Instruction (DI) model still produced the best gains. Later follow-up studies (Becker & Gersten, 1982; Gersten, Keating, & Becker, 1988) were completed over the following 10 years, and added support to the argument that the superiority of the Direct Instruction model was real and significant.

To expect gains to endure over such a long period might be considered unrealistic, but Chall (1979) had argued that if children could master the decoding stage " ... the knowledge and skills acquired are usually sufficient to become self-generative. That is, further growth can be achieved with practice on one's own" (p. 47). This concept was extended by Share (1995) when he described phonological recoding as a mechanism enabling self-teaching of the decoding of novel letter combinations. Stanovich (1986) emphasised the role of practice by citing it as the major determinant of vocabulary growth after about Year 4, and even important in subsequent intellectual development. Thus, the positive findings in the follow-up studies imply that early skill mastery led to a continued interest and involvement in reading for those disadvantaged students who graduated from the Direct Instruction model. The DI model has been criticised (Schweinhart, Weikart, & Lerner, 1986) for its strong emphasis on teacher-directed, scripted lessons, alleging a consequential over-reliance on teachers, and an inability to self-direct learning. However, follow up studies of the DI students showed "strong consistent long term benefits in reading" three, six, and nine years after students completed Follow Through (Gersten et al., 1988, p. 326). The effects were evident in higher achievement, fewer grade retentions, and more college acceptances than in comparison groups that had traditional education in the same communities.

Becoming a Nation of Readers

In 1985, *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985), a report of the Commission on Reading examined the teaching of reading, reading problems, and likely solutions. It favoured approaches that included a strong, early, explicit phonics emphasis. The approach is enhanced when first individual sounds are taught along with procedures for teaching continuous ("mmmaaannn") rather than discontinuous ("mmm-aaa-annn") blends. Weisberg and Savard (1993) pointed out that, of eight major beginning-reading programs, only the Direct Instruction program - *Reading Mastery* (Engelmann & Bruner, 1988) - makes explicit to teachers the importance of promoting continuous blends. Their study highlighted the benefits to students of this strategy, and lamented the popular programs that either ignored its importance, or recommended a discontinuous blending strategy.

The argument about the constituents of effective phonics teaching is currently being revisited, as there is now a developing acceptance of the importance of phonic strategies in beginning reading. While some Whole Language theorists still believe that any emphasis on phonics is unfruitful, or even harmful - "The rules of phonics are too complex, ... and too unreliable ... to be useful." (Smith, 1992, p. 438), the major disagreement now revolves around the mode of teaching - not *if* phonics, but *how* phonics. Some acknowledge a role for phonics, albeit a secondary one. "Almost by definition, we can say that good readers are ones who use context efficiently, to reduce their reliance on visual cues and grapho-phonemic knowledge." (Weaver, 1988). Of those Whole Language advocates who see a role for phonics in a reading program, most argue that any word analysis skill development should occur only in the context of reading connected text (Weaver, 1988). See Iverson and Tunmer (1993) for a fuller discussion of this issue.

Thus, the sort of systematic explicit phonics teaching envisaged by the report of the Commission on Reading is unlikely to be found in a modern Whole Language classroom. It is not that such teaching could not be included, but that currently it is proscribed by the major writers in that field (Edelsky, 1990; Goodman, 1986, 1989; Weaver, 1988). Henry (1993) argues that Whole Language's lack of explicitness regarding phonics militates against at-risk learners as they are the least likely to develop their own phonic generalisations. A further problem for such students is that such unsystematic access to useful phonic principles leaves them without a firm basis for mastery, or with enough massed and spaced practice for incorporation to occur. A fuller discussion of the important elements of phonics approaches may be found in Foorman, 1995; Groff, 1990; Henry, 1993; Stahl, 1992.

The *Becoming a Nation of Readers* report was clear about the need for explicit instruction:

While questions during the preparation and discussion phases of a reading lesson are important, these do not substitute for active, direct instruction. In direct instruction, the teacher explains models, demonstrates, and illustrates reading skills and strategies that students might be using. There is evidence that direct instruction produces gains in reading achievement beyond those that are obtained with less direct means such as questions (Anderson et al., 1985, p. 56).

Becoming a Nation of Readers defined as state of the art a direct teaching model with a phonics emphasis. It was critical of much of existing practice in beginning reading, in methods of teaching comprehension, and in a lack of systematic formal and informal assessment. In common with a number of recent commentators and researchers, this report did not consider that the early inclusion of phonics instruction precluded a parallel emphasis on meaning, and the use of authentic literature.

The Impact of Research on Practice

In the years following, researchers have begun to look more closely at specific elements of curriculum content, and methods of instruction to allow a more fine-grained analysis of what works best for whom, and at what stage. Thus far the lack of impact of this research on educational practice has been of concern to many in the educational community (Hempenstall, 1996; Stone, 1996). In earlier times, research findings were rarely conclusive, and it is understandable that such research results were not a major force in educational policy formulation. There is now a consensus among empiricist researchers about a number of issues crucial to reading instruction, and these are discussed in Chapter 4. However, the currently dominant model of reading instruction, known as whole language does not support an explicit phonics emphasis in beginning reading, and it is this model that is discussed in the next chapter. Its importance lies in the influence it has on the extent of reading success or failure among students in Australia.

CHAPTER THREE: THE WHOLE LANGUAGE APPROACH TO READING

When Australian state education departments, and major teacher associations take the unusual step of endorsing a particular model of teaching, such as Whole Language, one would anticipate that the decision would have been made with due gravity, including careful consideration of evidence supporting the model as worthy of such acclamation. Not only should such a model be well-credentialled, theoretically and empirically, but it should be sufficiently flexible to accommodate the diverse range of learners dependent on classroom experiences for the majority of their learning opportunities. This chapter examines the philosophy and practice of Whole Language, highlighting the flaws that make it an inappropriate model for such endorsement, and argues that its impact on 'at-risk' students is deleterious rather than supportive.

Historically, the consideration of learning disability has emphasised within-person factors to explain the unexpected difficulty that academic skill development poses for students with such disability. Unfortunately, the impact of the quality of initial and subsequent instruction in ameliorating or exacerbating the outcomes of such disability has received rather less exposure until recently. Over the past decade an approach to education with strong philosophical underpinnings, Whole Language, has become the major model for educational practice in Australia (House of Representatives Standing Committee on Employment, Education, and Training, 1992). There has been increasing controversy, both in the research community (Eldredge, 1991; Fields & Kempe, 1992; Gersten & Dimino, 1993; Liberman & Liberman, 1990; Mather, 1992; McCaslin, 1989; Stahl & Miller, 1989; Vellutino, 1991; Weir, 1990), and in the popular press (Hempenstall, 1994, 1995b; Prior, 1993) about the impact of the approach on the attainments of students educated within this framework. In particular, concern has been expressed (Bateman, 1991; Blachman, 1991; Liberman et al., 1989; Yates, 1988) about the possibly detrimental effects on "at-risk" students (including those with learning disabilities).

Whole Language: History

The Whole Language approach has its roots in the meaning-emphasis, whole-word model of teaching reading. Its more recent relation was an approach called "language experience" which became popular in the mid-1960's. The language experience approach emphasised the knowledge that children bring to the reading situation - a position

diametrically opposed to the Lockian view of "tabula rasa" (the child's mind as a blank slate on which education writes its message). In this language experience approach there is a firm link between oral language and written language, between reading and writing. "Anything I can say, I can write; anything I can write, I can read" (Weaver, 1988, p. 44).

The teacher uses the prior experiences and school excursions that a child has had to enable the child to dictate a story that the teacher records. The teacher and child read and re-read this story until the child can do so alone. Any skill teaching must derive from the child's story, hence the expression -teaching only from a meaningful context. There is the possibility within this framework that teachers will provide structured learning experiences around fortuitous opportunity but no clear recommendation that they should.

Whether the Whole Language approach represents an evolution from language experience (Stahl & Miller, 1989) or is sufficiently different to be considered an entirely separate model (McGee & Lomax, 1990), it is clear that they have commonalities and differences. Both emphasise the relevance of the language and knowledge that children bring to reading and that helps to link oral and written language. Both object to subskills teaching in isolation from the context of meaningful literature. In Whole Language, however, teachers are less likely to write children's dictated stories and more likely to encourage the children to write their own stories using invented spelling (Schickedanz, 1990). Language experience stresses the inter-relatedness of reading, writing, speaking and listening but, unlike Whole Language, delays the introduction of writing until the child has mastered a reasonable number of sight words (Allen, 1976; Stauffer, 1969, cited in Stahl & Miller, 1989). Weaver (1988) makes it clear that the developmental process for writing follows a scribbling - invented spelling - mature writing sequence, and hence writing should be a natural part of the language process from the beginning stages of reading development.

Goodman (1986) describes Whole Language as a philosophy rather than as a series of prescribed activities. Thus, Whole Language teaching consists of those activities a teacher with a thorough understanding of the philosophy would use. The teacher aims to provide a proper environment that will encourage children to develop their skills at their own developmentally appropriate pace.

This makes it difficult to describe what actually occurs in a Whole Language classroom, or whether there is any consistency from classroom to classroom that would enable an observer (other than one imbued with the philosophy) to recognise that the approach was indeed Whole Language. This vagueness is still evident in a selection of recent

journal articles (Smith, 1991; Newman, 1991; Johnson & Stone, 1991). There is a strong emphasis on principles, for example, the benefits of a natural learning environment (Goodman, 1986), and of exposure to a literate environment (Sykes, 1991). Mills and Clyde (1990, cited in Johnson & Stone, 1991) provide an outline of the Whole Language philosophy as evidenced in classrooms.

Highlight authentic speech and literacy events; provide choices for learners; communicate a sense of trust in the learners; empower all participants as teachers and learners; encourage risk taking; promote collaboration in developing the curriculum; be multimodal in nature; capitalise on the social nature of learning; encourage reflection. (p. 103)

Assumptions of the Whole Language Model:

Naturally Unfolding Development.

The abovementioned prescriptions do give the flavour if not the substance of what may occur in classrooms, and are consistent with a view of child development that combines a Rousseauian perspective of naturally unfolding development with an assumption that learning to read is essentially equivalent to learning to speak. Rousseau believed that children had an innate developmental script that would lead them (though perhaps at differing rates) to competence. Thus, unfettered maturation would allow the child to develop knowledge unaided (Weir, 1990). His ideas gained scientific respectability in the 19th Century when they were seemingly supported by a theory of evolutionary biology. This long since discredited theory asserted that the evolutionary journey from amoeba to human infant was replayed in every pregnancy, and the wisdom and knowledge of the parents (and of necessity, beyond) was present in the brain of the new generation. In Rousseau's view, humans were naturally good but could be turned bad by societal interference. His argument that society should not interfere in the natural development of children was paralleled by his view of the role of education. "Give your pupil no lesson in words, he must learn from his experience" (Rousseau, 1964 cited in Weir, 1990, p. 28). The Whole Language philosophy noted above, which assigns to the teacher the role of concerned facilitator and which decries teacher directed instruction as harmful or unproductive, can be readily sourced to the Rousseauian view.

Weir (1990) is critical of the foundations and practice of Whole Language which she argues has led to an increase in illiteracy, and the shifting of blame for poor achievement from the school to the home. She believes that advocates of this approach have a responsibility to provide evidence for naturally unfolding development to justify the use of indirect process-oriented education. Weir considers that Frank Smith and the Goodmans have dominated educational policies without an acceptable research base for their theories. Delpit (1988) is especially concerned about the effects of progressive education on minority groups. Rather than it being supportive of personal growth she sees the approach as being disempowering. "Adherents of process approaches ... create situations in which students ultimately find themselves held accountable for knowing a set of rules about which no one has ever directly informed them" (p. 287).

Reading as a Natural Process.

The model also assumes that reading (and writing) are natural parts of the same language process that enables the development of speech. Learning to read and write would be just as effortless and universal if the tasks were made as meaningful as is learning to talk. While the vast majority of children learn to speak with reasonable facility, a sizeable proportion of children do not learn to read well. In the USA, the figure is usually put at between 20 and 25 per cent of the school population (Stedman & Kaestle, 1987). In a recent study (Prior et al., 1994), 16 per cent of Year Two children in a representative Victorian community sample were considered reading disabled. One can recognise the principle of naturally unfolding development in Goodman's (1986) explanation for the disparity in ease of acquisition between speaking and reading. According to Goodman, it is the breaking down of what is naturally a wholistic process into subskills to be learned and synthesised that causes the gulf between expertise in speaking and reading.

Lieberman and Liberman (1990) do not accept that the fault lies with the unnecessary or harmful intervention of society through the education system. They argue that reading and speaking are qualitatively different activities, and cannot be expected to be mastered in the same epigenetic manner. They highlight a number of differences: all humans have developed language systems but only a minority a written form; while speech has a history as old as the species and appears to be biologically driven, written codes, or more accurately, alphabets have a cultural basis and a relatively short history (about 4000 years); speech all around the world is produced in a similar fashion using a limited range of sounds, while scripts are artificial systems that differ enormously across different cultures; while speech develops merely through exposure to speech, reading usually requires formal assistance. Liberman and Liberman conclude that learning to speak and learning to read are qualitatively different. Treating the two forms of language development as similar involves a false assumption, and they argue, the practices that derive from that assumption are part of the cause of reading failure. Stanovich (1986) agrees and cites a number of prominent researchers who accept the characterisation by Gough and Hillinger (1980) of reading as an "unnatural act." p. 396.

The Induction of the Alphabetic Principle.

Recognising the phonological basis of our language system is vital for it allows us to generate an infinite number of words from a limited range of sounds. Without it we would be reduced, as are animals, to a range of meanings equal to the number of distinct sounds we can produce. It is phonology (along with syntax) that distinguishes human language systems from other forms of natural communication. Children must have a wonderful capacity for

managing the phonology of language - by the age of 6 years the average vocabulary is 13,000 words (Miller, 1977, cited in Liberman & Liberman, 1990). The key to translating this ability to reading lies in the child's understanding of the alphabetic principle, the basis of English spelling. Because script is composed of graphemes that are roughly similar to the phonemes of spoken words, children must learn how spoken language maps onto written language (Griffith & Olson, 1992). In grasping the alphabetic principle the child must have some degree of phonemic awareness (the conscious realisation that words can be decomposed into discrete single sounds (phonemes), and letter/sound knowledge (Byrne & Fielding-Barnsley, 1991). This phonemic awareness helps children make sense of instruction about what sounds each letter makes in a word. The child is able to separate out those individual sounds (segmenting) when they are presented in the context of the word's other sounds. Without phonemic awareness the child is forced to memorise complete word patterns but is unable to manage novel words. As the memory demands escalate, memorising the letter landscape will become a less and less reliable strategy, and the child will become unduly reliant upon less effective strategies such as context cues.

Research continues to explore the significance of a range of phonological processes, but there is already an enormous weight of evidence that deficits in the area of phonemic awareness are responsible for the discrepancy between the ease of learning to talk and learning to read (Ball & Blachman, 1991; Tangel & Blachman, 1992). What makes the alphabetic principle difficult for some children is that while written words consist of a sequence of discrete graphemes, the spoken word consists of co-articulated sounds blended into a continuous rapidly-produced stream. Some children have great difficulty with the analysis of these co-articulated phonemes. The folding together of vowels and consonants alters their individual sounds, permitting speaking rates of 10-20 phonemes per second (Liberman & Liberman, 1990) effortlessly, automatically, seamlessly, and unconsciously. Someone must have first noticed that words like "cat" and "bat" shared some similarity, and that they could be represented more economically by sharing that similarity in the written form also. This was a significant linguistic discovery because it allows each phonological element to be recognised by a special shape, and anyone who knew the shape and consciously understood the internal structure of words could read. This is the discovery every beginning reader must make - unless somebody tells him or her. Whole Language approaches assume that children will discover the alphabetic principle through exposure to print, and through their writing experiences. In homes where early literacy experiences include an interest in the

structure of language, it is likely that children are not unduly disadvantaged by this failure to make explicit the importance of our language's structure. Unfortunately, when phonemic awareness is emphasised neither at home nor at school, children are unnecessarily placed at risk of failing at the task of reading.

While invented spelling, as used in Whole Language writing activities, can be a useful step on the way to phonemic awareness and literacy, a rationale that precludes corrective feedback (and assumes closer and closer approximations to accurate spelling will occur naturally) may lead to over-optimism about the utility of the strategy. Bryant and Bradley (1985) point out that children initially read and spell words in quite different ways, and hence invented spelling activities may contribute little to reading progress. Similarly, Thompson, Fletcher-Finn and Cottrell (1991, cited in Tunmer & Hoover, 1993) found that any knowledge of phoneme-to-letter correspondences acquired through invented spelling activities did not automatically transfer as knowledge of letter-to-phoneme correspondences in reading.

Many researchers (Stahl & Miller, 1989; Stanovich, 1986; Prior et al., 1994; Blachman, 1991; Grossen & Carnine, 1990; Byrne & Fielding-Barnsley, 1989; Groff, 1990) consider the notion of learning by "discovery" cavalier, and prejudicial to the progress of at-risk students - those least likely to induce the alphabetic principle, and who make up the majority of the children who do not learn to read adequately. Perhaps because of the distaste for quantitative research displayed by many Whole Language advocates (Groff, 1990) few empirical studies have been published to support the Whole Language assumption that the alphabetic principle will be induced. One study (Klesius, Griffiths Zielonka, 1991) compared a traditional basal approach and a Whole Language approach at Year 1 level. The basal approach did not have a synthetic phonics basis or teach phonemic awareness. The results indicated that although the Whole Language group achievement was lower than the traditional instruction group on all measures, none of the differences was significant. Unfortunately, those who began the year with low phonemic awareness skills remained so, and showed slower reading progress. This finding is in line with arguments that not only Whole Language programs but meaning-emphasis and analytic phonics-based programs that do not make explicit the alphabetic principle are ineffective for at-risk students (Chall, 1987; Bateman, 1991; Grossen & Carnine, 1990; Vellutino, 1991). "What they need to know, and what their experience with language has not taught them, is no more and no less than the alphabetic principle" (Lieberman & Lieberman, 1990, p. 72). More recently (e.g., Foorman, Francis, Beeler, Winikates, & Fletcher, 1997) there have been studies indicating the

superiority of phonics-emphasis beginning reading programs over the Whole Language approach.

Can Whole Language and Phonics be Reconciled?

The problem of unsystematic and indirect teaching of phonic skills proving ineffective for some students was addressed by Eldredge (1991). He compared a number of first grade programs using a Whole Language approach with a similar cohort using the same programs supported by 15 minutes of synthetic phonics. The modified program group scored significantly higher on all literacy measures after one year. To the extent that a well-designed phonics program can enable the development of the alphabetic principle, the addition of instruction in phonics should enhance the outcomes in Whole Language classes, and there is increasing evidence that it does so. In order for Whole Language advocates to adopt such strategies an adjustment to the philosophies behind their practices would be required. Thus far, however, Whole Language philosophy has been relatively impervious to the results of research. In fact, McCaslin (1989) warns that a major problem for the future development of Whole Language is its assumption that an empirical research perspective is responsible for inappropriate practice.

Ball (1993) also notes the conflict between the Whole Language philosophy's lack of attention to the structure of language and the consistent research on the causal link between metalinguistic awareness and reading development. In her view, the pedagogical battle between code-emphasis and Whole Language supporters is reflective of a broader debate evident in many of the social sciences. The major debate is between those who support a reductionist, positivist philosophy of science and those who rebel against that position - adopting a holistic, post-positivist, relativistic stance. In Groff's (1990) view, the reading dispute narrows down to the question of what constitutes the reality of reading behaviour. To relativists such as Weaver (1988), all empirical research is futile in determining teaching practice, because in performing the research we cannot avoid affecting the outcome - thereby confounding results. Relativists view reality as phenomenological, that is, it has no existence independent of our unique individual perspective. They tend to favour ethnographic approaches, such as case studies and classroom observation, as the appropriate means of enquiry, because those strategies do not interfere with naturally occurring processes. Empiricists view reality as "essentially cognitive transcending" (Rescher, 1982 cited in Groff, 1990), and see ethnographic research as useful for raising, rather than answering, questions about teaching practice.

In a comprehensive examination of the philosophical underpinnings of the education system in the USA, Stone (1996) decries the influence of developmentalism which he considers pervades classrooms and teacher training institutions to the detriment of students. Stone describes the history of developmentalism as reaching back to Rousseau, and includes Dewey, Piaget, Hall, Gesell, James, and Vygotsky as major contributors to the primacy of naturally occurring development, and to the suspicion accorded to all interventive approaches that harm is the inevitable outcome of interference with the natural order.

If decisions are to be made about state-supported approaches to reading then the question of who will evaluate claims of the two sides becomes critical. Keith Stanovich (1994), one of the foremost researchers and commentators on reading, argues that the weakness of educational decision-making is its vulnerability to faddish swings, a view also supported by Stone (1996). In Stanovich's view, it is the failure of policy makers to base decisions on empirical research, and their uncritical acceptance of the glib assurances of gurus, that has led to the current dissatisfaction in the wider educational community. He proposes that competing claims to knowledge should be evaluated according to three criteria. Firstly, findings should be published in refereed journals. If research is to be useful it must be well designed, and able to justify its findings. When peer review is part of the process of research, the well-known taunt "research can prove anything you want" becomes less valid. Poorly designed studies are rejected (often to appear in unrefereed journals). Secondly, reported results should be replicated by independent researchers. One feels more comfortable when research findings are repeated in studies where the researchers have no particular stake in the outcome. Thirdly, there is a consensus within the appropriate research community about the reliability and validity of the findings. This last criterion requires considerable reading across the field, but the frequency with which a particular study is cited, and accepted as legitimate, in journal articles provides one measure.

Whilst the use of these criteria cannot guarantee infallibility, it does offer reasonable consumer protection against spurious claims to knowledge. For example, were such tests used over the past 15 years to determine best practice, the claim would never have been accepted that learning to read is as natural and effortless as learning to speak; or that good readers use contextual cues to guide their reading, using print only to confirm their predictions. Yet these unsubstantiated (and demonstrably false) claims were accepted and a generation of teachers pressured through initial teacher-training and subsequent Ministry sponsored in-service to implement practices derived from them. Such erroneous practices have been especially

damaging to vulnerable students - those who aren't self-sustaining, who can't afford ineffective strategies, who rely on teachers rather than their parents to educate them.

It is clear that the sheer weight of evidence running counter to basic Whole Language postulates is having an impact at a policy level. In the USA, the Report of the Commission on Reading, *Becoming a Nation of Readers* (Anderson et al., 1985) supported the empirical approach "The trend of the data favours explicit phonics" (p. 42). In 1986, the US Congress contracted Marilyn Jager Adams to write a book about the critical elements in teaching beginning reading. Her book, *Beginning to read: Thinking and learning about print* (1990), is a milestone in that it synthesises from a variety of fields research that impinges on reading development. These research areas include education, psychology, linguistics, neurology and physiology. Her book is potentially very influential, recommends early and sustained intervention in teaching the structure of our language to beginning readers, has been roundly condemned by Whole Language supporters (Goodman, 1991), but has been difficult to ignore. It at least represents a scholarly focus for debate, and perhaps, dialogue.

In recent times, California has become the second least successful state in the union in the reading achievement of its students. As a result of the outcry that followed this finding, California has recently developed new guidelines for acceptable approaches to teaching reading, and has proscribed its formerly strongly-embraced Whole Language approach. The state insists that empirically supported approaches that include attention to the structure of language (that is, models emphasising phonemic awareness and phonics) be adopted in all schools.

On October in the USA, occurred a most significant event in the long history of the debate on the teaching of reading. The Reading Excellence Act (1997) was passed in the federal House of Representatives. The importance of the Act resides its mandating that any federal funding for programs in future must be based on the program being able to demonstrate reliable and replicable research support. This means that only objective, valid, scientific studies can be used to validate the approaches proposed in any project if funding is sought.

For many professions such an expectation would be unlikely to raise eyebrows; however, education has been a profession steeped in mythology, alchemy and magical thinking. This Act represents a revolution in education and its effects are likely to be felt soon in Australia, given our State and Federal governments concern with accountability in

education, and the statements of Minister for Education and Training about the allocating of funding on the basis of results.

Groff (1990) first suggested a commission of disinterested scholars who would determine firstly whether empirical research is admissible as a valid means of enquiry, and further, would judge quality. This is now the approach adopted in the US through the Reading Guarantee Act (1997) and its likely liaison with the National Institute of Child Health and Human Development. Unfortunately, this has not de-polarised the debate, with many Whole Language supporters incensed that decisions about reading are being taken out of the hands of teachers. Despite this outcry, it seems likely that a consequence of the this Act will be a reduction in the pendulum swings that have plagued education for such a long time. The effect this direction should ensure that the novelty of an prospective approach (without empirical support) will not be considered sufficient reason for its adoption. This is likely to have a damping action on such fashion swings, and simultaneously to direct developers' attention to the need to investigate the impact of their product before attempting its promulgation.

The Impact of Whole Language in Australia

In Australia, in 1993, a National House of Representatives Committee released a report "The Literacy Challenge", noting Whole Language has Australia-wide support and " virtually all curriculum guidelines on primary school literacy teaching produced are based on this approach. ... Virtually all teachers have undertaken the inservice training course, Early Literacy Inservice Course (ELIC), which is also based on a Whole Language approach to learning and literacy" (p. 25). While the Committee heard much evidence in support of the teaching of phonics, its recommendations did not include such an emphasis, finishing rather lamely, "The Committee accepts the arguments that there is no single correct method which will suit all children" (p. 27). Their recommendations were similarly vague. "All literacy training include specific instruction in the range of teaching strategies" p. 30. Interestingly, in an appended dissenting report five of the twelve members asserted that "All literacy training include specific instruction in decoding, skill acquisition and spelling" p. 64. It would seem that the pervasive influence of developmentalism described by Stone (1996) is as applicable to Australia as to the USA.

Given the degree of penetration of the Early Literacy Inservice Course it is instructive to examine it in more detail, and in particular in its views on the method and content of reading instruction.

In 1988, the Victorian Ministry of Education released the English Language Framework P - 10 "Language for Living". This document advocated a Whole Language approach to English teaching, and, although its recommendations were not compulsory, it was widely adopted in that State. In order to assist teachers to put the model into practice, literacy consultants from the Ministry's School Support Centres were enlisted to provide in-service teacher training. Of the courses offered the Early Literacy Inservice Course (ELIC) (Education Department of South Australia, 1984) was the most widely promoted. A ten unit program developed in South Australia, it was designed to be undertaken by groups of teachers after school for 1/2 hour each week with an additional 1 hour per week for between-unit activities and professional reading. The ten topics were: young children learning language, observing children reading, interpreting and using running records, matching children with books, encouraging reading development, the writing process, encouraging writing development, teaching writing, making programming decisions. The unit texts provide illustrations of appropriate activities, and Unit 5: Encouraging Reading Development is of interest for its title, and for the absence of any reference to teaching. The experiences considered worthwhile are: shared book experience, listening to stories, dictating and writing own stories, frequent silent reading, responding to stories. Further encouragement for the child-centred, discovery nature of the approach appears in the same Unit booklet. "Children's reading development, like their oral language development, largely depends on their establishment of a self regulating and self improving system" (Badger, 1984, p. 19).

Whilst this description of the function of the teacher highlights one major difference between the Whole Language and code emphasis/direct teaching approaches, another is the role of phonic skills in learning to read.

Whole Language Philosophy in Practice

Semantic, Syntactic and Graphophonic Cues.

Proponents of Whole Language either: disparage phonics, "Phonics is incompatible with a Whole Language perspective on reading and therefore is rejected" (Watson, 1989, p. 132); submerge phonics, "phonic information ... is most powerfully learned through the process of writing" (Badger, 1984, p. 19); or argue that phonic skills are taught within the context of three systems used to extract meaning from print (Cambourne, 1979). In this latter view, the graphophonic system is considered a fall-back position to be used when semantic and syntactic (the other two systems) fail (Weaver, 1988). Graphophonic cues refer to the correspondence between graphemes (the symbols in print) and phonemes (the speech sounds

they represent). Semantic cues involve incorporating the meaning of what is being read to assist with decoding words about to be read, that is, the next word should make sense in the context of the sentence's meaning. Syntactic cues arise because of the logic of our system of sentence construction: the next word is constrained by the rules of grammar. Syntactic and semantic cues are broadly described as context cues, as they may be used to predict a word without recourse to visual inspection. Goodman (1979) described skilled reading as a *psycholinguistic guessing game*. He considers reading a sophisticated guessing game driven largely by the reader's linguistic knowledge, and as little as possible by the print. Smith (1975) expresses this view succinctly. "The art of becoming a fluent reader lies in learning to rely less and less on information from the eyes" (p. 50). It was argued (Cambourne, 1979) that the speed of skilled reading could not be accounted for if the reader looks at every word. The hypothesis was that the good reader used contextual cues to predict words initially, and then confirm the word's identity using as few visual features as possible.

Holdaway (1980, cited in Hornsby, Sukarna, & Parry, 1986) provides this strategy. When word recognition is the problem readers should "(a) go back and read from the beginning of the sentence and/or read further on; (b) check the first letter or letter cluster; (c) make a prediction (an informed guess)" (p. 104).

The results of eye movement studies have not supported the skipping hypothesis. These studies (see reviews in Rayner & Pollatsek, 1989; Stanovich, 1986) using modern eye movement technology indicate that skilled readers do process all the print - they do not skip words, or seek only some features of words. Thus, the techniques of contextual prediction that are emphasised in Whole Language classrooms are based on an untenable hypothesis. It is unsurprising that Rayner (1989), perhaps the most notable of the researchers on eye movement studies, considers that the major failing of Whole Language is its lack of recognition that graphophonic cues are "more central or important to the process of learning to read than are the others" (p. 351). Bruck (1988) reviews research indicating that rapid, context-free automatic decoding characterises skilled reading. In fact, the word recognition of skilled readers provides them with the meaning even before contextual information can be accessed. Rayner and Pollatsek (1987), cited in Liberman and Liberman (1990), argue that it is only beginning and poor readers who use partial visual cues, and predict (or guess) words. This view is echoed by Stanovich (1986) who refers to a significant number of studies in support, and a further list of such studies can be found in Solman and Stanovich (1992).

The second rationale for presuming that contextual cues should have primacy in skilled reading was based on a flawed study by Goodman (1965, cited in Nicholson, 1986). Goodman found a 60-80% improvement in reading accuracy when children read words in the context of a story rather than in a list format. He argued on the basis of this study that the contextual cues provided marked assistance in word identification. There has always been acceptance that context aids readers' comprehension, but despite contention in the literature over Goodman's finding concerning contextual facilitation of word recognition, his study is still regularly cited as grounds for emphasising contextual strategies in a Whole Language classroom. The study was flawed in two ways. The design was not counterbalanced to preclude practice effects. That is, a list of words taken from a story was read, and then the story itself was read. Secondly, the study ignored individual differences in reading ability, so it was not possible to determine whether good, or poor readers (or both) derived benefit from context. Studies by a number of researchers including Nicholson (1985, 1991a), Nicholson, Lillas and Rzoska (1988), Nicholson, Bailey and McArthur (1991) have discredited Goodman's argument, and found that good readers are less reliant on context clues than poor readers. Poor readers attempt to use context because they lack the decoding skills of the good readers. Nicholson (1991a) argues that encouraging reliance on contextual cues confuses children, and he expresses concern at the rate of reading failure in New Zealand where Whole Language is endemic. A further problem involves the accuracy of contextual guesses. In a study by Gough, Alford, and Holley-Wilcox, (1981, cited in Liberman & Liberman, 1990) well educated, skilled readers given adequate time could only guess correctly one word in four from context. Schatz and Baldwin (1986) pointed out that low frequency words, and information-loaded words, are relatively unpredictable in prose. Finally, psychometric studies indicate that it is not measures of semantic and syntactic ability that predict word identification facility but rather alphabetic coding ability (Vellutino, 1993). Whole Language theorists would anticipate the converse being true.

Prior et al. (1994) in their study of more than 1600 Victorian children agreed that guessing is not an adaptive strategy, and that its promulgation disadvantages at-risk children. They argue that reading-handicapped children, in particular, need intensive training in phonetic analysis. This argument is also supported by numerous influential researchers (Chall, 1989; Bateman, 1991; Groff, 1990; Solman & Stanovich, 1992; Tunmer & Hoover, 1993; Adams, 1990; Byrne & Fielding-Barnsley, 1989; Ball, 1993; Blachman, 1991;

Eldredge, Quinn, & Butterfield, 1990; Nicholson, 1991a; Yates, 1988). Whole Language supporters do not accept this view.

If one accepts the empiricist position that learning to read is not a natural process equivalent to learning to talk, then the view that most language activities are equally helpful to reading development becomes dubious, as does the related assertion that children will master reading by being exposed to a literate environment. The literature on direct instruction (Rosenshine & Stevens, 1984) provides convincing evidence that students learn to read best when the allocated time for reading is spent directly on reading activities rather than on activities once or twice removed from reading. This literature also highlights the necessity of systematic teaching, careful monitoring and continuous feedback. Thus, it is not only the philosophy of the Whole Language approach, but the practices that derive from it, which do not have adequate research support.

Practices Recommended in Whole Language Programs.

In a similar vein if one accepts that the value of contextual strategies has been vastly over-rated and the value of phonic skills similarly under-rated, then one must query the value of the classroom activities that follow from contextual primacy. Hornsby, Sukarna, and Parry (1986) suggest:

(a) Teachers emphasise shared-book experience.

Nicholson (1985) criticises this activity because it bypasses a reader's decoding problem instead of directly addressing it. The presumption is that with the crutch provided by the shared-book experience students will be able to solve their own decoding problem. He compares this approach to attempting to teach a rat about mazes by wheeling it through the corridors in a trolley.

(b) Teachers use Cloze activities. They are designed to encourage children to use just enough visual information, for example, the first two letters of a word to assist word prediction, and the intention is to increase reading rate without cost to comprehension. However, skilled readers perceive and use all the letters in a word to decode (it is faster and more accurate than prediction and confirmation), thus this activity is unproductive, even counter productive.

Given the Whole Language emphasis on deriving cues about meaning from as many sources as possible, it is unsurprising that picture books may form a part of the reading program for beginning readers. Of course, picture books have been evident in classrooms long before Whole Language became prevalent but have been incorporated as a useful

element in a Whole Language program (ELIC, Unit 4, 1984). Studies by Solman and colleagues (Solman, 1986; Singh & Solman, 1990; Solman, Singh, & Kehoe, 1992) have cast considerable doubt on the wisdom of this strategy if the goal is to improve decoding. In fact, the presence of pictures, regardless of their salience to the words, impedes rather than assists word identification.

This finding highlights a problem with models that are philosophically rather than pedagogically driven. Just because a practice is consistent with a philosophical position does not mean that it will be effective in the classroom. It may even, as in this case, be counterproductive. Unfortunately, the view of empirical research expressed by Weaver (1988) "... it is impossible to conduct empirical research without affecting the outcome" (p. 220) is common among Whole Language advocates, and what a teacher does can become a moment-by-moment decision based on some intuitive understanding of the needs of the immediate situation.

The ELIC program (Unit 3, Interpreting and Using Running Records) highlights the importance of self-correction rates, and exhorts teachers to spend considerable time and energy in assessing the self-correction rates of all their students regularly. Clay (1969, cited in Share, 1990) noted that good readers self-corrected errors at a higher rate (once to every three or four errors) than did poor readers (once to every eight to twenty errors). She considered high rates were indicative of good text cue integration, which in turn was a measure of reading progress. The value of this activity has been questioned by Share (1990), and Thompson (1981, cited in Share, 1990). They found that self-correction rates are confounded with text difficulty. When text difficulty was controlled in reading level-matched designs, the rates of self-correction became similar. That is, when text is very difficult one is more likely to make errors, and increase the rate of self-correction. This is true for good readers and poor readers. Hence, an increased rate of self-correction could be interpreted as indicative of too difficult text. The conclusion that there is no direct support for self-correction as a determinant of reading progress makes the activity of recording such ratings for students of questionable value.

Assessment Techniques Used in Whole Language Classrooms.

Miscue analysis is a major procedure for assessing what strategies children are using in their reading. Goodman and Burke (1970, cited in Allington, 1984) were interested in a qualitative analysis of readers' errors. They were concerned only with errors that caused a loss of meaning; the number of errors was less important than the immediate impact on

comprehension. Hence decoding errors such as reading "ship" for "boat" were indicative of the student using contextual cues appropriately, and a signal for satisfaction about reading progress. The Reading Miscue Inventory (RMI) they developed did not focus on the graphemic and phonemic aspects of oral reading, but children who made errors based on graphemic similarity, for example, "boot" for "boat", would be considered to be over-relying on phonic cues, and in need of encouragement to rely more on context. Given the current knowledge about reading, the interpretation of the results of the RMI is not helpful to future planning for young readers. It is now considered (Stanovich, 1986) that a reader has a certain amount of attentional capacity to devote to the reading task. Good readers because of their relatively error-free, automatic, context-free decoding skills are able to devote most of their attention to comprehension. Conversely, most of the attentional capacity of struggling readers is used in battling the code, and focussing on less helpful strategies like context cues. The consequence of this expensive use of attention is that such students have relatively little capacity left for comprehension. The implication of these findings is that the qualitative analysis of reading errors is largely superfluous to planning. Decoding errors of whatever type are best addressed at the level of decoding instruction. Thus, the student who makes errors based on contextual strategies, and the student who makes errors based on inadequate graphophonic skills both require decoding instruction, and practice sufficient to enable effortless reading at the appropriate level of text difficulty.

The final problem for the Reading Miscue Inventory is its inadequacy as a psychometric instrument (Allington, 1984). Describing Leu's (1982) review of oral reading error analysis, Allington presents a number of deficiencies:

- (a) Vague definitions of the boundaries of the error categories;
- (b) An absence of theoretical justification for the categories;
- (c) A failure to allow for the effects of passage difficulty. When passage difficulty is controlled (i.e. similar error rates), reliance on context occurs at least as much for less skilled as for skilled readers (Allington & Fleming, 1978; Batey & Sonnenschein, 1981; Biemiller, 1970, 1979; Cohen, 1974-5; Coomber, 1972; Harding, 1984; Juel, 1980; Lesgold & Resnick, 1982; Perfetti & Roth, 1981; Richardson, Di Benedetto, & Adler, 1982; Weber, 1970; Whaley & Kibby, 1981; cited in Stanovich, 1986);
- (d) The ambiguity resulting when categorising multiple-source errors.

The Reading Miscue Inventory has had considerable influence in instructional texts and in classrooms (Allington, 1984), and is still influential among Whole Language theorists

(Weaver, 1988). Weaver also describes a revised version - RMI: Alternative Procedures (Goodman, Watson, & Burke, 1987). The rationale for the revision appears unchanged - " ... it is best to avoid the common sense notion that what the reader was supposed to have read was printed in the text" (Goodman et al., 1987, cited in Weaver, 1988, p. 340). Given the problems with theory, design and implications of the Reading Miscue Inventory its widespread acceptance in the education community is difficult to fathom.

Providing Corrective Feedback.

Teacher response to error is an area of instructional methodology in which Whole Language is in conflict with much empirical evidence. Corrective feedback, as defined by Kameenui and Simmons (1990) is "the instructional procedure that directs ... attention to incorrect responses and provides correct information" (p. 234). It is an integral element of Direct Instruction programs (Gersten, Woodward, & Darch, 1986), effective teaching principles (Yates, 1988; Good & Brophy, 1987), and considered of particular importance to students involved in special education (Hendrickson & Frank, 1993; Fields & Kemp, 1992). Whole Language theorists stress the importance of students taking responsibility for their own learning and of being prepared to take risks. They also see correction as an unnecessary interruption to the comprehension process (Goodman, 1970, 1973; Kemp, 1987; Smith, 1971, cited in Fields & Kemp, 1992), and hence are less supportive of the process. This is sometimes carried to extremes when learners' errors are quite acceptable and "celebrated" (Goodman, 1986, p. 47, cited in Liberman & Liberman, 1990), and further, considered "charming indications of growth towards control of language processes" (p. 19). The underlying philosophy of naturally occurring development is evident here. A concern that teachers may be ignoring this important instructional strategy was confirmed in a study by Fields (1991, cited in Fields & Kemp, 1992). Of 110 primary teachers employing a Whole Language approach, error correction was the least used of 31 instructional practices described. In a follow up study (Fields & Kemp, 1992), 66 Queensland state primary teachers, who had received formal training on one or other Whole Language course (e.g. ELIC), and whose approach to teaching met at least nine of the following Whole Language characteristics, were invited to participate. The characteristics were chosen from descriptions by Reutzel and Hollingsworth (1988), and Slaughter (1988), cited in Fields and Kemp (1992).

1. Indirect instruction (the teacher acts as a collaborator and facilitator);
2. Child centredness (the child's level of development and readiness is considered very carefully);

3. Dialogue and teacher scaffolding (tasks involve frequent teacher-pupil discussion and, where necessary, teacher assistance and support, to solve problems that the child cannot solve);
4. An informal classroom environment;
5. Whole Language used in context;
6. Intact literacy events (not an emphasis on substeps or specific skills);
7. Learn by doing;
8. The child's own writing;
9. Authentic oral language (not controlled or modified in any way);
10. Meaning dominated interactive discourse;
11. Pupil-pupil collaboration.

The teachers were provided with descriptions of the oral miscues of 6 hypothetical students and asked what corrections, if any, they would provide. In the majority of cases, self-correction oriented cues were provided, for example, delaying a response, asking the child to re-read, and requesting a meaning check. The authors noted that although the content of the feedback would more usefully have been code-based rather than context-based; nevertheless, these teachers were prepared to offer corrective feedback despite their training. In their ELIC course they would have been informed that "no amount of explanation, correction, or instruction has any immediate impact on children's language because they direct what they will learn and when they will learn it" (Badger, 1984, p. 16). They raise the possibility that some teachers, at least, are aware of "what works" in their classrooms, and pragmatically incorporate aspects of different models into their reading program. Vellutino (1991), in a review of reading instruction, agrees that good teachers quickly become aware of the limitations of a Whole Language philosophy. If this is so, then it is possible that those teachers who claim to be Whole Language teachers are, in fact, offering an eclectic program without the deficiencies in the purist model. Unfortunately, little is known about the existence or prevalence of such classrooms, although some Whole Language theorists believe it would be problematic if such eclecticism occurred. Newman (1991) despaired that the theoretical and political beliefs supporting Whole Language have not been accepted by some teachers who may only be "teaching Whole Language in the afternoons" (p. 73). She argues that only by being thoroughly imbued with the spirit can the "moment-by-moment judgments" (p. 74) needed in teaching be made appropriately. Mather (1992), like Pearson (1989), believes that good teachers will use what is effective, but is concerned about inexperienced teachers, and

those who are less analytic about their practices. She sees many students in Whole Language classrooms as victims of "poor programs produced in the heat of intense ideological debate" (p. 93). Ultimately, it is not enough to hope that teachers can make the right decisions in the classroom despite inadequacies in their training. An approach that has been found to be flawed fundamentally must either be revised or replaced.

Whither Whole Language?

Vellutino (1991) and other contemporary researchers (Ball, 1993; Bateman, 1991; Blachman, 1991; Byrne, 1991; Byrne & Fielding-Barnsley, 1989; Eldredge, 1991; Gersten & Dimino, 1993; Groff, 1990; Liberman & Liberman, 1990; Nicholson, Bailey, & McArthur, 1991; Rayner & Pollatsek, 1989; Solman & Stanovich, 1992; Stahl & Miller, 1989, Tunmer & Hoover, 1993 Weir, 1990;) are in agreement that Whole Language is not a comprehensive approach to reading instruction. Given that it is not just one approach among many, but is a model endorsed and promulgated in Australia and elsewhere by government education bodies, the disparity between its wide acceptance and the vast contrary evidence is alarming. While some authors (Groff, 1991; Liberman & Liberman, 1990) find little to recommend it, others believe that with modification to its methods of teaching, and to the content included, it could be recast into a generally acceptable and comprehensive approach (Chaney, 1990; Gersten & Dimino, 1993; Heymsfield, 1989; MacGinitie, 1991; Prior et al., 1995; Spiegel, 1992). Some (e.g., Stahl & Miller, 1989) consider it a valuable introduction to reading, but of less value beyond an orientating function, while others (Ball, 1993) fear that the differences may be so fundamental to make rapprochement impossible without a change in the basic philosophy of Whole Language.

Given the large body of evidence in support of phonemic awareness and the alphabetic principle as major determinants of reading success, it is hard to imagine that Whole Language can remain immune and unyielding, and still maintain credibility as a model of reading acquisition endorsed by state governments. Perhaps the reasonableness of the position taken by Foorman (1995) and Heymsfield (1989), or the improved student outcomes obtained by adding code instruction to a Whole Language program as described by Castle, Riach, and Nicholson (1994), Eldredge (1991), Heymsfield (1992), and Uhry and Shepherd (1993) will enable the evolution of the Whole Language approach into a more comprehensive and effective model, better able to meet the educational needs of the diverse group of learners in our classrooms. Certainly if one examines empirically accepted findings such as summarised by Vellutino (1991), it is difficult to accept the status quo.

(a) The most basic skill in learning to read is word identification; (b) an adequate degree of fluency in word identification is a basic pre-requisite to successful reading comprehension; (c) word identification in skilled readers is a fast acting, automatic, and in effect modular process that depends little on contextual information for its execution; (d) even skilled readers can accurately predict no more than one word out of four in sentence-contexts, indicating that the predictive role of context must be extremely limited; (e) because of limited facility in word identification, beginning and poor readers are much more dependent on context than are more advanced and good readers; (f) facility in alphabetic coding is critically important to the acquisition of skill in word identification; (g) phoneme awareness and facility in phoneme analysis are critically important to the acquisition of skill in alphabetic coding. Each of these generalisations is contrary to the approach to reading instruction currently advocated by Whole Language proponents (Vellutino, 1991, p. 442).

Newly elected conservative governments in Australia have demonstrated an increasing, if controversial, interest in the establishment of state and national testing programs. In addition, such governments have shown a distinct preparedness to examine the effectiveness of programs that compete for the scarce education dollar. It would be ironic, if in a time of decimation (in the true sense of the word) of the education system, one positive outcome was a shift towards accountability as objectively assessed by student outcome. One of the oft-heard complaints from researchers in this field is that educational decision-making is too often driven by ideology, or uncritically accepted innovation. There may well be an opportunity now for those of an empirical bent to influence such result-driven policy makers towards educational practices with legitimate theoretical and research support. Even a cursory reading of the popular media over recent years indicates that there is a real and growing dissatisfaction with the state of literacy in Australia, and that this dissatisfaction is centred on the manner in which it is being taught in our schools. Who is prepared to take up the issue with the decision-makers to create the structural changes necessary to rescue our system? Researchers have traditionally shied away from such overt involvement in the process of exerting influence. Yet they are an important part of an assembly that should also include teachers, parents, teacher educators, speech pathologists, school consultants, such as educational psychologists, and any other interested parties. Evidence, numbers, conviction, energy and political (and media) influence are all elements needed to create change in a system. For the sake of those not well served by the current system, who are unable to influence their predictably bleak future, it is surely time to stop fiddling around the edges of

the problem. It is time to address the core issue: the manner in which we approach beginning reading instruction.

This issue is addressed in the next chapter.

CHAPTER FOUR: THE ROLE OF PHONEMIC AWARENESS IN READING

In recent years there has been an abundance of research highlighting the pre-eminent status of phonemic awareness in the development of reading capability. This chapter examines the concept of phonemic awareness, and reviews recent findings pointing to its causal links with reading. The importance of linking reading instruction with phonemic awareness is explored, as is the significance of early identification and intervention. Finally, obstacles to systemic implementation are considered, in conjunction with the bleak prognosis for students with a history of early reading failure.

Phonemic Awareness: What Does it Mean?

Over the past two decades, but particularly in the last 10 years, there has been a burgeoning consensus about the critical importance of phonemic awareness to beginning reading success, and about its role in specific reading disability or dyslexia (Hatcher et al., 1994; Share, 1995; Stanovich, 1986). Phonemic awareness has also been described as phonological awareness, acoustic awareness, phonetic awareness, auditory analysis, sound categorisation, phonemic segmentation, phonological sensitivity, and phonemic analysis.

There has been some discussion about how best to define phonemic awareness. Ball and Blachman (1991) refer to the ability to recognise that a spoken word consists of a sequence of individual sounds. Stanovich (1986) defines it as the "conscious access to the phonemic level of the speech stream and some ability to cognitively manipulate representations at this level" (p. 362). Later, he suggested (1992, 1993b) that the terms "conscious" and "awareness" themselves have no acceptable definitions, and recommended phonological sensitivity as a generic term to cover a continuum from shallow to deep sensitivity. This term acknowledges the wide range of tasks used to assess levels of sensitivity. Read (1991) too was concerned about the term awareness, but because it implies a dichotomy rather than a continuum. He preferred the term *access to phonological structure*. As these alternatives have not yet gained currency, phonemic awareness will continue to be used here, accepting that the definition has limitations.

What is clear is that phonemic awareness concerns the structure of words rather than their meaning. To understand the construction of our written code, readers need to be able to reflect on the spelling-to-sound correspondences. To understand that the written word is composed of graphemes that correspond to phonemes (the alphabetic principle), beginning

readers must first understand that words are composed of sounds (phonemic awareness) rather than their conceiving of each word as a single indivisible sound stream. This awareness appears not to be a discrete state, but rather a sequence of development ranging from simple to complex, or as Stanovich (1992, 1993b) would prefer - from shallow to deep.

Phonemic awareness is more complex than auditory discrimination, which is the ability to perceive that *cat* and *mat* are different speech productions, or words. To be able to describe *how* they are similar but different, however, implies some level of phonemic awareness. The first entails hearing a difference, the second entails a level of analysis of the constituent sounds. Young children are not normally called upon to consider words at a level beyond their meaning, although experience with rhymes may be the first indication for children that they can play with the structure of words. For young children, the realisation that spoken sentences (a rather continuous stream of sound without clear pauses) are separable into discrete words is a pre-requisite for the recognition that words can be decomposed into segments (Liberman & Liberman, 1990).

Adams (1990), and Blachman (1984) warn that word consciousness (the awareness that spoken language is composed of words) should not be assumed even in children with several years schooling, though they report evidence that it may be readily taught even at a pre-school level. That school age children can lack such fundamental knowledge may be difficult for adults to accept, but it highlights the need in education to assume little, and assess pre-requisite skills carefully. Their warning also challenges the view, held by some Whole Language advocates (Goodman, 1979, 1986; Smith, 1975, 1992), that speaking and reading involve equivalent "natural" processes for all children. The implications of the Whole Language view are that the same environmental conditions that occur during the development of speech are those best provided for children learning to read. Liberman and Liberman (1990) have provided a forceful rebuttal of this position.

Having discovered that sentences are composed of words, the next logical unit of analysis is at the syllable level. However, syllables can be represented by any number of letters from one to eight. The word *understand* has three syllables, each of a different number of letters. *Un* has two, *der* has three. and *stand* has five letters. This variability makes the syllable unit of limited value in analysing the reading task (Bradley, 1990).

Rhyme and Alliteration

The recognition of rhyme may be the entry point to phonemic awareness development for many children (Bryant, 1990). To be aware that words can have a similar end-sound

implies a critical step in metalinguistic understanding - that of ignoring the meaning of a word in order to attend to its internal structure. This leads to a new classification system, one in which words are classified according to end-sound rather than meaning. Bryant (1990) points to the considerable amount of evidence indicating that children as young as three or four years can make judgments such as - when words rhyme, and when they begin with the same sound (alliteration). He argues that sensitivity to rhyme makes both a direct and indirect contribution to reading. Directly, it helps students appreciate that words that share common sounds usually also share common letter sequences. The child's subsequent sensitivity to common letter sequences then makes a significant contribution to reading strategy development. Indirectly, the recognition of rhyme promotes the refining of word analysis from intra-word segments (such as rhyme) to analysis at the level of the phoneme (the critical requirement for reading).

Studies by Bryant, Bradley, McLean, and Crossland (1989) showed a very strong relationship between rhyming ability at age three years and performance at reading and spelling three years later. A number of such studies have reinforced the value of such early exposure to rhyming games (e.g., Kirtley, Bryant, Maclean, & Bradley, 1989). That rhyming and phoneme awareness are related (through their common characteristic of requiring listening for sound similarities and differences) was supported by an interesting finding of a study by Lamb and Gregory (1993). They showed that children who were capable of good discrimination of musical pitch also scored highly on tests of phonemic awareness. Since pitch change is an important source of information in the speech signal (Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), it may be that sensitivity to small frequency changes involved in phoneme recognition is an important aspect of successful reading. Lamb and Gregory (1993) raise the interesting possibility that musical training may represent one of those pre-reading, home-based experiences that contribute to the marked individual differences in phonemic awareness with which children commence school.

Onsets & Rimes

Treiman (1991) has described a further stage in the development of phoneme awareness: the intra-syllabic units - onset and rime. The onset of a syllable is its initial consonant(s), and the rime is its vowel and any subsequent consonants in the syllable. Thus, in the syllables *sip-slip*, the onsets are *s* and *sl*, and the common rime is *ip*. Treiman's research has indicated a stage between syllable awareness and phoneme awareness when children are much more sensitive to the onset-rime distinction than the phoneme distinction. It has been

argued that this research holds promise for programs of educational intervention in reading disability because of the greater regularity of onset-rimes over individual letters (Felton, 1993). Thus, rime phonograms such as *ing*, *ight*, *ain* have much more regularity than the letters that form them. Knowing that *strain*, and *drain* rhyme, allows for reading *main* and *brain* by analogy. This has led some researchers (Bowey, Cain, & Ryan, 1992; Hulme & Snowling, 1992) to suggest that an emphasis on onset-rime may be an especially valuable approach to teaching dyslexics who tend to have relatively weak phonological skills.

Bowey and Francis (1991) also consider onset and rime the most effective focus for phonological activities intended to promote beginning reading and spelling for all children. They note that since most onsets in English are single consonants, the use of the intra-syllabic onset/rime distinction as the major unit in the early study of word structure is likely to hasten the development of awareness at the more difficult phoneme level. Treiman (1991) has argued convincingly that the onset/rime division is a natural one. Bradley (1990) too agrees, and considers that it is because rhymes correspond to rimes that most children develop such facility with them at a relatively early age. The awareness of these larger sublexical skills are viewed by Bruck (1992), Goswami and Bryant (1990) Tunmer and Hoover (1993) as prerequisites to initial reading acquisition, their difficulty level lying between that of syllable awareness and phoneme awareness (Bowey et al., 1992; Bowey & Francis, 1991; Bruck & Treiman, 1990; Kirtley et al., 1989). Spector (1995) perceives onset/rime as a potentially useful stage in the development of oral segmentation skills. She recommends, for children who have difficulty in segmenting complex syllables, the strategy of breaking such words into onset/rime as an intermediate step towards phonemic segmentation.

There appears to be a developmental sequence of phonological awareness. It begins with awareness of words as a unit of analysis, then proceeds to the awareness that words can share certain ending properties that we call rhyme; to an awareness that words can be decomposed into syllables, then more finely into sub-syllabic units called onsets and rimes, and then (and most importantly for reading) into awareness of individual phonemes, the smallest unit of sound analysis. A further developmental sequence involves the movement from a recognition of such properties to a capacity to produce examples of them. Thus, at one level one can nominate which pairs of words rhyme when presented orally; at a higher level one can produce examples.

If this is the developmental sequence, then the approach to effective teaching should take account of this sequence. The empirical question that arises is whether an emphasis on

teaching such an onset-rime distinction (rather than at the phoneme level) is more productive in initial (and, perhaps, remedial) reading instruction. The computer program developed by Wise, Olson and Treiman (1990) has focussed on onset-rimes in teaching beginning reading skills to normally-developing and dyslexic children. In the Wise et al. (1990) and the Olson and Wise (1992) studies, the authors noted an advantage for the children taught in this manner over an approach that segmented words after the vowel. The effect however was ephemeral, and least pronounced in the more disabled students. Ehri and Robbins (1992) findings were similar in that the poorer readers did not use sub-syllabic units larger than the grapheme. This led them to suggest that the onset-rime distinction is really the province of the more skilled reader. Goswami's research (Goswami & Bryant, 1990) had suggested that, for young children, words that share rimes are more readily decoded by analogy than are words that share onsets or vowels. Bruck and Treiman (1992) provided some support for that view, but as in the Wise et al. (1990) study, the measured advantage was lost within a day. In fact, a day later the rime group demonstrated poorer performance than the group taught onsets, and poorer than the group for which vowel analogy was emphasised. Nation and Hulme (1997) question the value of an early emphasis on onset-rime as skill at such tasks is not predictive of reading and spelling success.

These findings do not imply that struggling readers cannot be taught to make use of the strategy, nor does it mean that reading words by analogy is an unproductive strategy. However, the results of research presented above suggest caution regarding calls for introducing an initial emphasis on onset-rime distinctions for beginning readers. It would be judicious to ensure that beginners (and disabled readers) have or develop a grounding in grapheme-phoneme relationships, either before (or simultaneous with), such onset-rime emphasis (Munro, 1995). It is still unclear whether the generally accepted developmental sequence necessarily provides the optimum guidance for instruction. The instruction question should be answered empirically, and a number of researchers are attempting more fine-grained analysis to assist in providing clearer instructional direction. Olson (in press, cited in Snowling, 1996) reported a study indicating that adequate phonemic awareness skill was necessary if children were to benefit from onset-rime instruction. When dyslexic readers were provided with phonemic awareness training through Auditory Discrimination in Depth (Lindamood & Lindamood, 1969), simultaneously with onset-rime computer-based training, reading results were markedly improved. The ADD program emphasises phonemic awareness through a variety of oral/aural tasks, and by teaching students awareness of kinaesthetic cues

(mouth, tongue, lip position, breath usage). Nation and Hulme (1997) argue that it is likely to be more profitable to emphasise phoneme awareness even from the beginning reading stages. As is often the case, when several options are available and the evidence is not adequate to clearly support one or the other, the emphasis is most judiciously placed on the alternative that is most closely related to the reading process.

Thus, studies to now have raised more questions than answers about the instructional usefulness of onset-rime as a means of gently approaching the difficult phoneme concept.

Phoneme Awareness

Awareness at the level of the phoneme has particular significance for the acquisition of reading because of its role in the development of the alphabetic principle - that the written word is simply a means of codifying the sound properties of the spoken word. In order to decode the written word, one needs to appreciate the logic of the writing system, and as a prerequisite, the logic of oral word production.

There are two requirements of beginning reading for which phonemic awareness becomes immediately relevant: phonemic analysis and phonemic synthesis. For most children, the ability to produce the finer discrimination of phonemes begins in about Year I of their schooling (Ball, 1993). Individual phonemes are more difficult to specify because their acoustic values vary with the phonemes that precede and follow them in a word (a phenomenon called co-articulation), whereas syllables have relatively constant values in a word and hence are more readily recognised. The fact that consonants are "folded" into vowels can be understood by noting the different tongue positions for the beginning /d/ sound when it is followed by /oo/ and by /i/.

In most children the ability to synthesise (blend) sounds into words occurs earlier than analytic (segmentation) skills (Bryen & Gerber, 1987; Caravolas & Bruck, 1993; Solomons, 1992; Torgesen et al., 1992; Yopp, 1992). Thus, it is easier to respond with the word "cat" when presented with the sounds c - at or c-a-t, than it is to supply c-a-t when asked to tell what sounds you hear in "cat".

Tasks used to assess beginning (or shallow) phonemic awareness tend to emphasise sensitivity to rhyme and alliteration; for example, finding a word that begins or ends with the same sound as the stimulus word. A more complex task would involve the manipulation, or separation of sounds in a word, for example, What is the first sound you hear in "cat"? What word is left if you remove the /t/ from "stand"? (Torgesen et al., 1994). The shallow level of awareness typically develops during the pre-school years, the degree dependent on language

experiences, and perhaps, a genetic component (Olson, Wise, Connors, Rack & Fulker, 1989; Rack, Hulme, & Snowling, 1993). Other tasks used for assessment may include counting the sounds in words, adding, deleting or manipulating sounds, and categorising sounds at the beginning, middle, or end of words. Most of the tests available thus far are informal and without norms, but see Torgesen and Bryant (1994a) for a normed test for young children. Whereas the research findings are very impressive, there is inevitably a delay before comprehensive, valid, and reliable tests are constructed and promulgated. There are, as yet, no recognised tests that are able to delineate clearly the developmental stages, the skill levels of sensitivity and manipulation, and the at-risk from the normally progressing student.

As indicated above, deeper levels of awareness (i.e., at the phoneme level) tend to develop during Year (or Grade) 1 upon exposure to reading instruction. This raises the possibility that phonemic awareness may be a consequence of learning to read rather than a causal factor (Morais et al., 1987; Morais, 1991). The issue is by no means resolved; however, there is increasing consensus that the data are best explained by considering the relationship between phonemic awareness and reading development as a reciprocal one (Stanovich, 1992).

Phonemic Awareness: Its Relationship to Reading Development

Adams (1991), having published an authoritative text on beginning reading (Adams, 1990), was further moved to write “To my mind, the discovery and documentation of the importance of phonemic awareness ... is the single most powerful advance in the science and pedagogy of reading this century” (p. 392). Her enthusiasm for this area of research has been increasingly shared by researchers across a wide range of disciplines - education, special education, cognitive, educational and developmental psychology, and linguistics, judging by the number of published articles on phonemic awareness over the last 10 years.

Correlational Studies

The interest in this area is unsurprising when one considers that phonological abilities (of which phonemic awareness is a subset) are the most powerful predictors of reading success. A number of researchers have noted that the predictive power of measured phonological abilities exceeds that of more general cognitive abilities such as intelligence, vocabulary, and listening comprehension (Adams, 1990; Bradley & Bryant, 1983; Juel, 1988; Wagner & Torgesen, 1987; Yopp, 1988). This finding has been demonstrated not only for the English language but also for Swedish, Spanish, French, Italian, and Russian (Adams, 1990). The many correlational studies (see Wagner & Torgesen, 1987 for a review) that support this

link cannot, however, provide evidence of causality. It is known, for example, that knowledge of letter names prior to reading instruction is a strong predictor of success. Yet for children who do not know their letter names, teaching such names does not improve their reading prognosis (Wagner, Torgesen, & Rashotte, 1994). The early letter knowledge is merely a marker for other individual differences such as IQ, attention span, or early literacy experience; however, Walsh, Price and Gillingham (1988) provide a more optimistic view of the value of teaching letter names to a stage of automaticity.

A major problem for correlational studies, as argued by Felton (1992), is their facility for predicting good reading outcomes, but inability to shed light on just which children will not make progress. Thus, she reiterates Mann's (1984, cited in Felton, 1992) finding with a heterogeneous population in which a combination of phonological tasks: naming speed, phonetic recoding in working memory, and phonological awareness, assessed in the first year of school, accounted for 74% of the reading variance a year later. In contrast, for an at-risk sample, Felton and Brown (1990) found the same series of tests accounted for 43% of the reading variance a year later. The extent of the variance explained is impressive in either case, but also indicates that much variance is still unexplained.

Training Studies

In addition to the correlational evidence indicating that phonemic awareness is strongly predictive of reading attainment, there have accumulated a number of longitudinal training studies showing that the relationship between phonemic awareness and reading progress is indeed causal. This second finding is of great significance, for without it one could argue that phonemic awareness is purely a consequence of reading development, or alternatively merely related to a third variable (the true cause) such as intelligence, or social class.

The most famous of these studies (Bradley & Bryant, 1983) was described by Coltheart (1983) as being the first to provide an understanding of the cognitive processes involved in beginning reading acquisition. They were able to overcome the design problems that had limited the salience of the findings of many studies. Their landmark study became a model design (Bowey & Francis, 1991), and a stimulus to the now burgeoning research in this area. Bradley and Bryant developed a combined longitudinal and training study, because a longitudinal study alone cannot guarantee causality but does demonstrate genuine relationships, whereas training studies alone can demonstrate cause-and-effect relationships but may be "... arbitrary; one cannot be sure that such relationships exist in real life" (Bradley

& Bryant, 1983, p. 419). Thus, in such a combined approach, the limitations of either type tend to be cancelled out, without compromising the advantages of each.

Bradley and Bryant (1983) studied the predictive ability of sensitivity to rhyme and alliteration. They were interested in whether high levels of sensitivity were associated with later reading success and low levels with reading difficulty over the next four years. They screened 118 three year olds, and 285 five year old children for reading ability, and none was able to read any words on a reading test. This is an important safeguard because of the potential influence of reading facility on phonemic awareness (Bowey & Francis, 1991). They were able to demonstrate strong correlations between the original sound categorisation scores and students' reading and spelling over three years later. They selected 65 of the students with low (below 2 SD from the mean) phonemic awareness scores, and randomly assigned them to a training, and a non-training group. The first group was taught (in 40 sessions over two years) to attend to the sound structure of words, while the second was taught to categorise words in terms of their meaning. The children received normal reading instruction in school and at the end of the project were re-assessed. The training group had made significantly more progress in reading - an effect specific to reading as the two groups were similar in a standardised maths test.

Bradley (1990) retested the original experimental and control groups five years after the training was completed. The differences, quite remarkably, were still present in all four reading and spelling tests. The children who received less than seven hours individual assistance at age six or seven had maintained the advantage over a five year period. The value of early intervention in sound categorisation is obvious. Those children in the experimental group who were also taught letter-sound correspondences, and how sound and letter patterns are connected, performed far better than all other groups.

Linking Phonemic Awareness to Reading

Subsequent intervention studies may be divided into those that have, and those that have not emphasised the connection between phonemic awareness activities and letter-sound knowledge. Lundberg, Frost, and Petersen (1988) in an intensive pre-school study taught phonemic awareness activities (though not letter-sound knowledge) over a whole school year. At the end of the first and second year of school the experimental group demonstrated phonemic awareness, reading, and spelling superior to those of the controls.

Thus, the activities that involved no print had a beneficial effect when reading instruction occurred in the following year. Cunningham (1990) obtained similar results in her

study of first and second year students. Her study, however, was difficult to interpret as the students may have been receiving letter-sound training in their school reading program, an activity that might alternatively explain the improvement. Studies that have separated phonemic awareness in time from reading instruction are important in demonstrating that phonemic awareness could not be purely a consequence of reading instruction; however, a number of studies have shown the value of adding letter-sound training to phonemic awareness activities. Ball and Blachman (1988, 1991) taught phonemic awareness and letter sounds to an experimental group, general language activities and letter sounds to a control group, and also had a no-intervention control group. The groups of four to five students met four times per week for 15-20 minutes over a seven week period. The experimental group outperformed the control groups in phonemic awareness, reading and spelling. The two groups receiving letter-sound instruction did not differ in letter-sound knowledge, thus letter-sound knowledge alone did not have an impact on phonemic awareness, reading or spelling at the time of the posttest.

Byrne and Fielding-Barnsley (1989, 1990, 1991b, 1993, 1995) also conducted studies in which the teaching of both phonemic awareness and letter-sound correspondence to pre-readers was necessary to establish the alphabetic principle. Using a program they had developed, *Sound Foundations* (Byrne and Fielding-Barnsley (1991a), they examined whether young pre-school students could be taught the alphabetic principle by emphasising how different words may begin or end with the same sound (phoneme identity). It was their belief (as earlier Bradley and Bryant, 1983 had argued) that phoneme identity was the ideal vehicle for promoting alphabetic insight, as its attainment implies the presence of segmentation skills. Their approach was more parsimonious than that of a number of more broadly-based programs that had included a significant emphasis on segmentation training. Their concern was not whether omnibus phonemic awareness programs were effective in developing phonemic awareness, and subsequently, reading skills. The question for them was whether a well aimed, but minimal (in time and resources), intervention could achieve a similar outcome. The minimalist attitude extended to teaching only a representative range of sounds (seven consonants and two vowels) over a twelve week period. Compared to a control group their students gained in phonemic awareness (even to untrained sounds) and knowledge of the alphabetic principle.

Follow-up research (Byrne and Fielding-Barnsley, 1993, 1995) indicated that this advantage extended to superiority in pseudo-word decoding in Years One and Two, and in

reading comprehension in Year Two. They concluded that children who enter school with some understanding of the structure of words (prior to significant print experience) find it easier to master the task of converting written to spoken language.

Similar findings, this time with respect to invented spelling performance, which is a good measure of phonemic awareness in beginning readers (Mann, Tobin, & Wilson, 1987) were reported by Tangel and Blackman (1992). Hatcher, Hulme and Ellis (1994) refer to this approach of integrating phonemic awareness and letter-sound instruction as the "phonological linkage hypothesis" (p. 42). Hatcher et al. compared the progress of six and seven year old students with reading difficulties, exposed over a 20 week instructional period, to one of three conditions. The matched groups were taught reading, or phonological skills, or both. The children in the reading-with-phonology group demonstrated most improvement in reading and spelling at the conclusion, and at a nine month follow-up. The effect appeared to imply a synergism, as the phonology-alone group had more phonology improvement than the integrated group, but no significant reading improvement. Further, albeit oblique, support for the phonological linkage hypothesis comes from a study by Iversen and Tunmer (1993) in which children who were in a Reading Recovery program showed quite dramatically accelerated progress when phonemic awareness activities were added to the established regimen.

Studies, such as that by O'Connor, Jenkins, Leicester, and Slocum (1993), that have focussed on at-risk, reading disabled, or learning disabled students have noted either a slow response to phonemic programs or failure to generalise phonemic skills (within and across tasks) - or both of these. In fact, the Wagner, Torgesen and colleagues (1993, 1994) studies noted at-risk students' high levels of resistance to progress. The authors warn that programs will need to be more intensive, and perhaps more extended, than those currently trialed. To that, Blachman (1994) adds a concern for both treatment components and treatment timing. In a recent study, Blachman, Ball, Black, and Tangel (1994) showed the usual group mean improvement of children taught phonemic awareness and letter sounds in their first year of school. They further examined the intra-experimental group differences, and provided additional phonemic awareness and letter-sound instruction for the first 12 weeks of the second year of school to those students who had made minimal progress in the experimental group. Additionally, their reading program had a strong phonics emphasis to build upon the phonological development. The results were very pleasing, and highlight the need for continuous progress assessment, with intervention applied as student need dictates. Berninger

and Abbott (1994) consider such resistance to usually effective teaching programs as their preferred system for classifying students as learning disabled, eschewing the commonly used aptitude-achievement discrepancies.

There has been a marked increase in the number of programs published recently for training phonemic awareness in young children. Some programs currently available include: Byrne and Fielding-Barnsley (1991a); Catts and Vartiainen (1993); Lindamood and Lindamood (1969); Solomons (1992); Torgesen and Bryant (1994b); and Trelease (1989). The nature of tasks varies but may include:

(i) beginning speech-sound awareness activities, such as listening to nursery rhymes or alliteration sequences;

(ii) making judgments or producing rhymes or alliteration, sounds games (e.g., "I spy..."), and answering questions about word structure (e.g., Do these words rhyme?, Which word is longer?, and Which of these words starts the same as *cat*?);

(iii) blending/segmentation activities (e.g., counting or tapping syllables, pronouncing syllables, or isolating syllables, onsets, rimes, and phonemes, deletion tasks (e.g., What word is left when *s* is removed from *sit*?) and addition tasks (e.g., Say *it* with *s* at the beginning).

Despite the genuine cause for optimism that the phonemic awareness research evokes, there remain numerous questions regarding program content, the age of intervention, the method of content delivery, and the identification of those most in need. Increasingly studies are emphasising a more fine-grained analysis of the structure of phonemic awareness (Høien, Lundberg, Stanovich, & Bjaalid, 1995; Wagner, Torgesen, Laughon, Simmons, & Rashotte, 1993; Wagner, Torgesen, & Rashotte, 1994) in an attempt to resolve these pressing questions.

Other Phonological Abilities.

Phonemic awareness is only one, albeit critical, member of a class of phonological processing skills that involve the use of the sound structure of oral language in learning to read. (Adams, 1990; Badian, 1993; Cornwall, 1992; Crowder & Wagner, 1992; Felton & Brown, 1990; Torgesen, 1993; Torgesen, Wagner, & Rashotte, 1994; Wagner & Torgesen, 1987, Wagner et al., 1993; Wagner et al., 1994).

Another phonological skill, besides phonemic awareness, which has been implicated in reading progress is speed of lexical retrieval, also known as phonological recoding in lexical access. It is assessed through the ability to name rapidly colours, letters, numbers and objects. It is considered relevant to reading because it is indicative of how readily children can gain access to a sound, sound-sequence, or a word meaning (Bowers & Swanson, 1991;

Cornwall, 1992; Davis & Spring, 1990). Both naming speed and sight word reading depend on automatic, rapid symbol retrieval, and Wolf (1991) argues an important connection between naming speed for both letters and numbers, and word recognition.

Readers must apply a conversion from the print into one of:

(i) a phonological representation constructed through oral reading or subvocalization.

This process allows appropriate selection of the word's meaning via the access to the phonologically coded lexicon, the link having been developed through oral language.

Or (ii) employing a visual representation of the printed word to gain direct access to the lexicon. This system represents the most common strategy for skilled readers, but is useful only when the earlier phonologically-based system has been practised sufficiently to achieve automaticity. (Adams, 1990).

In the early stages of reading, a child who relies on visual strategies needs to find a unique visual cue for each new word - a strategy doomed to failure as the vocabulary requirements become overwhelming (Freebody & Byrne, 1988; Tunmer & Hoover, 1993).

There has been debate about the relationship between phonemic awareness and naming speed. Wagner and Torgesen (1987) considered them both a reflection of a unitary phonological process; however, other research (Badian, 1993; Cornwall, 1992; Felton & Brown, 1990) found no correlation between the two skills. Recent studies by Torgesen, Wagner and colleagues (Wagner et al., 1993; Wagner et al., 1994; Torgesen et al., 1994) avoided some of the methodological problems plaguing earlier longitudinal-correlational and training studies. They employed multiple measures across a range of phonological processing tasks in longitudinal and cross sectional studies. Confirmatory factor analysis revealed five distinct but correlated phonological processing abilities. Their assessment involved multiple measures of each construct, and they used the resulting latent variables (representing the common variance among the measures) to preclude task specific, or error-variance.

They found two relatively uncorrelated latent abilities through their naming speed tasks, depending on whether the presentation was in a serial-trial, or isolated-trial format, that is, whether response-time was to digits (or letters) flashed serially onto a screen, or response time to name each of a group of digits (or letters) presented on cards. The significance of two such abilities is as yet unclear; however, it is consistent with other findings highlighting the predictive power of naming speed tasks (Bowers & Swanson, 1991; Catts, 1991; Cornwall, 1992; Davis & Spring, 1990; Felton, 1992; Tunmer & Hoover, 1993) for later reading ability.

It is also generally accepted that slow naming speed is characteristic of dyslexics (Wagner & Torgesen, 1987). This does not imply, however, that one can improve reading by providing practice at naming various items quickly. The focus on "underlying process variables" (Blachman, 1994) has been largely discredited (Arter & Jenkins, 1979). The appropriate question is not how to improve naming speed, but rather, how to improve reading in children with problems in accessing phonological information from their mental lexicon. In a small study involving both good and poor readers Rubin, Rotella, Schwartz, and Bernstein (1991) found that teaching phonological awareness skills to third grade children also improved their naming ability. While this result has no direct implications for improved reading it does support the view of Wagner, Torgesen and colleagues that their five phonological processing variables are related. Various researchers have examined this question, and Table 1 gives an indication of the correlations found in a selection of recent studies.

Table 1
Correlations Reported In Recent Studies

Variables	r	Study
Word Attack & Phonemic Awareness	0.56	Vandervelden & Siegel, 1995.
	0.54	Mann, 1993.
	0.53	Stage & Wagner, 1992.
	0.50	Wagner et al., 1994.
	0.48	Badian, 1993.
Word Attack & Digit Span	0.43	Bowers & Swanson, 1991.
	0.44	Bowey, Cain, & Ryan, 1992.
	0.32	Wagner et al., 1994.
Word Attack & Picture Naming	0.28	Bowers & Swanson, 1991.
	0.35	Bowers & Swanson, 1991.
	0.35	Wagner, et al., 1994.
	0.27	Vandervelden & Siegel, 1995.
Word Attack & Spelling	0.20	Badian, 1993.
	0.83	Vandervelden & Siegel, 1995.
	0.77	Stage & Wagner, 1992.
	0.70	Shankweiler, et al., 1996.
	0.63	Cornwall, 1992.
	0.62	Greenberg, et al., 1997.
Phonemic Awareness & Picture Naming	0.58	McDonald & Cornwall, 1995.
	0.14	Wagner, Torgesen, & Rashotte, 1994.
Phonemic Awareness & Digit Span	0.42	Bowers & Swanson, 1991.
	0.35	Wagner, et al., 1994.
	0.32	Bowey, 1996.
Phonemic Awareness & Spelling	0.59	Shankweiler, et al., 1996.
	0.57	Stage & Wagner, 1992.
	0.49	Vandervelden & Siegel, 1995.
Digit Span & Spelling	.2 - .3	Wagner, et al., 1993.

Another latent phonological ability is that of phonetic recoding in working memory. The beginning reader has to be able to decode a series of graphemes, and temporarily order them in a sound-based store in order to carry out the cognitively expensive task of blending. The efficiency with which this storage is performed optimises or diminishes the attentional capacity available for these blending and subsequent word-, and sentence-, comprehension tasks. The Wagner, Torgesen et al., studies used digit span (oral and visual), sentence

memory, and a distracter memory task to assess this ability. Gathercole, Willis, and Baddeley (1991) suggest that non-word repetition may be a purer measure, as it avoids the possibility of using lexical and semantic cues to assist recall. Wagner and Torgesen (1987), in their review of research, note that coding items phonetically is the major memory problem for poor readers; the deficit is a specific memory problem not a general one. The view that phonetic recoding in working memory is an important determinant of early reading success is supported by a number of researchers. (Catts, 1991; Felton, 1992; Hurford, Darrow, Edwards, Howerton, Mote, Schauf, & Coffey, 1993; Lindamood, Bell, & Lindamood, 1992; Shapiro, Nix, & Foster, 1990; Webster & Plante, 1992).

Gathercole et al. (1991) replicated their previous finding that phonological memory skills were also significantly associated with vocabulary knowledge. In their view the efficiency of the short term phonological store is a major determinant of ease of retrieval of a sound sequence from long term memory. Interestingly, the Wagner et al. (1994) longitudinal study found that the development rate for phonological memory paralleled that of vocabulary development in the first three years of schooling.

As with phonological coding in lexical access (or naming speed) it is not yet apparent how (if at all) weaknesses in this area might be addressed. Wagner and colleagues conclude that attempting to improve this skill through memory training, or mnemonic strategies has not been, nor is it likely to be, fruitful though they raised the interesting possibility that phonetic recoding in working memory may improve as reading skill develops. Their longitudinal study (Wagner, Torgesen, & Rashotte, 1994), however, failed to find any such trend. Although the rates of development across the five phonological processing abilities were somewhat uneven over the first three years of schooling, phonological memory was the slowest of the five. Nevertheless, there was considerable stability across the 5 variables over time, lending support to the view that they are causal to beginning reading, and not ephemeral individual differences soon submerged under the effects of schooling. This is not to argue that reading itself plays no role in enhancing phonological processing - only that it is not an overwhelming role (Wagner et al, 1993).

The two remaining latent phonological abilities (those most strongly related to later reading skill) comprise phonological awareness. They are phonological analysis (or segmentation), and phonological synthesis (or blending). It has been argued (Torgesen et al., 1992; Yopp, 1992) that synthesis develops earlier than analytic skills. Solomons (1992), and Caravolas and Bruck (1993) consider segmentation quite difficult for children below age five

or six, whereas Bryen and Gerber (1987) suggest that only by age six can 70% of children succeed in phonemic segmentation tasks. Certainly in the Torgesen et al. comparison of two phonological awareness training programs, blending skills (What word is this: /k/, /a/, /t/?) were more readily taught to first year students than were segmentation skills (Which of these three words begins the same as cat?). Their intervention study highlighted the need to teach both skills if promotion of decoding is the objective.

The Consequences of Phonemic Unawareness

There is ample evidence that students who do not make good initial progress in learning to read find it increasingly difficult ever to master the process. Stanovich (1986, 1988a, 1993a) outlines a model in which problems with early phonological skills can lead to a downward spiral, one in which even higher cognitive skills can be affected by slow reading development. Support for this model has been provided by a number of studies (Berninger, Thalberg, DeBruyn, & Smith, 1987; Bishop & Butterworth, 1980; Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1993; Nicholson, 1991b; Vellutino et al., 1994).

Stanovich uses the label Matthew effects (after the Gospel according to St. Matthew) to describe how, in reading, the rich get richer and the poor get poorer. Children with a good understanding of how words are composed of sounds (phonemic awareness) are well placed to make sense of our alphabetic system. Their rapid development of spelling-to-sound correspondences allows the development of independent reading, high levels of practice, and the subsequent fluency which is critical for comprehension and enjoyment of reading

Unfortunately children without good phonemic awareness tend to fall into a downward spiral of achievement in which initial lack of success in reading can develop into widespread cognitive deficits (Ceci, 1991). Contrary to the hope that initial slow progress is merely a maturational lag to be redressed by a developmental spurt at some later date, typically even relatively minor delays tend to become increasingly major over time (Stanovich, 1993a). A study by Juel (1988) reported a probability that a poor reader in Year 1 would still be so classified in Year 4 was .88. Jorm, Maclean, Matthews, and Share (1984) in their longitudinal study noted similar outcomes. Other studies by Hill (1995), and Shaywitz, and colleagues (1997), have supported the view that, without assistance, the prognosis is bleak for struggling beginning readers.

The implications of these findings are both disturbing and instructive. That there may be a specific cause of most inadequate reading progress is encouraging. Early intervention has the potential to significantly reduce failure, with its attendant personal and social cost. That an initially modular (phonological) deficit may broaden into further language, intellectual, and motivational deficits (Stanovich, 1986) is worrying for those attempting to alleviate the reading problems of students in mid-primary school and beyond. In these cases the consequences of the reading failure may remain even if the cause of the reading problem was successfully addressed. For teachers trying to provide effective remedial assistance to such pupils the Matthew effects help explain

- (a) why progress can be painfully slow,
- (b) why there may not be a significant change in general classroom performance consequent upon improved reading,
- (c) why teaching phonemic awareness to older children may not necessarily have as great an impact as anticipated.

Early Identification and Intervention.

Many researchers (Adams, 1990; Ball, 1993; Ball & Blachman, 1991; Blachman, 1994; Bradley & Bryant, 1983; Byrne & Fielding-Barnsley, 1989; Catts, 1991; Cunningham, 1990; Felton, 1993; Foorman, Francis, Novy, & Liberman, 1991; Hatcher, Hulme, & Ellis, 1994; Juel, 1993; Torgesen et al., 1994; Simmons, 1992; Stanovich, 1986, 1988b, 1992, 1993a) have noted the cost-beneficial effects of early intervention, and stressed the importance of primary prevention, for a variety of reasons - from the purely pragmatic or economic, to issues of social justice. Early intervention has long been regarded as logical, yet programs as intensive as the massive 1960's early intervention program, Head Start, have not achieved the success that was sought and anticipated. The value of empirical research since that time has been in the narrowing of the focus of the early intervention for reading - from a broad range of "readiness" activities that were largely peripheral to reading development - to a specific language area called phonemic awareness.

If early intervention is to be feasible, it is necessary to determine with a reasonable degree of accuracy the students who will comprise the problem reader cohort if they are not provided with appropriate additional assistance. It is important for at least two reasons. Ideally, screening devices should not miss many at-risk students (i.e., they should have few false negatives) because the students in need deserve assistance. Secondly, screening should not include many students who would cope well without additional help (i.e., false positives) because scarce resources need to be applied where they will have the optimum impact.

Tests measuring phonological skills are beginning to assume importance because of their capacity to add discrimination power to screening batteries (Badian, 1994; Cornwall, 1992; Felton, 1992; Hurford et al., 1993; Hurford, Schauf, Bunce, Blaich, & Moore, 1994; Spector, 1992). Some of these recent studies have demonstrated excellent results by including phonological tests in a battery to predict problems in reading-acquisition. Hurford et al. (1994) assessed 170 school beginners, and predicted with 100% accuracy which students would be diagnosed with a reading disability two years later. They used phoneme deletion,

phonological discrimination, IQ, and pseudo-words. Badian (1994) assessed 118 preschoolers mid-year and successfully predicted 91% of good or poor readers two years later. She used phonological awareness, naming speed, and an orthographic matching task. Majsterek and Ellenwood (1995) noted that of 17 procedures frequently used to identify preschoolers with learning disability, none specifically targeted phonological awareness. In their study two measures, sound blending and rhyme detection, were significantly related to word attack skills three years later, at the end of Year 2. Stuart (1995) found that sound to letter matching at the start of school predicted 93% of reading progress at the end of Year 1, and seven months later. Mann (1993) used a simple phoneme segmentation test in the first year of school, and found it a good predictor of reading progress a year later.

Research and Education Systems.

Research has brought us to the point that early phonemic awareness intervention in kindergarten and beyond may preclude the debilitating effects of early failure for many students (Foorman, Francis, Beeler, Winikates, & Fletcher, 1997; Mazzocco, et al., 1997). In practice, the necessity for heavy investment in one-to-one remedial programs (such as Reading Recovery) could be markedly reduced. It has been estimated that on average 16% (Prior et al., 1994) of our students could be classified as reading disabled. The recent studies have suggested that, with early phonemic awareness intervention, this could be reduced to a figure that would allow seriously reading disabled students to be provided with more intensive (and extended) assistance, and reduce the debilitating Matthew effects.

Unfortunately such emphasis on the structure of our language does not sit easily with many of those who support the predominant model of teaching, Whole Language (Ball, 1993). Leaving aside philosophical objections, there are no insurmountable reasons why such emphases could not form part of an integrated Whole Language program (Castle, Riach, & Nicholson, 1994; Vellutino, 1991). Research has shown that all children of at least low average intellectual ability (Adams, 1990; Bateman, 1991) can be taught to read given the strategies and the will. For example, Felton (1993) followed the progress of at-risk students in their first and second year of school. They had phonological processing problems (either awareness, or naming deficits, or both) and were randomly assigned to two reading programs emphasising code, or context. Results unambiguously favoured the code-emphasis regardless of the phonological processing deficit. Felton's work emphasised the mutually facilitative effects of phonemic awareness emphasis, and code-emphasis reading instruction. She concluded that:

- (i) at-risk children should be identified in their first school year.
- (ii) phonemic awareness training should be available for those students, and taught using a direct instruction approach.
- (iii) structured code emphasis teaching should follow, using controlled vocabulary. Explicitly taught strategies such as blending (rather than guessing strategies) should be promoted.
- (iv) a significant portion of the school day should be assigned to direct instructional activities.
- (v) teaching the onset-rime distinction will hasten students progress from letter-by-letter decoding to skilled reading.

- (vi) reading, writing and spelling instruction should be integrated, with correct spelling emphasised.
- (vii) it should be recognised that at-risk students may need three years of direct instruction in basic reading skills.
- (viii) it must be recognised that teaching to mastery is insufficient, and provision should be made for adequate opportunities for the practice necessary to achieve automaticity.

Felton's important conclusions represent a confluence of the research in phonemic awareness, and that in effective teaching. It represents a position on early reading instruction that is vastly different (in almost every respect) from the nationally popular Whole Language approach discussed earlier. It also assumes that teachers have the necessary training and understanding of phonemic awareness to allow for its implementation. Research by Lindamood (1993), and Moats (1994b) suggests that this assumption may not be warranted, and that current teacher training priorities do not allow for pre-service instruction in these areas. Hence, teachers may need to be retrained if the results of research into beginning reading are to be put into practice successfully.

Apart from problems of ensuring that teachers are trained to make use of the optimum strategies for reading instruction, there is concern that students at-risk may be resistant to attempts to improve their levels of phonemic awareness (Wagner, Torgesen, et al., 1993, 1994). This suggests that there is a considerable amount of empirical research to be completed regarding the optimum methods of teaching the phonological skills necessary for reading success. The next chapter reviews the approach to basic skill teaching called Direct Instruction, an approach likely to play a significant role in future effective programs.

CHAPTER FIVE: THE DIRECT INSTRUCTION MODEL

The failure of the school system to provide effectively for the basic skill development of each its pupils is of concern to both the general and research communities. It is especially salient for those inclined towards empiricism as there are behavioural approaches to teaching with excellent research support that could make a major contribution to the prevention and alleviation of this distressing problem. Unfortunately, the evidence for the effectiveness of such programs has been largely ignored by educational decision-makers. One example of this group of behavioural-based models is known as Direct Instruction. It is contrasted with the approach to teaching called Whole Language (described in Chapter 3), one with little empirical support and major theoretical weaknesses. A broader issue, considered within the context of this educational problem area, concerns researchers' responsibility for the dissemination and application of their work within the community.

One of the most thoroughly researched educational models is Direct Instruction. There is ample evidence of its effectiveness for a wide range of student learning problems. It differs from Whole Language in its assumptions about the teaching process, about learner characteristics, and about the means of syllabus construction; in fact, it could be described as the antithesis of Whole Language, but has had very little impact upon the Australian school system.

Although their [Whole Language] theories lack any academically acceptable research base they continue to dominate educational policy. Direct Instruction models are ignored notwithstanding the huge body of research that indicates that direct instruction is vastly superior if basic skills and knowledge are the goal (Weir, 1990, p.30).

The Direct Instruction model lauded in Follow Through had its beginnings in the early 1960's through the work of Carl Bereiter and Siegfried Engelmann. The subsequent involvement of Wes Becker and Doug Carnine among others led to the publication of a number of teaching programs in 1969. The programs share a common teaching style readily observable to any classroom visitor. The instruction takes place in small groups with a teacher directing activities with the aid of a script, and students are actively involved in responding to a fast paced lesson during which they receive constant feedback. Programs are designed according to what, not whom, is to be taught. Thus, all children work through the same sequence of tasks directed by a teacher using the same teaching strategies. Individual

differences are accommodated through different entry points, reinforcement, amounts of practice and correction strategies (Gregory, 1983).

Characteristics of the Direct Instruction Model

There are a number of important characteristics of Direct Instruction programs (Becker, 1977). It is assumed that all children can learn and be taught, thus failure to learn is viewed as failure to teach effectively (Engelmann, 1980). Children whose progress is restricted must be taught to learn faster through a focus on features of teaching designed to improve efficiency of instruction. These features derive from the design of instruction, and from process variables such as how the curriculum is implemented. Curriculum is designed with the goal of "faultless instruction" (Engelmann, 1980), that is, sequences or routines for which there is only one logical interpretation. The designer's brief is to avoid ambiguity in instruction - the focus is on logical-analysis principles. These principles allow the organisation of concepts according to their structure and the communication of them to the learner through the presentation of positive and negative examples.

Engelmann (1980) highlights four design principles:

(i) Where possible teach a general case, that is, those skills which, when mastered, can be applied across a range of problems for which specific solutions have not been taught, for example, decoding regular words. These generalisations may be taught inductively, by examples only, or deductively, by providing a rule and a range of examples to define the rule's boundaries.

(ii) Teach the essentials. The essentials are determined by an analysis of the skills necessary to achieve the desired objective. There is an underlying assertion that, for reading, it is possible to achieve skilled reading by task analysis and the teaching of subskills within a cumulative framework. Advocates of a "Whole Language" perspective would disagree with the possibility or desirability of teaching in this manner.

(iii) Keep errors to a minimum. Direct Instruction designers consider errors counter-productive and time-wasting. For remedial learners a high success rate is useful in building and maintaining motivation lost through a history of failure. This low error rate is achieved by the use of the instructional design principles elucidated in Theory of Instruction (Englemann & Carnine, 1982) and by ensuring students have the pre-skills needed to commence any program (via a placement test).

(iv) Adequate practice. Direct Instruction programs include the requirement for mastery learning (usually above 90% mastery). Students continue to focus on a given task

until that criterion is reached. The objective of this strategy is the achievement of retention without the requirement that all students complete the identical regimen. The practice schedule commences with massed practice, shifting to a spaced schedule. The amount of practice decreases as the relevant skill is incorporated into more complex skills. Advocates of Direct Instruction argue that this feature of instruction is particularly important for low-achieving students and is too often allowed scant regard (Engelmann, 1980). Whereas, this emphasis on practice may be unfashionable, there is considerable supporting research, and a number of effective schools are increasingly endorsing its importance (Rist, 1992). "The strategies that have fallen out of style, such as memorising, reciting and drilling, are what we need to do. They're simple - but fundamental - things that make complex thinking possible" (p. 19).

Roots of the Direct Instruction Model

It is these principles of instructional design that sets Direct Instruction apart from traditional and modern behavioural approaches to teaching. However, the model does share a number of features with other behavioural approaches (e.g., reinforcement, stimulus control, prompting, shaping, extinction, fading), and with the effective teaching movement (mastery learning, teacher presentation skills, academic engaged time, and correction procedures). These latter features have been researched thoroughly over the past 20 years, and have generally been accepted as comprising "direct instruction" (Gersten, Woodward, & Darch, 1986).

Rosenshine (1979) used the expression to describe a set of instructional variables relating teacher behaviour and classroom organisation to high levels of academic performance for primary school students. High levels of achievement were related to the amount of content covered and mastered. Hence the pacing of a lesson can be controlled to enhance learning. Academic engaged time refers to the percentage of the allotted time for a subject during which students are actively engaged. A range of studies (Rosenshine & Berliner, 1978) has highlighted the reduction in engagement that occurs when students work alone as opposed to working with a teacher in a small group or as a whole class. The choral responding typical of DI programs is one way of ensuring high student engagement. The author once counted 300 responses in the 10 minutes of teacher directed decoding activity in a Year 7 reading group (Hempenstall, 1990).

A strong focus on the academic was found to be characteristic of effective teachers. Non-academic activities, while perhaps enjoyable or directed at other educational goals, were consistently negatively correlated with achievement. Yet, in Rosenshine's (1980) review of studies it was clear that an academic focus rather than an affective emphasis produced classrooms with high student self-esteem and a warm atmosphere. Less structured programs and teachers with an affective focus had students with lower self esteem. Teacher centred rather than student centred classrooms had higher achievement levels. Analogously, teachers who were strong leaders and did not base their teaching around student choice of activities were more successful. Solomon and Kendall (1976) cited in Rosenshine (1980) indicated that permissiveness, spontaneity and lack of classroom control were " ... negatively related, not only achievement gain, but also to positive growth in creativity, inquiry, writing ability, and self esteem for the students in those classrooms" (p. 18).

The instructional procedure called demonstration-practice-feedback (sometimes model-lead-test) has strong research support (Rosenshine, 1980). This deceptively simple strategy combines three elements of teaching strongly related to achievement in one general model. It comprises an invariant sequence in which a short demonstration of the skill or material is followed by guided practice during which feedback is provided to the student (and further demonstration if necessary). The second phase usually involves response to teacher questions about the material previously presented. It would appear that the overlearning this phase induces is particularly valuable. The third phase, that of independent practice, is evaluated by the teacher.

Medley's (1982) review indicated the efficacy for low SES students of a controlled practice strategy involving low cognitive level questions, a high success rate (above 80%), and infrequent criticism. Thus, the popularity among teachers of high cognitive level question implicit in discovery learning models is difficult to justify empirically. These high level questions require students to manipulate concepts without having been shown how to do so. Research on discovery approaches has indicated a negative relationship with student achievement. Winnie's (1979) review of 19 experimental studies on higher order questions made this point very strongly, as does Yates (1988).

To summarise the findings of research into teacher variables with a positive impact on student learning, Rosenshine and Berliner (1978) provide a definition for direct instruction, a concept providing part of the theoretical basis for Direct Instruction.

Direct instruction pertains to a set of teaching behaviours focused on academic matters where goals are clear to students; time allocated for instruction is sufficient and continuous; content coverage is extensive; student performance is monitored; questions are at a low cognitive level and produce many correct responses; and feedback to students is immediate and academically oriented. In direct instruction, the teacher controls the instructional goals, chooses material appropriate for the student's ability level, and paces the instructional episode (p. 7).

Direct Instruction has developed into a comprehensive system of instruction covering many skill areas: reading, mathematics, language, spelling, microcomputing, writing, reasoning, and a variety of other school subjects including chemistry, critical reading, social studies, history. Thus, the approach that initially restricted its emphasis to basic skills is now expanding into higher order skills (Kinder & Carnine, 1991), has a strong research base, and has unfulfilled promise as part of a solution to the problems of illiteracy in our community.

Evaluation of the Direct Instruction Model

Chapter 2 contained a description of a very large national evaluation of different approaches to teaching, entitled Operation Follow Through. This evaluation indicated that the Direct Instruction approach was particularly effective. Additional to the Follow Through data, evaluation of Direct Instruction programs has been very intensive. For example, Fabre (1984) compiled an annotated bibliography of almost 200 studies completed prior to 1984. For the most part, research findings have been very impressive. Notable positive reviews of outcome research are provided by Gersten, 1985; Gregory, 1983; Kinder and Carnine, 1991; Lockery and Maggs, 1982; White, 1988. See later for contrary views.

Whereas, Direct Instruction was originally designed to assist disadvantaged students, its emphasis on task characteristics and effective teaching principles may transcend learner characteristics and be valuable across a range of learners. Lockery and Maggs (1982) reviewed research indicating success with average children, those with mild, moderate or severe skill deficits, those in resource rooms, withdrawal classes and special classes in regular schools, disadvantaged students (including aboriginal and children whose first language is not English), students in special facilities for mild, moderate and severe intellectual disability, and physical disabilities.

Gersten (1985) in his review of studies involving students with a range of disabilities concluded that Direct Instruction tended to produce higher academic gains than traditional approaches. He also suggested that the mastery criterion (in excess of 90%) may be particularly important for special education students, and called for more formative evaluation where only one instructional variable is manipulated, and also, for more instructional dimensions research to highlight those variables alone or in company that are associated with academic gains. Gersten referred to the Leinhardt, Zigmond, and Cooley (1981) study with 105 learning disabled students. The authors noted that three teaching behaviours were strongly associated with student progress in reading - the use of reinforcers, academic focus, and a teacher instruction variable involving demonstration, practice and feedback. Each of these is critical to the definition of direct instruction (Rosenshine, 1979) and supports the assertion that there are teacher behaviours that transcend student characteristics. This study was the first to demonstrate that specific direct instruction principles have value for learning disabled students.

White's (1988) meta-analysis of studies involving learning disabled, intellectually disabled, or reading disabled students restricted its focus to those studies employing

equivalent experimental and comparison groups. White reported an effect size of 0.84 standard deviation units for the DI over comparison treatments. This is markedly above the 0.25-0.33 standard for educational significance of an educational treatment effect (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977). White concluded that " ... instruction grounded in Direct Instruction theory (Engelmann & Carnine, 1982) is efficacious for both mildly and moderately/severely handicapped learners, and in all skill areas on which research has been conducted" p. 372.

Further support for the approach comes from Kavale (1990). His summary of research into direct instruction and effective teaching concludes that they are five to ten times more effective for learning disabled students than are practices aimed at altering unobservable learning processes such as perception. Binder and Watkins (1990) describe Direct Instruction (along with Precision Teaching) as the approaches best supported by research to address the problems of teaching found in the English-speaking world.

Recently Hendrickson and Frank (1993) provided the bold prediction that

The decade of the 1990's will witness, in classrooms serving students with mild mental retardation, the implementation of a group of instructional methods often referred to as effective teaching practices or direct instruction, if we heed the literature published in this area over the past 15 years. (p. 11)

Criticisms of Direct Instruction

Despite the long history of empirical support for Direct Instruction, unsurprisingly there have also been criticisms. These have been based on a number of different grounds:

- (a) DI is an IBM conspiracy to oppress the masses (Nicholls, 1980).
- (b) It causes delinquency (Schweinhart, Weikart, & Lerner, 1986. Further, its "side effects may be lethal" (Boomer, 1988, p. 12).
- (c) Its view of the reading process is wrong (Gollash, 1980).
- (d) It is incompatible with other more important principles:
 - (i) Normalisation (Penney, 1988).
 - (ii) The wholistic nature of reading (Goodman, 1986; Giffen, 1980)
 - (iii) A naturalistic educational paradigm (Heshusius, 1991).
 - (iv) Flexible reciprocal child-teacher interaction (Ashman & Elkins, 1990).
 - (v) Teacher professionalism (McFaul, 1983).
- (e) The success of DI is illusory, based on tests that do not measure real reading (Cambourne, 1979).
- (f) Other approaches are more effective, for example, Whole Language (Weaver, 1991), discovery learning (Bay, Staver, Bryan, & Hale, 1992); or as effective as DI (Kuder, 1990; O'Connor et al., 1993).
- (g) It may be inappropriate for certain sub groups.
 - (i) Those in special education (Heshusius, 1991; Kuder, 1991; Penney, 1988).
 - (ii) Those with certain learning styles, for example, those with an internal locus of control (McFaul, 1983; Peterson, 1979).
 - (iii) Those of high ability (Peterson, 1979).
- (h) Its use is best restricted to basic skill development (Peterson, 1979).
- (i) It is best used in conjunction with other approaches (Delpit, 1988; Gettlinger, 1993; Harper, Mallette, Maheady, & Brennan, 1993; Spiegel, 1992; Stevens, Slavin, & Farnish, 1991).
- (j) Students might not find it acceptable (Reetz & Hoover, 1992).

Of the literature critical of the model, much is based on philosophical issues concerning reality and power; on theoretical issues such as the nature of the learning process, the role of teaching, or issues of measurement. Of the few studies in which alternative approaches have proved equivalent or superior, issues of treatment fidelity have arisen. It is rarely made clear whether the model described is the Direct Instruction model or a direct

instruction clone of unknown rigour. Nor is it usually specified whether the teachers of any Direct Instruction program have been provided with the training required to ensure the programs are presented according to the presentation protocols.

It is of interest that the debate on Direct Instruction has become much more widespread in recent years. An issue of *Education and Treatment of Children* (Becker, 1988) was devoted to Direct Instruction. The National Reading Conference in the USA has regular sessions on the pedagogical impact and appropriateness of Direct Instruction (Kameenui & Shannon, 1988). The *Journal of Learning Disabilities* (1991) devoted two issues (Vol 24, Nos 5, 6) to "sameness analysis" - an instructional design principle central to Direct Instruction (Englemann & Carnine, 1982). In recent years writers of texts on teaching (Becker, 1986), special education (Cole & Chan, 1990; Gable & Warren, 1993, Greaves & McLaughlin, 1993; Scruggs & Wong, 1990; Wolery, Ault, & Doyle, 1992), and educational psychology texts (Joyce, Weil, & Showers, 1992; Kameenui & Simmons, 1990; Tuckman, 1991) have included Direct Instruction as a legitimate approach to a range of educational problems. This represents the increasing academic acceptance of the model that until the mid-1980's was virtually ignored by researchers and writers other than advocates from, or influenced by, the University of Oregon. From one of the most respected writers and researchers on the problems of learning disability (a term coined by Kirk and Bateman in 1962) comes the highest praise. "The documented success of Siegfried Engelmann and his colleagues' direct instruction reading programs with thousands of hard-to-teach and high risk children is unsurpassed in the annals of reading history" (Bateman, 1991, p.11).

Despite the controversy, Direct Instruction research and program development continues. It no longer has a sole emphasis on instructional design for basic skills such as reading, spelling, maths, language, and writing--but has broadened its area of application to include higher order skills, for example, literary analysis, logic, chemistry, critical reading, geometry and social studies (Carnine, 1991; Casazza, 1993; Darch, 1993; Grossen & Carnine, 1990b; Kinder & Carnine, 1991). Use has been made of technology through computer-assisted instruction, low cost networking and videodisc courseware (Kinder & Carnine, 1991); and, researchers have begun to test the model in non-English speaking countries, for example, third world countries (Grossen & Kelly, 1992), and Japan (Nakano, Kageyama, & Kioshita; 1993). It has also shown promise in recent research on teaching a most challenging group of students--school aged children with TBI, traumatic brain injury (Glang, Singer, Cooley, & Tish, 1992).

There seems little doubt that it will continue to be a viable and productive model throughout the 1990's, although there remains a question mark over the extent of adoption by the school system. The major hurdle continues to be its lack of attractiveness for educators, and resultant absence of penetration into classrooms. Part of that problem relates to the popularity of the Whole Language philosophy of teaching. The two models are very disparate, presenting polar extreme views on both the content of a reading program, and the approach to instruction (Gersten & Dimino, 1993).

Problems of Acceptance in Education

Hendrickson and Franks's prediction is brave because, despite impressive research support, DI has made little impact in regular or special education. Maggs and White (1982) wrote despairingly, "Few professions are more steeped in mythology and less open to empirical findings than are teachers" (p. 131). Murphy (1980) considered that behavioural consultants should be agents of change, but are generally naive about the politics of change in organisations. He suggests that an improved understanding of organisational contingencies would enhance the likelihood of successful implementation. Barnes (1985) suggested five popular perceptions for the approach's lack of acceptance in education.

- (a) Its phonic basis conflicts with the popular "Whole Language" philosophy.
- (b) Its highly structured scripted lessons are an insult to trained teachers.
- (c) It over-emphasises basic skills and ignores higher order goals.
- (d) Its emphasis on the teacher's responsibility for learning outcomes threatens those teachers holding the view that student performance is largely determined by the child's genetic or family history.
- (e) The structure implies a crushing routine that bores teachers. Students become bored either for the same reason or due to the teacher's resultant lack of enthusiasm.

Barnes does not accept the validity of these objections but highlights them as obstacles to be overcome.

Fields (1986) posits the "practicality ethic" as the determining feature of programs likely to be readily adopted. Can the recommendation be easily translated into practice, that is, in the classroom? Is the recommendation congruent with the teacher's philosophy or goals? How difficult in time and effort is implementation? Fields sees problems for DI in each of these areas and recommends a fall-back position. He suggests differing levels of implementation - from the total package for schools with a major priority to lift student

achievement and an acceptance of the model - through to a simple acceptance of a more active teaching style to be adopted in their classrooms.

Ruddell and Sperling (1988) express a general concern at the gulf between literacy research findings and teachers' practice. They call for research aimed at discovering why empirically proven practices are "thwarted, undermined, or ignored in the classroom" (p. 319). The concern is even more impelling if one accepts Roger's (1983, cited in Ruddell & Sperling) assertion that there is often a period of 25 to 35 years between a research discovery and its serious implementation.

Solity (1991) notes the aspects of Direct Instruction unappealing to teachers. He views the problem within the wider context of the negative view many teachers have of behavioural approaches in general. He considers the method of introduction of behavioural concepts as crucial to acceptance, and cites examples of "softer" language being more acceptable. Gersten and Guskey (1985) argue that teachers' methods have evolved largely through experiences in their own classroom, and a model that requires a significant change from that practice will evoke reluctance. In their studies, teachers' philosophies which were generally antithetical to Direct Instruction became consonant with those of DI following successful program implementation. Hence attitude change followed rather than preceded behaviour change. They argue that trying to change attitudes through, for example, presenting research data is unlikely to be successful. On the other hand, a well organised pilot program in the school, run by a respected teacher with good consultant support, is likely to produce gains difficult to ignore in children personally known to the teachers. The salience of change in known children combined with strong instructional leadership from the school administration may lead to a change in teacher behaviour. As in Gersten and Guskey's study, the initial reluctance may be transformed into a new energy-giving direction in teaching.

Lindsley (1992) is quite scathing in addressing the question of why effective teaching tools aren't widely adopted. He considers that teachers have been:

... seduced by natural learning approaches. ... Most educators have bought the myth that academic learning does not require discipline - that the best learning is easy and fun. They do not realise that it is fluent performance that is fun. The process of learning, of changing performance, is most often stressful and painful. (p. 22)

Gable and Warren (1993) have also noted that the potential role of behavioural science in general, but with particular emphasis to education, has been largely ignored by decision-makers and even by many practitioners. Carnine (1991) laments that decision-makers lack a

scientific framework, and are inclined to accept proposals based on good intentions and unsupported opinions. Meyer (1991, cited in Gable & Warren), however, blames the research community for choosing restricted methodology (e.g., single subject design), and for being too remote from classrooms. She believes greater attention will be paid when the credibility of research is improved.

On the other hand, perhaps it is the tendency of empiricists to place caveats on their findings, as opposed to the wondrous claims of ideologues and faddists unrestrained by scientific ethics, that makes decision-makers wary. Fister and Kemp (1993) consider several likely obstacles, important among them being the absence of an accountability link between decision-makers and student achievement. Such a link seems unlikely without a regular mandated state or national test program. They also apportion some responsibility to the research community for failing to appreciate the necessity nexus between research and its adoption by the relevant target group. The specific criticisms include a failure to take responsibility for communicating findings clearly, with the end-users in mind.

Researchers have often validated practices over too brief a time-frame, and in too limited a range of settings to allow general program adoption across settings. Without considering the organisational ramifications (such as staff and personnel costs) adequately, the viability of even the very best intervention cannot be guaranteed. The methods of introduction and staff training in innovative practices can have a marked bearing on their adoption and continuation.

Fister and Kemp (1993) argue that researchers have failed to meet their own criterion by not incorporating research-validated staff-training procedures, and organisational analysis in their strategies for promoting program adoption. Their final criticism involves the rarity of the establishment of model sites exemplifying excellent practice. When prospective adoptees are able to see the reality rather than the rhetoric of a program they are arguably more likely to take the (often uncomfortable) steps towards adoption. In addition, it is possible to discuss with on-site teachers the realities of being involved in the innovation. Woodward (1993) points out that there is often a gulf between researchers and teachers. Researchers may view teachers as unnecessarily conservative and resistant to change, while teachers may consider researchers as unrealistic in their expectations, and lacking in understanding of the school system and culture. Teachers may also respond defensively to calls for change because of the implied criticism of their past practices, and the perceived devaluation of the professionalism of teachers (in that other professions are determining their teaching practices).

Thus, there are three groups whom researchers need to be able to influence if their innovations are to be adopted. At the classroom level, teachers are the focal point of such innovations, and their competent and enthusiastic participation is required if success is to be achieved. As schools become increasingly self-governing, school principals and school councils are also in a position to influence policy within their setting. At the broader system level, decision makers presumably require different information and assurances about the viability of change of practice (cost being fundamental). Given that many researchers have neither the funding, the interest, and perhaps the skill to promulgate their findings, it is clear that the relationship between science, school practice, and government policy-making will remain vexed.

CHAPTER SIX: STUDY RATIONALE AND RESEARCH QUESTIONS

Study Rationale

It is acknowledged that phonemic awareness is a necessary condition for mastery of reading in an alphabetic system. A degree of phonemic awareness may be present before instruction, as in children with a supportive background of literacy-promoting activities. It may be taught to young pre-school or prep aged children prior to the introduction of formal literacy training. It may be taught to children simultaneous with such literacy instruction.

An alternative yet to be adequately explored is the employment of a code-emphasis reading program (without a dedicated phonemic awareness program), with subsequent monitoring of the development of phonemic awareness as the program is implemented. Phonemic awareness is sometimes mistakenly seen as synonymous with code emphasis or phonics. Phonemic awareness refers to conscious access to the phonemic level of the speech stream, and some ability to manipulate cognitively representations at this level. Phonemic awareness has also been referred to as: phonological awareness, acoustic-phonetic skill, auditory analysis, sound categorisation, phoneme segmentation, and phonemic analysis. Phonics is a set of rules of grapheme-phoneme correspondence that provides clues to the pronunciation of the written word, that is, the sounds the letters make. It also refers to a method of teaching reading that emphasises such rules.

This latter approach represents the lowest cost option, but necessitates progressive or continuous assessment to indicate which, if any, children are failing to make adequate progress purely by exposure to the phonics program.

An important research question then arises concerning the prediction of which students will make progress in phonemic awareness purely by exposure to the reading program, and which students will require phonemic awareness training in addition to their reading program. The ability to discriminate accurately would reduce the cost in time and money of providing phonemic awareness training to the large cohort of students entering school with low phonemic awareness, and to schools wishing to provide remedial assistance to older struggling students.

If the use of a phonics program is sufficient to induce the alphabetic principle in some students, then it may be that there are differing degrees of resistance to phonemic awareness,

and the limits of environmental manipulation should be ascertained before assuming that the absence of phonemic awareness is due necessarily to neurological deficit.

Can phonemic awareness be induced through a code-emphasis program? How much phonemic awareness? There may be a threshold level of phonemic awareness necessary to take advantage of a code emphasis program - "Phonics instruction is not effective unless children already have (or quickly develop) some phonemic awareness at the beginning of first grade" (Juel, 1993, p. 97). What is the optimum strategy for older children? Phonemic awareness and code-instruction? One or other? A hybrid? For whom will it be successful? Can we predict for whom?

This study charts the progress of 206 students identified by nine schools as making particularly slow progress in reading skills. As such, it may not be a representative sample of all students with reading difficulties. In particular, the ratio of boys to girls identified by teachers is often of the order two, three or four to one, whereas it is generally acknowledged that the true incidence is closer to one to one (Prior et al., 1995). There were 150 males (100 experimental and 50 control) and 56 females (34 experimental and 22 control) in total.

Additional to the 9 individual school cohorts, was a charitable organisation, Orana Family Services, that provides an educational service to surrounding schools. Students attend the Education Resource Centre for four sessions per week to participate in the same reading program as did students from the other 9 schools.

The students' progress is compared to that of students with similar levels of reading difficulty (as determined in the same manner by their schools) who are on a waiting list to be involved in the program at a later time. The comparison groups are drawn from the same schools participating in the reading program, thus reducing the chances of socio-economic or other differences confounding the comparison. The group is best described as a non-equivalent control group (Cooke & Campbell, 1979) as the students are not randomly assigned to their respective groups, but are convenience samples.

The cohorts identified by their classroom teachers were given a placement test to determine the level of their reading skills, as groups for any given level of the reading program need to be relatively similar in their entry skills. The placement test ensures first, that student and teacher time will not be wasted by providing a program to a student who already has mastered the outcome objectives; and second, ensures that students have sufficient entry skills to achieve initial and sustained success in the program. Placement tests

were performed at the school by the relevant designated teacher, under the supervision of the author.

The *Corrective Reading: Decoding* program is a Direct Instruction program sequence designed to improve the skills of students who have thus far failed to make adequate progress in reading. It is a remedial rather than a beginning reading program.

The Research Questions

Question 1

- (a) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a statistically significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a statistically significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a statistically significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- A. For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a statistically significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison
- (e) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a statistically significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 2: For *Corrective Reading Level A* alone

- (a) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a statistically significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

- (b) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a statistically significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a statistically significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (d) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a statistically significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (e) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a statistically significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 3: For *Corrective Reading Level B* alone

- (a) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a statistically significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a statistically significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a statistically significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (d) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a statistically significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (e) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a statistically significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 4: Are there similar outcomes for the *Teach Your Child to Read in 100 Easy Lessons* program?

- (a) For disabled readers, does participation in the *100 Lessons* program lead to a statistically significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *100 Lessons* program lead to a statistically significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *100 Lessons* program lead to a statistically significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (d) For disabled readers, does participation in the *100 Lessons* program lead to a statistically significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (e) For disabled readers, does participation in the *100 Lessons* program lead to a statistically significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Are the effects of educational significance?

These questions examine the same processes, but seek to establish whether any observed changes are of educational importance through the examination of effect sizes.

Question 5: For *Corrective Reading Level A and B Together*:

- (a) For disabled readers, does participation in the *Corrective Reading* (Levels A and B considered together) program lead to an educationally significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *Corrective Reading* (Levels A and B considered together) program lead to an educationally significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *Corrective Reading* (Levels A and B considered together) program lead to an educationally significant increase in naming speed

(another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(d) For disabled readers, does participation in the *Corrective Reading* (Levels A and B considered together) program lead to a educationally significant increase working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(e) For disabled readers, does participation in the *Corrective Reading* (Levels A and B considered together) program lead to a educationally significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 6: For *Corrective Reading Level A*:

(a) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a educationally significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(b) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a educationally significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(c) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a educationally significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(d) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a educationally significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(e) For disabled readers, does participation in the *Corrective Reading Level A* program lead to a educationally significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 7: For *Corrective Reading Level B*

- (a) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a educationally significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a educationally significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a educationally significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (d) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a educationally significant increase in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (e) For disabled readers, does participation in the *Corrective Reading Level B* program lead to a educationally significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 8: Are there similar effect size outcomes for the *Teach Your Child to Read in 100 Lessons* program?

- (a) For disabled readers, does participation in the *100 Lessons* program lead to a educationally significant increase in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *100 Lessons* program lead to a educationally significant increase in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (c) For disabled readers, does participation in the *100 Lessons* program lead to a educationally significant increase in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (d) For disabled readers, does participation in the *100 Lessons* program lead to a educationally significant increase in working memory (another phonological process)

compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(e) For disabled readers, does participation in the *100 Lessons* program lead to a educationally significant increase in spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 9: Are there differential sex effects for the A and B groups considered together?

(a) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to sex differences in the phonemic awareness effect sizes measured ?

(b) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to sex differences in the phonological recoding (word attack) effect sizes measured?

(c) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to sex differences in naming speed (another phonological process) effect sizes measured ?

(d) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to sex differences in working memory (another phonological process) effect sizes measured ?

(e) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to sex differences in spelling (a partly phonological process) effect sizes measured ?

Question 10: Are there differential age effects for the A and B groups considered together?

(a) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to age differences in the phonemic awareness effect sizes measured?

(b) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to age differences in the phonological recoding (word attack) effect sizes measured?

(c) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to age differences in naming speed (another phonological process) effect sizes measured?

- (d) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to age differences in working memory (another phonological process) effect sizes measured?
- (e) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to age differences in spelling (a partly phonological process) effect sizes measured?

Question 11: Are there differential school effects for the A and B groups considered together?

- (a) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to school differences in the phonemic awareness effect sizes measured?
- (b) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to school differences in the phonological recoding (word attack) effect sizes measured?
- (c) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to school differences naming speed (another phonological process) effect sizes measured?
- (d) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to school differences in working memory (another phonological process) effect sizes measured?
- (e) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to school differences in spelling (a partly phonological process) effect sizes measured?

Question 12: Individual differences in the effects

- (a) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a higher percentage of students achieving a significant increase (1 SD) in phonemic awareness compared to similarly disabled readers in a waitlist comparison group who do not receive the program?
- (b) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a higher percentage of students achieving a significant increase (1 SD) in phonological recoding (word attack) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(c) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a higher percentage of students achieving a significant increase (1 SD) in naming speed (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(d) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a higher percentage of students achieving a significant increase (1 SD) in working memory (another phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

(e) For disabled readers, does participation in the *Corrective Reading* program (Levels A and B considered together) lead to a higher percentage of students achieving a significant increase (1 SD) spelling (a partly phonological process) compared to similarly disabled readers in a waitlist comparison group who do not receive the program?

Question 13: Were the students at Orana more severely reading disabled than the other students?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In naming speed (another phonological process)
- (d) In working memory (another phonological process)
- (e) In spelling (a partly phonological process)

Question 14: Were the effect sizes achieved at Orana greater than for other schools?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In naming speed (another phonological process)
- (d) In working memory (another phonological process)
- (e) In spelling (a partly phonological process)

Question 15: For disabled readers participating in the *Corrective Reading* program (Levels A and B in sequence), were the effect sizes comparable in the two programs?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In naming speed (another phonological process)

- (d) In working memory (another phonological process)
- (e) In spelling (a partly phonological process)

Question 16: For disabled readers participating in the *Corrective Reading* program, were there differences between the Level A students' pretest scores and those of the Level B students?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In naming speed (another phonological process)
- (d) In working memory (another phonological process)
- (e) In spelling (a partly phonological process)

Question 17: For disabled readers participating in the *Corrective Reading* program (either Level A or Level B), were the effect sizes comparable in the two programs?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In naming speed (another phonological process)
- (d) In working memory (another phonological process)
- (e) In spelling (a partly phonological process)

Question 18: For disabled readers participating in the *Corrective Reading* program (Levels A and B considered together), to what degree do the students' scores approach the normal (interquartile) range?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In working memory (another phonological process)
- (d) In spelling (a partly phonological process)

Question 19: For disabled readers participating in the *Corrective Reading* program (Level A), to what degree do the students' scores approach the normal (interquartile) range?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In working memory (another phonological process)

- (d) In spelling (a partly phonological process)

Question 20: For disabled readers participating in the *Corrective Reading* program (Level B), to what degree do the students' scores approach the normal (interquartile) range?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In working memory (another phonological process)
- (d) In spelling (a partly phonological process)

Question 21: For disabled readers participating in the *100 Lessons* program, to what degree do the students' scores approach the normal (interquartile) range?

- (a) In phonemic awareness
- (b) In phonological recoding (word attack)
- (c) In working memory (another phonological process)
- (d) In spelling (a partly phonological process)

Question 22: What is the relationship between the phonological variables measured?

Question 23: What is the nature of the latent variables underlying the dependent measures?

Question 24: Is success in the *Corrective Reading* program predicted by any of the pretest scores?

Question 25: What are the theoretical implications that follow? For example, are there implications for the phonological representation theory; or, for the reciprocal causation model, or for the best approach to teaching older disabled readers?

CHAPTER SEVEN: METHOD

Participants

The participants were primary school students attending five State (67 students) and four Catholic (120 students) schools in the Northern and Western suburbs of Melbourne. Additionally, 29 students from various other local State and Catholic schools attended the Orana program. In this study, the student cohort was initially developed through teacher referral. That is, the students included in the study were referred by teachers as those of concern to the school because they demonstrated slow reading progress. Not all students so referred were included in the study however, only those who fell within the skill band suitable for inclusion in the *Corrective Reading* program. Each student was individually assessed with the designated program placement test to ensure the presence of the entry skills and the absence of the program outcome skills. The assessments that produced the pretest results for this study were performed on a subset of the teacher-referred sample – only those within the *Corrective Reading* program skill span.

The relevant SES figures are found in Table 2 below. More than half of the students were from areas considered disadvantaged. The low mean index (995) compared to Victoria's mean (1027) indicates that the study areas have a high proportion of low income families, and more people (on average) with little training, and in unskilled occupations. The mean index of 995 corresponds to the 25th percentile, a high degree of disadvantage.

Table 2

Socio-economic Indices

Area	n	Disadvantage	
		Index	Percentile
Braybrook	7	790	10%
Coolaroo	53	950	10%
Craigieburn	34	1074	90%
Epping	4	1051	10%
Melton	17	1054	75%
Mill Park	26	1091	90%
Moomba Park	8	958	25%
Pascoe Vale	11	1002	50%
Thomastown	48	987	25%
		Study Mean	
		= 995	

Source: 1991 Census data from Castles (1994)

The age of the students who participated in the study varied from 7.8 years to 13.4 years ($M = 9.7$ years, $SD 1.2$ years), and the program period varied from 5 to 10 months ($M = 7$ months) to complete the 60-65 lessons. There were 15 dropouts whose scores were not included. In a number of cases the students had left the school during the progress of the program - a few students from Year 6 had transferred to secondary school before the program's completion the following year; whereas others transferred from their school during the school year. In a few cases, students were absent through extended illness.

Testing Procedure

Pretesting and posttesting were performed largely by the author, and some individual testing was performed by postgraduate (Masters by coursework) students who have been trained in all aspects of test administration. Their contribution to the assessment took place under the supervision of the author. The students had received specific training in the assessment package as part of the Clinic experience within their course, and also were able to observe the author testing, and receive feedback on their own practice sessions. A test manual with all standardised instructions was provided for each tester. The Picture Naming Test (Hempenstall, 1995a), Woodcock Word Attack (Woodcock, 1987), and the WISC III Digit Span (Wechsler, 1991) were administered individually. Stop watches were used to measure the 60 s period for the Picture Naming Test, and there was a risk of reliability problems given the number of testers involved. The student testers were unaware of which group they were assessing (experimental or control). Each of the testers was familiar with timed tests from their course and clinic work, and could reasonably be relied upon to measure accurately 60 s. In all the students administered the three individual tests to 140 students, in either the pretest or posttest phase.

The group tests were administered only by the author (a qualified teacher and educational psychologist). They were the Test of Phonological Awareness (TOPA), and the Brigance Spelling subtest, and were administered to students in groups from 4-20 in size, depending on school facilities. During both tests the children were taught to cover their work in the manner recommended in the TOPA. This procedure (to preclude copying) was referred to regularly during the assessment sessions.

Teachers

The program presenters were in each case qualified teachers in either the State or Catholic system. In the case of the students taught at Orana, the teachers were experienced in

the presentation of Direct Instruction programs. In other schools, teachers received a varying amount of support and training. In some cases, training was contracted from Orana, in which Orana teachers provided a total of three days instruction to new Direct Instruction teachers, including orientation, demonstration, supervised practice and feedback.

In other cases the author provided consultation, demonstration, supervised practice and feedback, though to a lesser extent than was available to the Orana-trained teachers. Some teachers elected to forego the author-offered assistance, and were content to commence the program with much less support. This lesser option involved consultation, the provision of an author-prepared manual, and the availability of a number of video-tapes of Direct Instruction classes as potential models. In addition, several Catholic schools received ongoing support from an educational psychologist from the Catholic Education Office, who was experienced in Direct Instruction. Another school received ongoing assistance from two Masters students under the supervision of the author.

Measures

Construct: Phonemic Awareness

A wide variety of tasks have been used to measure the construct of phonemic awareness. Yopp (1988) in a factor analytic study examined 10 such test-types to determine their validity and reliability. The types included: sound-to-word matching, word-to-word matching, recognition or production of rhyme, isolation of a sound, phoneme segmentation, phoneme counting, phoneme blending, phoneme deletion, specifying deleted phoneme, phoneme reversal, and invented spellings. She found that most of these were significantly and positively correlated, supporting the view that they tap a single construct, and hence adding weight to the construct's validity. In addition, each of the tests had a significant relationship with the criterion learning test, lending predictive validity to the construct. Stahl and Murray (1994), in their factor analysis of a variety of phonemic awareness tasks, took account of varying linguistic complexity. They found that their data were best accounted for by a single factor.

The Test of Phonological Awareness (TOPA) (Torgesen & Bryant, 1994) measures the capacity to use phoneme segmentation (or at least phoneme isolation). It provides analytic tasks that require children to be aware of, and be able to identify individual phonemes presented to them orally in a word. In the Early Elementary version they are required to isolate the last sounds in 20 words, and compare them to those of other words. The authors describe the test as one of phonological sensitivity (or shallow phonemic awareness) rather

than explicit (or deep) phonemic awareness because the test does not require the manipulation or pronunciation of the phonemes in the presented words (although one presumes such a capacity would be helpful). In Yopp's (1988) factor analytic study, two factors were noted within the phonemic awareness construct, each relating to the level of demand on working memory. This test provides line drawings of the words in order to reduce the working memory requirements of the test. The test is designed to be used for early identification or screening (it can be used as a group test). In this study, it is used as a measure of change in phonemic awareness over the period of the intervention. The maximum possible score is 20 (indicating mastery), and raw scores were used in all analyses.

The test manual argues that the TOPA meets the requirements for technical adequacy according to standards set forth by the American Psychological Association (1985, cited in Torgesen & Bryant, 1994). The TOPA manual reports three measures of reliability. Content sampling revealed an internal consistency of .88 average for the Early Elementary version (as used in this study). Using time sampling for the Kindergarten version over a 6 week time frame, a corrected stability estimate was .94. The score was lower (.77) for the Early Elementary version perhaps due to the longer interval (8 weeks) between tests, but probably also because of variations in the children's phonemic awareness response to the reading instruction that was taking place at that time in Year 1. The average standard error of measurement for the Early Elementary version was 5.2.

Measures of criterion-related validity provide strong support for the TOPA. It is correlated with other phonemic awareness tasks such as sound isolation (.66), and importantly, with the Word Attack (.66), and Word Identification (.60) subtests of the Woodcock Tests of Reading Mastery-Revised (Woodcock, 1987). As the Word Attack (pseudo-word) subtest is considered the most appropriate measure of phonological recoding (Hoover & Gough, 1990; Siegel, 1993; Wood & Felton, 1994), a measure that implies advanced or deep phonemic awareness, this correlation offers good concurrent validity.

The correlation with the Word Attack subtest (a reading measure administered at a later date) also provides predictive validity support. The correlation between the beginning Year 1 TOPA scores and the end Year 1 reading scores was .52. Further such support was provided by a study by Høien, Lundberg, Stanovich, Bryant, and Bjaarlid (1995) in which initial-phoneme and final-phoneme matching tasks (the tasks assessed in the TOPA) were by far the most potent predictors of reading acquisition. Naslund (1994) indicates that this same sort of oddity task has been successfully predictive of reading performance in English and

German. Nation and Hulme (1997) assert that tests of phonemic segmentation are the most sensitive measures in predicting the occurrence of reading problems.

Construct validity is established through the test's similarity of item types to other known tests of phonemic awareness. Further, the test is sensitive to changes in student performance consequent upon the implementation of a phonemic awareness program.

A difficulty inherent in the use of the TOPA in Australia relates to the use of the end sound /r/. The tenth item of the *Ending-Sound Same* subtest involves the stimulus word "chair", and a choice of responses from "sheep - can - jar" from which students choose the one with the same last sound. Pronunciation of final "r" is fairly distinct in American English, but is much less so in Australian. Bearing in mind that students are expected to use the sounds they hear in making judgements (not spelling knowledge), the item may differentiate students on grounds other than phonemic awareness, for example, spelling ability, or the ability to problem-solve by eliminating incorrect alternatives (sheep, can). The same problem arises in the *Ending Sound-Different* subtest in which students decide which one of "four - star - ball - pear" has a different end sound.

A decision was made to accentuate the /r/ in the American manner rather than risk breaching standardisation by altering the items. It was thought that the children's experience with the American accent readily evident in television programs would make the /r/ emphasis unremarkable for them. As only the author administered this test, consistency of presentation of the accent was not a concern.

Construct: Phonological Recoding in Lexical Access

Many studies (Share, 1995) have noted the higher error rate, and slower naming speed of disabled readers confronted with continuous lists of numbers, letters, pictured objects, and colours. The difficulty is independent of semantic abilities, remaining when skilled and less skilled readers are matched on receptive vocabulary (Jorm, Share, Maclean, & Matthews, 1986). Nor does it appear that the speed and error rates are due to visual perceptual processes, but rather to greater difficulty in establishing phonological representations (Share, 1995). The theoretical link between naming tasks and reading involves the requirement of retrieving the name for a stimulus presented in visual format. In practice, it has been the speed with which the task is completed that correlates most highly with both word recognition and comprehension (Wolf, 1991).

In a study of picture naming skills by Levy and Hinchley (1990, as cited in Lemoine, Levy, & Hutchinson, 1993) of good and poor readers there was a significant, regular speed difference between the groups (consisting of children from Year Three to Year Six). Vellutino et al. (1996) reported similar findings for younger children (Prep and Year 1). Wolf (1984, as cited in Crowder & Wagner, 1992) reported a correlation of .35 between picture naming speed and word recognition one year later.

There are two recognised formats for naming speed tasks - continuous (or serial) and discrete. In the discrete reaction time format the child names items presented individually in a rapid sequence, usually on a computer (which also times the response). It is considered a measure of pure retrieval because it removes the requirement of scanning and its associated visual and motor processes.

In the continuous naming format the child has the complete array of visual stimuli available to peruse sequentially. Because of the left to right sequencing and the associated requirement of engaging in the simultaneous naming of a previous item while scanning the next, the continuous format more closely approximates the reading task (Wolf, 1991). The two types are strongly correlated (Bowers, 1995).

A continuous picture naming test was developed (Appendix A) for this thesis to provide a simple test of rapid naming - one directly relevant to reading. The skill has been assessed in a number of forms, but usually involves naming of letters, numbers, colours, pictures, and objects. This test is a variant of the Rapid Automatised

Naming test (Denckla & Rudel, 1976). The RAN requires rapid sequential naming of colours, objects, or letters, and is measured by total elapsed time to complete. It has been argued that letter naming is the naming skill most salient to reading, which is unsurprising given that it directly involves an element of the reading process, and is accepted as a strong predictor of future reading success in beginning readers.

In this thesis, however, it was not assumed that all students were firm in their letter-sound knowledge. It was considered likely that a number of students, especially those eligible for the *Corrective Reading* program: Level A, would fall into this category. The rationale for naming speed tests is to determine ease of retrieval of information in the child's lexicon. If letter-sound knowledge is not firm then results of assessment would confound size of knowledge base with speed of retrieval. In that case the test would not be purely one of naming speed.

The Picture Naming Test in this thesis uses black and white line drawings of everyday objects and events. The pictures were chosen from a range of provided in the TOPA (Kindergarten version). The test comprised 60 pictures in 3 pages, and students were allowed one minute to name quickly as many as they could. They were further instructed to "pass" if they could not remember a name, to avoid losing time on any one item. The test was administered individually. Scores were kept for number correct, number incorrect, and number omitted. Denckla and Rudel (1976) had noted that "dyslexic" children were more likely to make circumlocution errors (explaining the picture rather than naming the object), while "non-dyslexic" poor readers were more likely to provide an incorrect name for the object.

Reliability figures (Hempenstall, 1995a) were obtained by using a test-retest protocol with an interval of 2 weeks, involving a class of 28 students from a northern suburbs primary school. The composite Year 3-4 class was tested individually in the identical format to the subsequent study. The ages of students ranged from 7.07 to 10.2 years. Pearson correlation was calculated at .77.

In terms of validity, the test claims to be a measure of picture naming speed, and asks students to name pictures rapidly; hence, it is reasonable to claim face validity. The relationship between picture naming speed and subsequent reading achievement noted above also provides predictive validity support.

In this thesis, the test was used to detect changes occurring over the period of the intervention, and raw scores were used for the analyses. The maximum possible score is 60, and raw scores were used in all analyses.

Construct: Phonological Recoding in Working Memory

Working memory may not be a major limiting factor in skilled reading because most words are recognised instantly, and comprehension occurs at the time of the word's fixation (Crowder & Wagner, 1992). For unskilled and novice readers, however, shortcomings in verbal working memory are likely to be exposed in the blending task, and in retaining the meaning of a sentence during its progressive decoding (Share, 1995). Disabled readers typically struggle to retain in working memory verbal material presented orally or visually (See Wagner & Torgesen, 1987 for a review). Such short-term memory problems for verbal material has been evidenced in a variety of memory tasks including digits, letters, groups of words or sentences, and in objects and nameable pictures (Share & Stanovich, 1995). The performance of these tasks requires the capacity to store information represented in a phonological code. The deficit appears specific to phonological representation, as in visuo-spatial tasks there is no similar deficit (Share, 1995). Thus, the relationship between memory span and reading is well established correlationally, but there is little evidence to support a direct causal role from memory to reading. Hulme and Roodenrys (1995) provide data to support the idea that short term memory is merely a marker for other phonological deficits (especially, the quality of phonological representations), also readily observed in speech rate measures.

Further, short term memory impairment has been noted prior to school commencement, and hence cannot be explained as merely a consequence of slow reading progress, although interestingly, the ability may be amenable to improvement as reading skill develops (Ellis, 1990; Goldstein, 1976, cited in Share, 1995). Pre and post testing of Digit Span may detect any such effects occurring during the intervention.

The measure chosen for phonological recoding in working memory was the Digit Span subtest of the Wechsler Intelligence Scale for Children-Third Edition (WISC-111). It (or a variant) has been used in studies by Ackerman and Dyckman, 1993; Bowers, 1995; Gathercole, Willis, and Baddeley, 1991; Lehto, 1995; Snowling, Goulandris, and Defty, 1996; Stahl and Murray, 1994; and, Stodhard and Hulme, 1992. The maximum possible score is 30, and raw scores were used in all analyses

Scores for Digits Forward and Digits Backward were collected (in addition to the total score), as there is a suggestion that they may not involve identical cognitive processes. Rudel and Denckla (1974, as cited in Wechsler, 1991) noted better Digits Forward than Digits Backward scores in children with developmental disorders involving the right hemisphere. Bowey, Cain, and Ryan (1992) consider Digits Forward to be indicative of articulatory loop capacity, as the student is able to rehearse continuously the sequence until its utterance. Digits Backwards also implies the presence of articulatory loop capacity, but in addition, the ability to monitor the sequence and manipulate its elements. Hence it also involves central executive functions. Rohl and Pratt (1995) using a multiple regression analysis asserted that backward repetition made contributions to reading and spelling that were independent of simple repetition tasks. There are some similarities between this task, and what Lindamood, Bell and Lindamood (1992) termed Comparator Function. She defined the term as “... the ability to compare two phonological structures by holding their phoneme, and/or syllable segments in mind, so any variations in the number, or order of their segments can be explicitly noted or represented” (p. 357). She sees this factor as one that limits the phonemic awareness of perhaps a third of the population, and of particular relevance to self-correction in reading and spelling.

The use of both scores enables a judgement about which constituent (if any) of working memory is most affected in a group with reading difficulty. Sattler (1992) considers raw score differences of three points between Digits Forward and Digits Backward to be noteworthy.

Split-half reliability coefficients are provided in the WISC-111 manual at an average of 0.85, and an average of 1.17 for standard error of measurement. The manual provides ample evidence of concurrent and predictive validity for the Full Scale through factor analytic evidence and correlations with other measures of intellectual ability. For the Digit Span subtest, a correlation of 0.74 with the WISC-R is provided.

Construct: Decoding

The Woodcock Reading Mastery Tests-Revised (Woodcock, 1987) is a comprehensive reading assessment tool frequently used in educational settings. The Word Attack subtest requires the student to decipher either nonsense words, or words that occur very rarely in our language. A correct response precludes the possibility of having used other than a phonological recoding strategy, or reading by analogy with similar real words.

There are two forms each containing 45 items in ascending order of difficulty. Testing is discontinued following six consecutive failures. As with the other subtests in the WRMT-R, it provides continuous-year norms, although for the purposes of this thesis gain scores were of more interest. The maximum possible score is 45, and raw scores were used in all analyses

A number of studies have used standard scores for the Woodcock Word Attack Test. However, in a study by McGuinness et al. (1995) there were significant correlations between age and standard scores on this test. Since the purpose of standard scores is to partial out the effects of age, a failure to do so makes such a transformation non-beneficial. It was decided then to perform analyses on raw scores.

Split-half reliability is reported in the WRMT-R manual as being at the median .87 with a standard error of measurement of between 3.3 and 5 for the age range of interest. Olson, Forsberg, Wise, and Rack (1994) supported the WRMT-R test-retest figures, reporting good correlations in their own study involving a four month, and even a four year interval.

Content validity is established by examining the scope and sequence of the subtest items, and by noting that the items are indeed relatively impervious to other than the reading strategies stated. Concurrent validity is evidenced through a comparison of this test with other recognised measures of reading. The total reading score provides correlations ranging from 0.78 to 0.92 with 5 other recognised reading tests across the age range chosen. The Word Attack subtest compared with another recognised word attack scale in the Woodcock-Johnson Reading Scale (Woodcock, 1978, cited in Woodcock, 1987) provides correlations from 0.64 to 0.9 across the age range chosen.

This subtest has been used in a number of studies to assess phonological recoding (e.g., Alexander, Anderson, Heilman, Voeller, & Torgesen, 1991; Bowers, 1995; Bowers & Swanson, 1991; Bowey, Cain, & Ryan, 1992; Lovett, Border, De Luca, Lacerenza, Benson, & Brackstone, 1994; Stanovich & Siegel, 1994; Vellutino et al., 1994; Vellutino, et al., 1996; Wagner, Torgesen, & Rashotte, 1994; Wood & Felton, 1994). Further, pseudo-word decoding is considered by many as the optimum means of assessing phonological recoding (Goulandris & Snowling, 1995; Paulesu et al., 1996; Share, 1995; Spear-Swerling & Sternberg, 1994; Stahl & Murray, 1994; Stanovich & Siegel, 1994).

Although Olson et al. (1994) accept the WRMT-R has adequate validity and test-retest correlations (even in their own study over a four month, and four year delay), they also make suggestions for improvements. They argue that improvements should include more complex

pseudo-words, consonant clusters within syllables; fewer words that can be read by analogy; and a measure of response time. They also consider silent phonological choice tasks delivered by computer have much to offer in assessment of this area. Vellutino, Scanlon, and Tanzman (1994) concur, and add that there are too few items at any level, thus leading to an overly rapid acceleration of difficulty.

The test is used here because it measures the degree to which students transfer phonemic awareness to the reading task. It also correlates strongly with word recognition and reading comprehension (Elbro et al. 1994; Vellutino et al., 1994), and thus can arguably provide a proxy for general reading progress.

Construct: Spelling Ability

There have been a number of approaches used to assess spelling. One obvious means is to assess spelling errors in the context of written expression; however, it is too complex a task to be realistic in terms of time and scoring issues. Another approach is to require the student to recognise deliberate spelling errors in a list or story (a proofreading task). In this thesis a dictated word list approach was adopted because students are familiar with such a format, for ease of assessment in a group setting, and because it is a generally accepted format. Lindamood (1994) agrees with Moats (1994c) that “... the primary construct for investigation of spelling should be the ability to write dictated words in lists” (p. 351).

The Brigance Comprehensive Inventory of Basic Skills (Brigance, 1992) spelling sub-test is primarily a criterion-referenced instrument of this type. It is based on words used at the various grade levels in five or more of nine published spelling programs. The test was presented in all cases by the author, following the manual’s instructions. Presentation involved a scripted introduction followed by a sequence of the word, the word used in a scripted sentence, then the word again. It was presented in a group format, and students used a sheet of white card to cover their answers as they were completed. The criterion for discontinuation was less than 60% correctly spelled words at any given grade level. All students commenced at the Year 1 level, and relatively few students were successful at the Year 4 level; thus in most testing sessions four groups of 10 words were presented to all students. The maximum possible score is 40, and raw scores were used in all analyses

The test has several limitations. For example, there have been no published reliability figures. Test-retest reliability was determined (Hempenstall, 1995c) in a class of 28 students in a northern suburbs primary school. The composite Year 3-4 class was tested in a group format, using blank sheets of paper to cover their work in order to preclude collaboration. The

ages of students ranged from 7.07 to 10.2 years. Pearson correlation was calculated at .97 (Statistical Package for the Social Sciences, Version 6.1, 1995).

The Brigance test does not have parallel forms, and hence the same form was presented at pre- and posttest. Given the time interval was in excess of four months in most cases, practice effects should not have had a major impact. In addition, practice effects should have occurred equally in experimental and comparison groups and are therefore controllable.

A potential disadvantage involves the high proportion of irregular words in the test. One may expect that any improvement would be as a result of improved capacity to relate sounds to letters and letter groups. This effect should be most evident in regular words, though one might anticipate errors on irregular words to more closely approximate the word's pronunciation. As scoring is dichotomous - either correct or incorrect, such change is not measured by the test. Foorman and Francis (1994) noted such an outcome in their study of Year One beginning readers exposed to letter sound instruction. Ehri (1993) also points to the value of assessing spelling growth in a more fine grained manner, as when a child improves his misspelling of "pickle" from *po* to *pikl*. Moats (1994c) describes such a spelling assessment system in which quality points (1-5) are assigned for degrees of spelling errors based on a specified set of criteria. The effect of a correct/incorrect dichotomy as used in this thesis would be to attenuate measured change by failing to note within-incorrect-category improvement, that is, underestimating spelling growth.

The Corrective Reading Program

The *Corrective Reading* program is a remedial reading program designed for students in Year 3 and above. It comprises two strands: Decoding and Comprehension, and within these strands are a number of levels. The Decoding strand was the focus of this thesis, having 4 levels (A, B1, B2, C) corresponding to the students' decoding capacity assessed with a placement test.

The program was chosen as the intervention program for this thesis because of the author's experience with it, and its record of success in improving the reading outcomes for children at-risk. This was noted in the empirical studies available in the research literature, and also in evaluations performed over a number of years by the author.

The *Corrective Reading* program has been evaluated on many occasions (both the 1978 and 1988 editions), though its effects on phonological processes have not previously been a focus. Most analyses have emphasised word recognition and reading comprehension as outcome variables, and results for a wide range of poor readers have been strong. Studies

have noted positive outcomes for learning disabled students (Holdsworth, 1984; Lloyd, Epstein, & Cullinan, 1981; Maggs & Murdoch, 1979), intellectually disabled students (Polloway & Epstein, 1986; Polloway, Epstein, Polloway, Patton, & Bell, 1986), maladjusted boys (Thorne, 1978), secondary students (Campbell, 1983; Gregory et al., 1982a; Gregory et al., 1982b; Sommers, 1995), adults (Herr, 1989), and gifted students (Noon & Maggs, 1980).

Selection

The placement test is administered prior to the program and consists of several passages of prose, the rate and accuracy of reading determining the program level for any given student.

The test is designed to assess ability at the word level. The story text is not amenable to contextual strategies, and the assessment criteria of rate and accuracy make it difficult for other than skilled decoders to pass unnoticed. In the author's experience it is capable of making the discrimination necessary to place students in any of the 4 levels (A, B1, B2, C), or to detect those whose skills are above or below the entry criteria. Used informally as a posttest measure it frequently has demonstrated that the student would now be correctly placed at the next higher level. This implies that the assessment device is closely related to the specified program objectives.

The placement test also ensures that student groups are relatively homogeneous in their decoding ability, and that they are neither over-challenged by the level of difficulty of the program, nor already competent at that level. The test is administered individually and takes about five to ten minutes. Detailed instructions are provided for administration and scoring.

In the school settings the reading group teacher performed the screening after discussion with the author on the details of administration and scoring. Typically the screening sample was derived from class teacher reports of students in the middle or upper primary school whose reading progress had been of concern. This teacher-identified group was then assessed with the placement test.

The possible outcomes of such assessments are:

the child's current decoding skill levels are below those of the lowest level of the program (Level A), and would be best addressed with a beginning reading program.

the child is appropriate for placement in one of the four program levels, or

the child has already mastered the decoding skills taught at each level, and any reading deficits are probably not in the area of decoding.

Depending on the range of Year levels included in the assessment cohort, it is possible that, meeting all the students' needs would require the provision of several of the levels, most frequently Levels A and B1. Schools then decide which group or groups they are able to supply with a program. In some cases schools decide to provide one program as a pilot, and plan subsequent programs after evaluating the first. This is a reasonable decision, but means that some of the identified students will not receive (immediate) assistance.

This decision usually causes some discomfort, and it is tempting to alter remedial direction and simply supply a little (usually ineffectual) aid to all of the identified students rather than select only a subset for the intensive program.

As all of the students who fall within the Program's range are equally in need of support, the basis for selecting one group must be on grounds other than differential need. Some schools decide to provide the Level B1 program initially, because the majority of such students are in Years Five and Six. Schools that make this choice place a high value on ensuring students to not leave primary school without their receiving some measure of remedial reading assistance.

Other schools choose to offer Level A, as the majority of the eligible students arrive from Year Three and Four. These schools consider such students able to make better progress (being younger), and also will be enrolled at the school long enough to participate in further levels subsequently, if that is deemed necessary. Obviously each of these options is a compromise as it involves excluding some students in need.

In some cases this exclusion is permanent as the senior group leaves the school at the end of that year. In other schools the identified-but-not-treated group will receive assistance in the next round of programs offered by the school. With most schools this latter sequence ensued. All schools were enthusiastic about extending their program involvement supported by objective and subjective evaluation of their pilot. On only one occasion was the program discontinued (albeit for one year), when school resources were inadequate to continue to provide the staff required.

The wait list group provided the source of the non equivalent control group students for this thesis. It is important for the internal validity of the thesis to note that the basis for selection in either the experimental or comparison group was not on the basis of greater need, but rather school values. All of the students identified were in similar need, and at each program level displayed a similar degree of reading deficit.

The students on the waiting list were told that they would be included for assistance later that year, and each was pretested and post tested in the same manner as the experimental group. In 12 cases, students in the wait list group were later included in the program, and were re-assessed after their program involvement; hence, they appear in the thesis as members of each group.

Program Design

There are two major features evident in the *Corrective Reading* program. They are the emphasis on decoding skills (phonics) and the Direct Instruction approach to teaching the phonics content. It includes work on both isolated words and connected sentences, but its major emphasis is at the level of word structure. It is made clear to students that the decoding of novel words involves careful word analysis rather than partial cue or contextual guessing. Students are continually prompted to take account of all letters in a word, and become sensitised to common (and often problematic) letter groupings, for example, those beginning with combinations *st, bl, sl, fl, pl, sw, cl, tr, dr*; or ending with *nt, nd, st, ts, mp, ps, cks, ls, ms, th, er, ing, ers, y*. The sentences provided are constructed in a manner that allows few clues for contextual guessing, but provides ample opportunities to practise what has been learned in the teacher-presented word-attack segment of the lesson.

Lessons are designed to be provided in groups of up to 15 students. In this thesis, most groups comprised about 10 students. The rationale for this reduction involved the lack experience of the teachers with the program, and the observation that in most groups of poor readers there are usually several students difficult to motivate, and maintain on task.

This first hurdle is difficult for those teachers used to a less directive model of teaching. Lessons are scripted, and most teachers report requiring at least 20 lessons before reasonable comfort with the approach is achieved. Teacher support is valuable in the early stages to assist in this skill development, and to preclude teacher initiated changes that may jeopardise program success. The level of support needed varies from teacher to teacher; however, it was not possible in these studies to provide the extensive teacher training model described by the program designers (Engelmann, Becker, Carnine, & Gersten, 1988). Woodward (1993) indicated that Follow Through teachers took at least 12 months to master the teaching skills involved in Direct Instruction programs.

The program designers claim that the model combines the benefits of 1:1 tutoring with the effectiveness of group instruction. This is achieved by the use of choral responses prompted by various signals (a new skill for most teachers). Not only must teachers follow a

script, but they must be able to signal reliably to students when to respond, and then pay attention to each student's response in order to monitor skill development and teaching effectiveness. The results of this monitoring process help determine lesson pacing by controlling the amount of repetition necessary for mastery. The larger the group, the more difficult it is to monitor continuously every student's progress - thus smaller group sizes are helpful for novice program presenters. As teachers' reliance on the script diminishes, and as their signalling improves, so their adroitness at student monitoring improves and they are better able to manage larger groups.

The issues of behaviour management are usually more demanding in secondary than primary schools, but may still present difficulties in middle and upper primary schools. Participation in the reading program involved parent, but not student consent; that is, students were not volunteers. Most schools considered the needs of the students too important to allow students the right of veto. To help motivate students whose history has made reading a non-preferred activity, the program includes a points system for each lesson segment. Most schools perceived the advantage of this system and incorporated it successfully into their plan. The potential for program disruption by a few disillusioned students was an additional reason for beginning with smaller group sizes.

Lessons typically range from 45 minutes to one hour, dependent on teacher lesson pacing. Typically, pacing improves with experience, but initially some teachers find it difficult to complete a whole lesson in the time allotted.

Program design specifies an optimum schedule of five lessons each week. This level of intensity has been found important for students with reading problems, as they tend to have difficulty retaining new skills and knowledge. For this reason, there is strong emphasis on massed practice for mastery, and spaced practice for retention. If the lesson frequency falls too low, retention may be jeopardised leading to a general progress deceleration. However, not all schools are able to timetable five lessons per week, and even those that do so find competing events sometimes force class cancellation.

The Level A program focuses attention on word structure through reviewing letter sound correspondence, and regular rhyming, blending and segmenting activities. It relates these phonemic awareness activities to the written word by initially emphasising regularly spelled words decomposable by using these skills. When this phonic approach is accepted by students as a viable (even valuable) strategy, common irregular words are introduced. In the authors' view this sequence is important to prevent the jettisoning of the generative decoding

strategies because of their apparent inconsistent results if irregulars are initially encountered at the high rate common in authentic literature.

Engelmann, Hanner, and Johnson (1988) describe the range of skills taught in Decoding A:

Letter/sound identification; sounding-out (segmenting) orally presented words, and then *saying them fast* (blending); decoding words of varying degrees of irregularity; reading whole words *the fast way*; reading short groups of words; sentence reading; spelling. Related skills such as matching letters, and common letter groupings (such as *ing*), word completion (for example, rhyming), and symbol scanning are included on the student worksheets.

The main objective in Decoding A is to impress upon students that there are regularly spelled words, words that are pronounced by blending the sounds of the letters in them. When students understand that the word identification relates to its spelling, irregularly spelled words, such as *said* and *what*, are introduced. It is explained that such words are spelled in one way but pronounced in a different way.

The sentence-reading exercises provide practice in reading words within a context. Most Level A students are not familiar with the concept and practice of decoding, and their problem is magnified when they try to read sentences. Usually, their sentence-reading strategy involves guessing, based on the syntax, the first letter or two, or the position of words within the sentence (e.g., they guess that the first word in the sentence is *the*). The objective of the sentence-reading activities is to retrain students in how to read words in sentences; achieved partly through ensuring contextual strategies will be unproductive, and through immediate correction of all decoding errors.

The next level of the *Corrective Reading* program builds on the curriculum presented in Level A. The typical Decoding B lesson is divided into four major parts. Word-attack skills take up about 10 minutes of the period. Students practise pronouncing words, identifying the sounds of letters or letter combinations, and reading isolated words composed of sounds and sound combinations that have been learned by the students.

Group story-reading follows immediately after word-attack skills. This part of the lesson takes approximately 15 to 20 minutes. Students take turns reading aloud from their storybook, while those who are not reading follow along. The stories are divided into parts, and when the group reads a story part within the error limit, the teacher presents specified comprehension questions for that part.

Individual reading checkouts follow, and take about 10 minutes. Assigned pairs of students read two passages, the first of which is from the lesson just read by the group; whereas, the second is from preceding lesson. Each member of the pair first reads the passage from the current story, then the passage from the preceding lesson. Points for the first passage are earned if the student reads the passage within a specified rate and error criterion. (For instance, the student must read 85 words in one minute, with no more than two errors).

Workbook activities are presented as the last 10 minutes of the lesson. These activities are integrated with the activities in the other sections to provide additional practice opportunities.

Data Analysis

Raw data were analysed using SPSS 6.1 (1995) procedures. Statistical significance testing was performed at two levels. A single factor between subjects multivariate analysis of covariance (MANCOVA) was performed with combined pretest scores serving as covariates, and combined posttest scores as dependent variables. Data were also analysed using two-way mixed multivariate analysis of variance (MANOVA), with follow-up testing involving simple main effects. These procedures were associated with Research Questions 1-4.

The second level of statistical significance testing involved a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as covariates, and posttest scores as dependent variables. Data were also analysed using two-way mixed analysis of variance, with follow-up testing involving simple main effects. These procedures were associated with Research Questions 1-4.

Assumptions of normality and homogeneity of variance were tested for all data used in ANOVA, ANCOVA, MANOVA, and MANCOVA analyses. Results of the Lilliefors test for normality, and Levene's test for homogeneity of variance are summarised in Tables 5, 18, 19, 36. When assumptions were violated appropriate transformations were assessed, and the transformed variables retained if subsequently assumptions were met. When no transformation assisted, the following convention recommended by Tabachnick and Fidell (1996) was adopted. Their response to such irretrievable violations of homogeneity of variance is to suggest a more stringent requirement for significance, and $p = .01$ was selected. This precaution may be unnecessary as Tabachnick and Fidell further argue that most tests of homogeneity of variance are unreasonably strict, and that if group sizes are reasonably matched (within a ratio of 4:1 largest to smallest cell size), a condition met in this thesis, then multivariate procedures are fairly robust in the event of such violations.

Effect size d (Cohen, 1988) was calculated for each dependent variable to provide information on the magnitude of the observed changes. The calculation of effect size was based upon the ratio of the difference between the group means at pre and posttest (separately for experimental and control groups) and the pooled standard deviation of that group at pre and posttest. All effect sizes were calculated using the Hunter-Schmidt error correction procedure (Hunter & Schmidt, 1990), which involves dividing the calculated effect size by the square root of the test reliability. These procedures were associated with Research Questions 5-11, 14, 15, 17.

Analysis of Research Question 12 involved a visual examination of Table 11.

Analysis of Research Question 13 involved visual examination of Table 12.

Analysis of Research Question 16 involved visual examination of Table 15.

Analysis of Research Questions 18-21 was performed through visual inspection of Figures 16-30, and 36-40.

To examine the relationship between variables at pretest and posttest, correlations, hierarchical regressions, and principal component analyses were also performed. These procedures were associated with Research Questions 22-24.

CHAPTER EIGHT: RESULTS

Objectives

This research was designed to assess the effect of participating in the *Corrective Reading* program on phonological processes (i.e., phonemic awareness, phonological recoding in lexical access, and phonological recoding in working memory), word attack, and spelling. The relationship between these variables was also of interest. To this end a number of statistical procedures were performed on the data derived from pretests and posttests of the relevant variables. In all, the data analysis involves correlation matrices, multivariate and univariate analyses of variance and covariance, multivariate, hierarchical, and simple linear regressions, principal component analyses, and effect size calculations.

This first section was designed to answer the research questions concerning the outcomes of the program: Did participation in the *Corrective Reading* program increase phonemic awareness, phonological recoding (word attack) skills, and other phonological processes (i.e., naming, working memory)? Did the *Corrective Reading* program effects generalise to spelling?

Descriptive Statistics

Tables 3 and 4 provide the raw and transformed data used for all Level AB analyses, whilst Table 5 is a correlation matrix incorporating the correlations between all variables at pretest and posttest for the combined AB experimental and control groups.

Table 3

Experimental vs Control Group: Mean Raw Scores

	<i>n</i>	Age	TOPA	Word Attack	PNT	Digit Span	Spelling
Control	72						
Pretest							
Range		83-153	5-20	0-29	6-54	4-17	5-39
<i>M</i>		115.72	14.57	11.38	33.43	9.28	17.85
<i>SD</i>		14.77	4.49	6.77	9.03	2.22	6.46
Posttest							
Range		90-160	6-20	0-28	14-52	5-15	4-34
<i>M</i>		122.61	15.28	12.31	34.57	9.57	19.53
<i>SD</i>		14.76	4.36	6.96	8.57	1.84	6.33
Experimental	134						
Pretest							
Range		92-161	1-20	0-31	12-53	3-15	1-31
<i>M</i>		115.27	12.84	10.16	33.03	9.34	15.47
<i>SD</i>		13.26	3.90	6.45	7.64	2.08	5.58
Posttest							
Range		97-167	6-20	0-41	9-55	2-18	7-41
<i>M</i>		122.24	17.04	19.54	37.39	10.31	20.99
<i>SD</i>		12.83	3.47	8.18	8.01	2.28	5.75

Table 4
Experimental vs Control Group: Mean Power Transformed Scores

	<i>n</i>	Word Attack	Digit Span	Spelling
Control	72			
Pretest				
Minimum		1.00	0.06	2.72
Maximum		2.49	0.26	9.76
<i>M</i>		1.90	0.12	5.91
<i>SD</i>		0.31	0.03	1.37
Posttest				
Minimum		1.00	0.07	2.37
Maximum		2.47	0.21	8.97
<i>M</i>		1.94	0.12	6.26
<i>SD</i>		0.30	0.02	1.33
Experimental	134			
Pretest				
Minimum		1.00	0.07	1.00
Maximum		2.53	0.35	8.47
<i>M</i>		1.83	0.12	5.40
<i>SD</i>		0.34	0.03	1.28
Posttest				
Minimum		1.00	0.06	3.35
Maximum		2.72	0.18	10.07
<i>M</i>		2.20	0.11	6.58
<i>SD</i>		0.30	0.02	1.12

Table 5
Correlations between Pretest and Posttest Scores

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. TOPA Pretest scores										
2. TOPA Posttest scores	0.59									
	<i>P</i> <.001									
3. Word Attack Pretest scores	0.45	0.29								
	<i>P</i> <.001	<i>P</i> <.001								
4. Word Attack Posttest scores	0.27	0.42	0.61							
	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001							
5. PNT Pretest scores	0.28	0.16	0.10	0.00						
	<i>P</i> <.001	<i>P</i> =.024	<i>P</i> =.174	<i>P</i> =.976						
6. PNT Posttest scores	0.28	0.23	0.22	0.22	0.70					
	<i>P</i> <.001	<i>P</i> =.001	<i>P</i> =.002	<i>P</i> =.001	<i>P</i> <.001					
7. Digit Span Pretest scores	0.34	0.34	0.23	0.11	0.28	0.20				
	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> =.102	<i>P</i> <.001	<i>P</i> =.004				
8. Digit Span Posttest scores	0.28	0.32	0.24	0.23	0.17	0.27	0.59			
	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> =.001	<i>P</i> =.001	<i>P</i> =.013	<i>P</i> <.001	<i>P</i> <.001			
9. Spelling Pretest scores	0.39	0.19	0.53	0.20	0.15	0.20	0.22	0.16		
	<i>P</i> <.001	<i>P</i> =.005	<i>P</i> <.001	<i>P</i> =.004	<i>P</i> =.037	<i>P</i> =.005	<i>P</i> =.002	<i>P</i> =.022		
10. Spelling Posttest scores	0.36	0.30	0.53	0.45	0.10	0.26	0.28	0.30	0.71	
	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> =.158	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001	<i>P</i> <.001	

Note. Coefficient / 2-tailed Significance

Reading Disability Criterion

It was argued in Chapter 1 that the major deficit facing the disabled reader is a difficulty in decoding single words, and that the primary basis for this difficulty is

phonological in nature. It has also been argued in Chapter 7 that a pseudoword decoding test is an appropriate tool for discerning such a difficulty. An analysis of the pretest scores of the combined experimental and control groups reveals that the average score on Word Attack meets each of the criteria that various studies (Felton, 1992; Lovett & Steinbach, 1997; Lovett et al, 1994; Lyon & Moats, 1997; Newby, Recht, & Caldwell, 1993; Prior, Sanson, & Oberklaid, 1995; Stanovich & Siegel, 1994; Vellutino et al, 1996) have adopted as defining reading disability. The average score of the cohort in this study can be converted into the frameworks described above as: the 5th percentile, a standard score of 75, and an average delay of 2.8 years as assessed on the Woodcock Tests of Reading Mastery (1987). A graphical representation of these figures may be found in *Figure 17*. Mean scores for Word Attack (A & B combined).

Multivariate Analyses

A single-factor between-subjects multivariate analysis of covariance (MANCOVA) was performed to indicate whether there was any difference between the experimental and control groups on the combined posttest scores for the five main dependent measures. The five corresponding pretest scores served as covariates. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(25, 707.32) = 2.33, p < .001$, so subsequent analysis required fitting separate slopes for each level of the treatment group factor. This analysis revealed that there was a significant multivariate relationship between the combined pretest scores and the combined posttest scores for both the control group, Wilks' $\eta^2 = .16, F(25, 707.32) = 18.08, p < .001$, and the experimental group, Wilks' $\eta^2 = .19, F(25, 707.32) = 16.13, p < .001$. With the pretest results partialled out separately for the two groups, there was a significant overall difference between the treatment and control groups, Wilks' $\eta^2 = .89, F(5, 190) = 4.75, p < .001$.

Results for the combined variables were also analysed using a two-way mixed multivariate analysis of variance (MANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, Wilks' $\eta^2 = .94, F(5, 200) = 2.59, p = .027, \text{power} = 0.79$, and for time, Wilks' $\eta^2 = .40, F(5, 200) = 60.55, p < .001, \text{power} = 1.00$, and for the group-by-time interaction, Wilks' $\eta^2 = .60, F(5, 200) = 26.85, p < .001, \text{power} = 1.00$. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, Wilks' $\eta^2 = .94, F(5, 200) = 2.61, p = .026$, multivariate effect size = .06, power = .80 and at posttest, Wilks' $\eta^2 = .84, F(5, 200) = 7.54, p < .001$, multivariate effect size = .16, power = 1.00. Further, a significant pre- to posttest difference was found for the control group, Wilks' $\eta^2 = .72, F(5, 67) = 5.22, p < .001$, multivariate effect size = .28, power = .98, and for the experimental group, Wilks' $\eta^2 = .22, F(5, 129) = 93.78, p < .001$, multivariate effect size = .78, power = 1.00, and the magnitude of

effect was substantially larger for the experimental group. The multivariate effect size ($1 - \eta^2$) can be considered large when it exceeds 0.15 (Cohen, 1988).

Univariate Analyses

This series of outcomes involved univariate analyses of the pretest and posttest data, and also included the effect size d . Under the Cohen (1988) convention, 0.2 constitutes a small effect size, 0.5 a medium effect size, and 0.8 a large effect size. Slavin (1990) argued that an effect size above 0.25 should be considered educationally significant. The rationale for the decision to adopt both ANCOVA and ANOVA procedures is discussed in Chapter 10, under *Choice of Analyses*.

Assumptions of normality and homogeneity of variance were tested for all data used in ANCOVA and ANOVA analyses, and data transformations were performed when necessary, as shown in Table 6. Power transformations were used for all experimental and control group, pretest and posttest data for Word Attack, Digit Span, and Spelling. Transformations were unnecessary for Picture Naming, and were unhelpful for TOPA scores (for which the more stringent requirement for significance $\alpha = .01$ was adopted).

Table 6

Tests of Normality and Homogeneity of Variance: Levels A and B Combined

Variable	Test	Group	Trans formation	Lilliefors' test of Normality	Levene's test of Homogeneity of Variance
Word Attack	Pretest	Control	Power	> .2	
		Treatment	Power	.03	.53
	Posttest	Control	Power	> .2	
		Treatment	Power	.08	.52
Digit Span	Pretest	Control	Power	.00	
		Treatment	Power	.00	.87
	Posttest	Control	Power	.00	
		Treatment	Power	.00	.98
Spelling	Pretest	Control	Power	> .2	
		Treatment	Power	> .2	.53
	Posttest	Control	Power	> .2	
		Treatment	Power	.00	.20
PNT	Pretest	Control	No	> .2	
		Treatment	No	> .2	.08
	Posttest	Control	No	> .2	
		Treatment	No	> .2	.45
TOPA	Pretest	Control	No	.01	
		Treatment	No	.05	.02
	Posttest	Control	No	.00	
		Treatment	No	.00	.00

Initial analyses were performed on the total sample (206 students). The results for the students in Level A and Level B were combined. The overall finding was that educationally significant change occurred in each of the measured variables, the size of the program effect varying from medium in the case of Digit Span, and Picture Naming, to large in Word Attack, TOPA, and Spelling.

Test Of Phonological Awareness (TOPA)

Results for TOPA were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202) = 14.15, p < .001$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pretest scores covaried significantly with posttest scores for both the control, $F(1, 202) = 127.84, p < .001$, and experimental groups, $F(1, 202) = 57.69, p < .001$. With the pretest results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202) = 31.73, p < .001$.

Results for TOPA were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204) =$

0.00, $p = .98$, but a significant main effect was found for time, $F(1, 204) = 172.29$, $p < .001$, power = 1.00, and the group-by-time interaction, $F(1, 204) = 53.75$, $p < .001$, power = 1.00, which is illustrated in Figure 1. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, $F(1, 204) = 8.23$, $p = .005$, $d = -0.48$, and at posttest, $F(1, 204) = 10.04$, $p = .002$, power = 1.00, $d = 0.53$. Further, no significant pre- to posttest difference was found for the control, $F(1, 204) = 3.41$, $p = .066$, $d = 0.18$, power = 0.451, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 204) = 222.63$, $p < .001$, $d = 1.29$, power = 1.00, and the magnitude of effect was large for the experimental group.

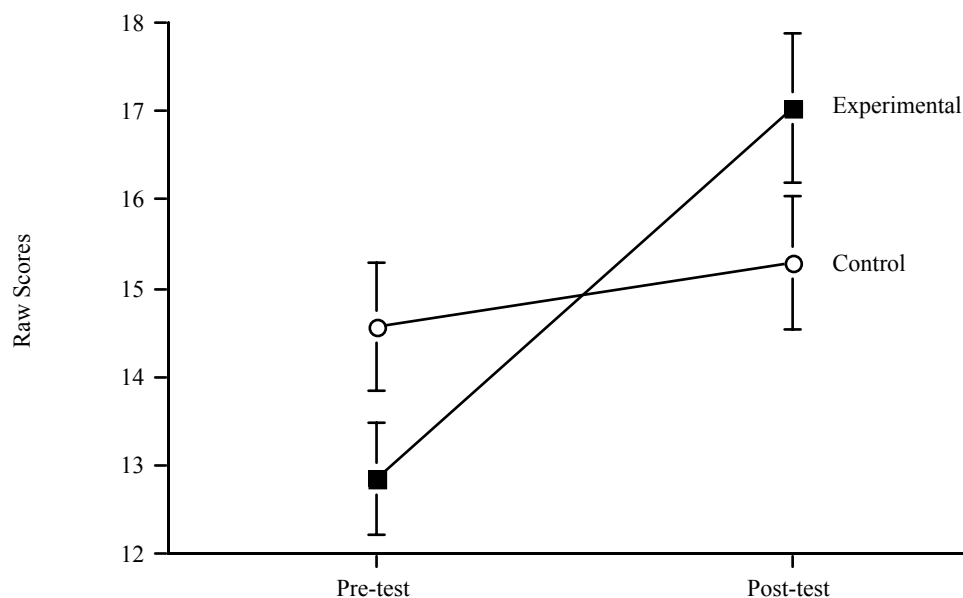


Figure 1. Interaction (+SE) between experimental and control groups at pre- and posttest for TOPA (A and B combined).

Word Attack

Results for Word Attack were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202) = 11.28$, $p = .001$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pretest scores covaried significantly with posttest scores for both the control, $F(1, 202) = 101.96$, $p < .001$, and experimental groups, $F(1, 202) = 85.88$, $p < .001$. With the

pretest results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202) = 23.55, p < .001$.

Results for the power transformed scores for Word Attack were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, $F(1, 204) = 4.79, p = .030$, power = 0.58, and for time, $F(1, 204) = 196.06, p < .001$, power = 1.00, and the group-by-time interaction, $F(1, 204) = 73.49, p < .001$, power = 1.00, which is illustrated in Figure 2. Follow-up testing of the interaction using simple main effects found a non significant difference between the experimental and control groups at pretest, $F(1, 204) = 2.01, p = .158, d = -0.20$, power = .29, but a significant difference at posttest, $F(1, 204) = 33.03, p < .001$, power = 1.00, $d = 1.00$. Further, no significant pre- to posttest difference was found for the control, $F(1, 204) = 1.86, p = .174$, power = .27, $d = 0.15$, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 204) = 267.69, p < .001$, power = 1.00, $d = 1.34$., and the magnitude of effect was large for the experimental group.

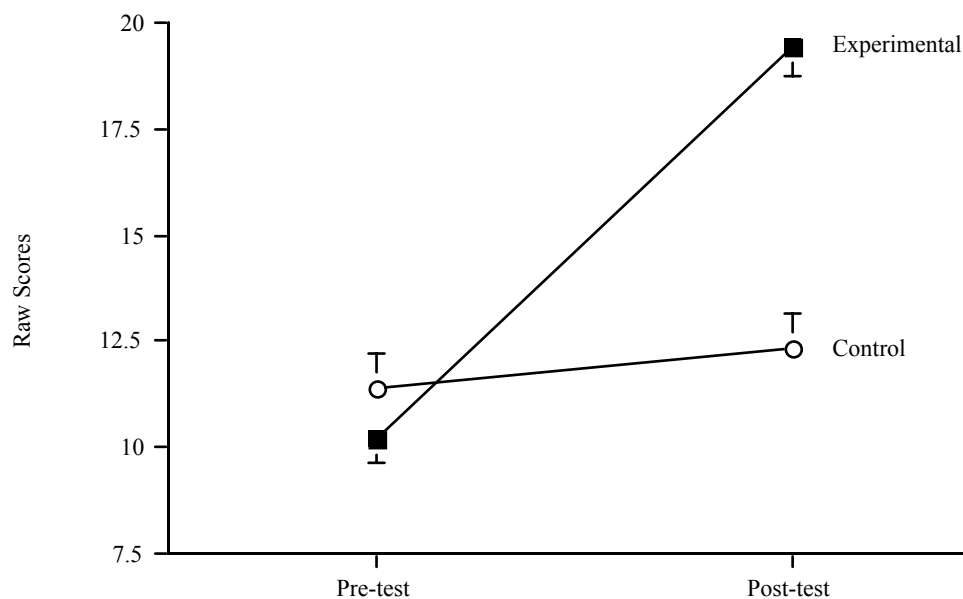


Figure 2. Interaction (+ SE) between experimental and control group at pre- and posttest for Word Attack (A and B combined).

Picture Naming Test

Results for Picture Naming Test were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed no violation of the assumption of

homogeneity of slopes, $F(1, 202) = 2.27, p = .134$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups $F(1, 203) = 10.48, p = .001$.

Results for Picture Naming Test were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204) = 0.92, p = .337$, power = 0.17, but a significant main effect was found for time, $F(1, 204) = 47.49, p < .001$, power = 1.00, and the group-by-time interaction, $F(1, 204) = 10.11, p = .002$, power = .88, which is illustrated in Figure 3. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 204) = 0.11, p = .737$, power = 1.00, $d = -0.06$, but a significant difference at posttest, $F(1, 204) = 4.22, p = .041$, power = .53, $d = 0.39$. Further, no significant pre- to posttest difference was found for the control group, $F(1, 204) = 2.28, p = .133$, power = .32, $d = 0.15$, but a significant pre- to posttest difference was found for the experimental group, $F(1, 204) = 55.31, p < .001$, power = 1.00, $d = 0.57$, and the magnitude of effect was medium for the experimental group.

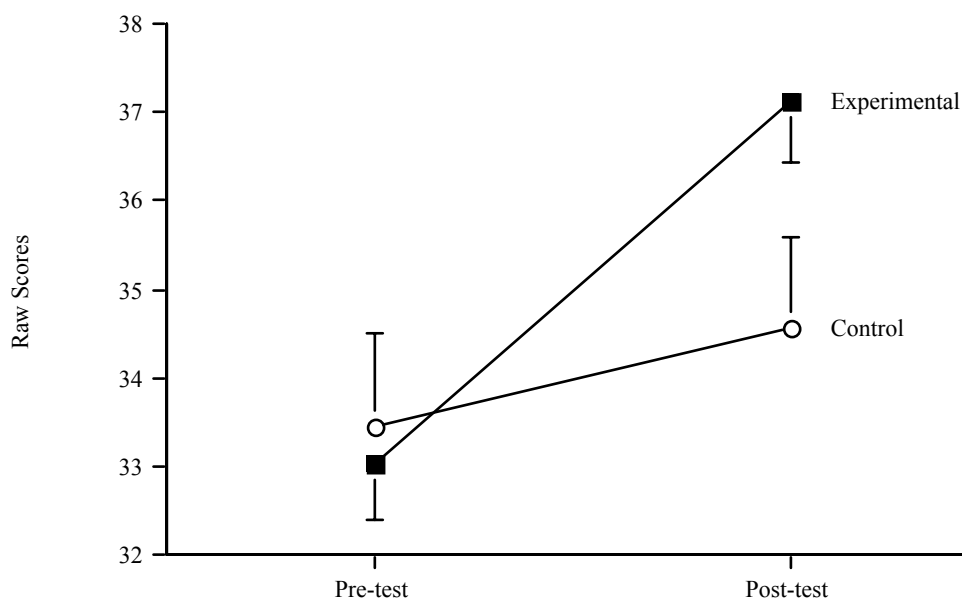


Figure 3. Interaction (+ SE) between experimental and control group at pre- and posttest for Picture Naming Test(A and B combined) .

Digit Span

Results for Digit Span were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 202) = 0.25, p = .621$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 203) = 7.92, p = .005$.

Results for power transformed scores for Digit Span were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204) = 1.5, p = .222, \text{power} = .23$, but a significant main effect was found for time, $F(1, 204) = 28.71, p < .001, \text{power} = 1.00$, and not for the group-by-time interaction, $F(1, 204) = 3.68, p = .056, \text{power} = .48$, which is illustrated in Figure 4. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 204) = 0.00, p = .947, \text{power} = .03, d = 0.03$, but found a significant difference at posttest, $F(1, 204) = 6.08, p = .015, \text{power} = 0.69, d = 0.38$. Further, no significant pre- to posttest difference was found for the control, $F(1, 204) = 2.62, p = .107, \text{power} = .36, d = 0.16$, but a significant difference was found for the experimental group, $F(1, 204) = 29.77, p < .001, \text{power} = 1.00, d = 0.48$, with a medium effect size for the experimental group.

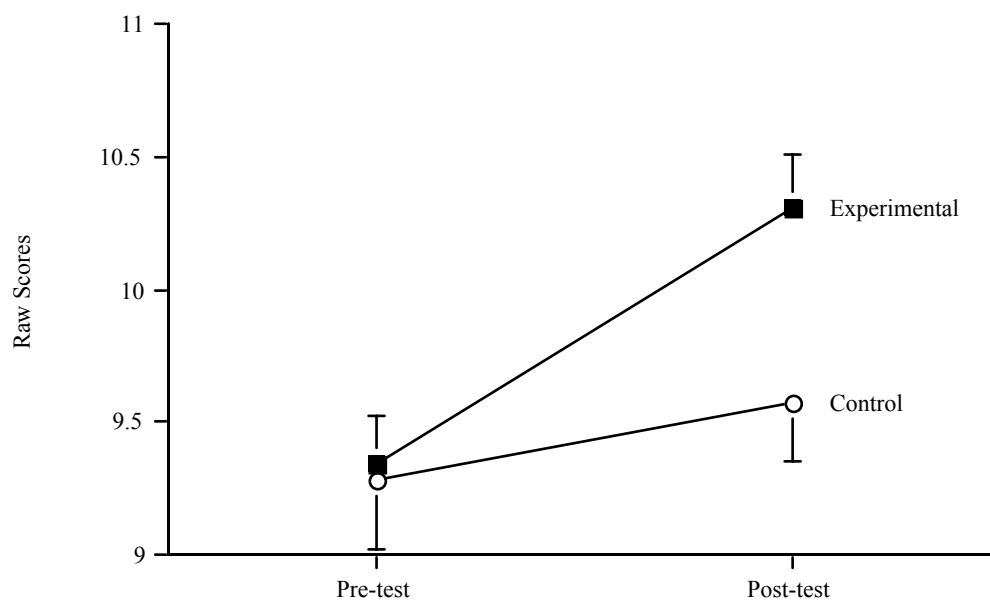


Figure 4. Interaction (+ SE) between experimental and control group at pre and posttest for Digit Span (A and B combined).

Brigance Spelling

Results for Brigance Spelling were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 202) = 5.37, p = .021$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pretest scores covaried significantly with posttest scores for both the control, $F(1, 202) = 126.58, p < .001$, and experimental groups, $F(1, 202) = 112.42, p < .001$. With the pretest results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202) = 12.26, p = .001$.

Results for the power transformed Spelling scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 204) = 0.30, p = .58, \text{power} = .038$, but a significant main effect was found for time, $F(1, 204) = 188.89, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 204) = 36.89, p < .001, \text{power} = 1.00$, which is illustrated in Figure 5. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, $F(1, 204) = 7.03, p = .009, \text{power} = .75, d = -0.42$, but not at posttest, $F(1, 204) = 3.32, p = .07, \text{power} = .44, d = 0.25$. Further, significant pre- to posttest differences were found for both the control, $F(1, 204) = 10.41, p = .001, \text{power} = .89, d = 0.27$, and experimental groups, $F(1, 204) = 215.38, p < .001, \text{power} = 1.00, d = 0.99$, however the magnitude of effect was large for the experimental group.

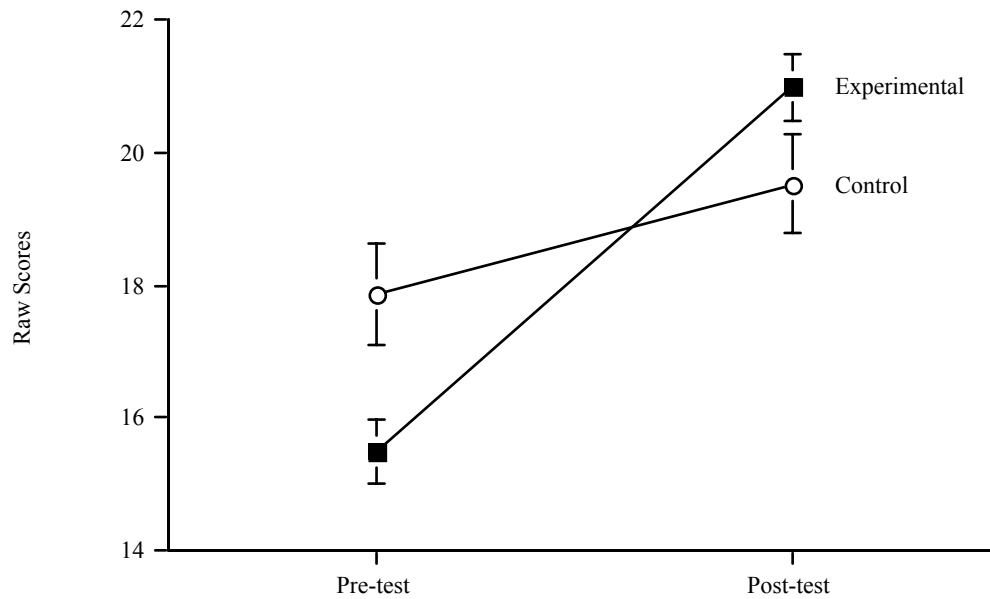


Figure 5. Interaction (+ SE) between experimental and control group at pre and posttest for Brigance Spelling (A and B combined)

The initial questions were: Did participation in the *Corrective Reading* program increase phonemic awareness, phonological recoding (word attack) skills, and other phonological processes (naming, working memory)? Did the *Corrective Reading* program effects generalise to spelling? The results presented for the combined A and B groups in the above sets of analyses indicated a clear pattern of educationally significant and educationally significant increases represented in the posttest scores for the experimental group. The effects varied from large (TOPA, Word Attack, Spelling) to moderate (Digit Span and Picture Naming).

How Widespread are the Effects?

The effects of the program on each of the processes assessed have been strong. It was also of interest to consider whether there were differential effects across age, sex, or school attended. In Table 7, it can be seen that only sex was a significant predictor of Word Attack, and this picture is enhanced by considering the effect sizes in Table 8.

Table 7
Summary of Linear Regression Analysis for Age, School, and Sex Predicting Word Attack Gains Scores (n = 134)

Variable	B	β	p
Age	-0.064	-0.131	0.123
School	0.354	0.129	0.128
Sex	-3.078	-0.206	0.016
(Constant)	18.396		0.001

Effect Size Calculation by Sex

As is evident from Table 8 the major findings regarding program effect sizes were similar though not identical for boys and girls. The effects were very large for boys' word attack compared to large for girls; for spelling the effect size for boys were large whilst for girls it fell into the medium range. Girls demonstrated greater improvement in naming, whereas TOPA and Digit Span results were quite similar to those of the boys.

Table 8

Effect Size (d) Calculation by Sex

	n	TOPA	Word Attack	PNT	Digit Span	Spelling
Boys						
Experimental	100	1.26	1.57	0.57	0.51	1.08
Control	50	0.17	0.19	0.14	0.17	0.22
Girls						
Experimental	34	1.40	0.91	0.87	0.62	0.71
Control	22	0.21	0.03	0.19	0.08	0.44

Effect Size Calculation by Age

Another interesting question involved possible age differences in the program's effects. It was apparent however that the beneficial outcomes were quite consistent across different ages.

Table 9

Effect Size (d) Calculation by Age

	<i>n</i>	TOPA	Word Attack	PNT	Digit Span	Spelling
8-						
Experimental	43	1.24	1.25	0.69	0.66	1.03
Control	19	0.06	0.10	0.21	-0.19	0.11
9						
Experimental	44	1.14	1.33	0.75	0.44	0.86
Control	25	0.27	0.08	0.24	0.40	0.26
10						
Experimental	30	1.21	1.39	0.51	0.38	0.96
Control	17	0.23	0.25	0.11	0.12	0.45
11						
Experimental	12	3.71	1.89	0.48	0.85	1.52
Control	7	0.04	0.50	0.19	-0.21	0.54
12+						
Experimental	5	3.87	1.78	0.36	0.70	1.97
Control	4	0.42	-0.14	-0.39	0.16	0.32

Effect Size Calculation by School

An important issue involved the degree to which good results for poor readers could be obtained across a number of sites, using different teachers. As is evident in Table 10, there were differences between schools (some with extreme effect sizes), although the overall picture was similar to the combined results. Although there may have been differences in students from school to school, the placement test results indicate similar reading rate and accuracy.

Note that some schools were absent because experimental group students were not represented in every school.

Table 10
Effect Size (d) Calculation by School

	<i>n</i>	TOPA	Word Attack	PNT	Digit Span	Spelling
Coolaroo						
Experimental	14	0.54	1.19	1.39	0.35	0.89
Control	8	0.14	0.17	-0.02	-0.06	0.09
Mill Pk						
Experimental	17	1.11	1.49	0.60	0.47	1.18
Control	9	0.70	-0.09	0.32	1.44	0.96
Moomba Pk						
Experimental	6	3.53	5.51	1.02	1.3	2.64
Control	0	X	X	X	X	X
Orana						
Experimental	22	2.42	1.72	0.33	0.69	0.98
Control	7	0.17	-0.25	0.20	-0.12	0.70
Our Lady's						
Experimental	17	0.63	0.99	0.40	0.37	0.62
Control	17	0.02	0.66	0.29	0.16	0.52
St Clares						
Experimental	28	1.50	1.30	0.41	0.53	1.49
Control	20	0.16	0.15	0.15	0.03	0.35
St Dominic's						
Experimental	17	1.64	2.14	1.03	0.70	0.62
Control	0	X	X	X	X	X
St Olivers						
Experimental	11	1.19	1.86	0.30	0.66	1.47
Control	0	X	X	X	X	X

Individual Differences in the Effects?

Barker and Torgesen (1995) described a convention for determining what percentage of students are assisted by an intervention. Such a measure adds to the information provided by mean score changes and effect sizes. These amalgamated measures, despite their usefulness, can partly disguise the overall picture when very large changes occur for some students, but negligible or even negative changes occur for a sizeable proportion of students. The threshold used by the authors (and others) was a raw score increase of two or more at posttest. This convention raises some concerns about the comparability of scores across tests. For example, the mean score for Digit Span at pretest was only nine. Thus, a change of two represents a relatively greater improvement than for other tests. The convention was altered in this study to allow comparability between tests. This was achieved by replacing the raw score changes with a change of one standard deviation. In any case, the major interest resided not in

the between-test comparisons, but rather with the experimental/control comparison for each test.

Table 11

Students whose Raw Score Increased by 1 Standard Deviation or More

	<i>n</i>	TOPA	Word Attack	PNT	Digit Span	Spelling
Experimental	134					
N of students		124	123	103	27	57
Percentage		92.5	91.8	76.9	20.1	42.5
Control	72					
N of students		3	47	46	9	6
Percentage		4.2	65.3	63.9	12.5	8.3

Orana: A Special Setting.

Were the students referred to Orana more severely reading disabled than those in the other participating schools? Table 12 provides a comparison of the phonological processing scores of the Orana students and those of other students. The table indicates that the Orana students were older, and their pretest scores were markedly lower on TOPA, Word Attack, Picture Naming, and Spelling, and marginally higher on Digit Span. The results are suggestive of a more severely disabled population referred to Orana.

Table 12

Mean Scores for Orana vs. Other schools (Experimental and Control Combined)

	Other schools: <i>n</i> = 177					
	Pretest			Posttest		
	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>
Age	83-153	114.43	12.77	90-160	121.44	12.53
TOPA	5-20	13.74	4.14	6-20	16.36	3.86
Word Attack	0-31	11.20	6.62	0-41	17.40	8.58
PNT	14-54	33.71	7.92	14-55	36.87	8.04
Digit Span	3-17	9.21	2.12	5-18	9.99	2.13
Spelling	1-39	17.00	5.96	4-41	21.21	5.89

	Orana <i>n</i> = 29					
	Pretest			Posttest		
	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>
Age	98-161	121.52	17.85	103-167	128.07	17.57
TOPA	1-20	11.69	4.18	6-20	16.83	4.09
Word Attack	0-17	6.83	4.87	0-27	14.62	7.70
PNT	6-48	29.86	8.77	9-50	33.55	9.24
Digit Span	7-15	9.86	2.12	7-13	10.59	1.68
Spelling	5-20	12.03	4.26	7-24	16.00	4.50

In Table 13 it is apparent that, despite having students lower on most of the measures, the effect sizes for students taught by Orana staff were higher on all measures except Spelling. Note that the number of students in Table 12 exceeds that in Table 11 because the Orana staff were contracted to teach a program in one of the participating schools.

Table 13

Effect sizes (d) Obtained by Orana vs. Other schools (Experimental Only)

	<i>n</i>	TOPA	Word Attack	PNT	Digit Span	Spelling
Other schools	98	1.23	1.35	0.57	0.54	1.16
Orana	36	1.45	1.45	0.83	0.53	0.78

Results for Students Who Participated in Consecutive Programs

There were five students (in the same school) who first participated in Level A, and then Level B. The effect sizes for the relevant programs are in Table 14. It was evident that progress continued for those students who participated in consecutive programs.

Table 14

Effect Sizes (d) for Level A and Level B

	ES (<i>d</i>)	
	Level A	Level B
TOPA	0.99	1.89
Word Attack	3.68	1.63
PNT	0.35	1.97
Digit Span	0.59	0.08
Spelling	1.24	2.22

In summary, were there differences in effect across different groupings? There were some differences across sex, but they were not uniform across measured variables. More importantly, the pattern of results tended to be similar - large effects for the literacy related variables (TOPA, Word Attack, Spelling), and moderate for the other phonological processes (Digit Span and Naming). Similarly, across age and school groupings the same pattern was readily discernible. At a special setting in which more severely reading disabled students were present, effect sizes were of similar order to those of their less disabled age-peers. In a small group of students completing consecutive program levels, the pattern of results was repeated at the additional level.

Are There Differences Between the Program Levels - A, B?

There were two levels of the *Corrective Reading* program for students, depending on their placement test result. How did the two groups of students initially differ with respect to phonological processes? Table 15 allows comparison on those variables. It is evident that the lower rate and accuracy scores on the placement test (the basis for assigning the students to the different levels) was also reflected in lower scores on some phonological processes tests. Level A students scored lower than Level B students on TOPA, Word Attack, and Spelling; however, they were not markedly different from Level B students on Picture Naming or Digit Span. The Level A students were older by three months, on average.

Table 15:

Differences in Phonological Processes between Levels

	<i>n</i>	Level A 119			Level B 87		
		Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>
Age							
Pretest		83-161	116.79	15.67	92-144	113.62	10.50
Posttest		90-167	123.45	15.38	99-148	120.98	10.33
TOPA							
Pretest		1-20	12.53	4.10	5-20	14.70	4.01
Posttest		6-20	16.43	3.83	6-20	16.38	3.99
Word Attack							
Pretest		0-29	7.58	4.99	0-31	14.69	6.31
Posttest		0-35	15.83	8.15	0-41	18.57	8.78
PNT							
Pretest		6-54	33.08	8.54	14-52	33.22	7.61
Posttest		9-54	35.69	7.91	14-55	37.20	8.64
Digit Span							
Pretest		3-17	9.08	2.21	6-14	9.60	1.97
Posttest		2-18	9.70	2.11	6-17	10.48	2.08
Spelling							
Pretest		1-39	13.90	5.28	6-31	19.56	5.37
Posttest		4-34	18.07	4.98	13-41	23.77	5.72

Level A and Level B Separately

Another element of the thesis involved a more detailed examination of the effects of the two levels of the *Corrective Reading* program (Level A, and Level B1). The summaries of raw and transformed data are in Tables 16, 17, 18, and 19. Note in Tables 17 and 19 that only the raw scores for Word Attack, Digit Span and Spelling were transformed.

Table 16:

Experimental vs Control Groups: Mean Raw Scores for Level A.

	<i>n</i>	Age	TOPA	Word Attack	PNT	Digit Span	Spelling
Control	34						
Pretest							
Range		83-153	8-20	0-29	6-54	4-17	5-39
<i>M</i>		121.18	14.15	7.97	35.24	9.12	15.38
<i>SD</i>		16.52	4.40	5.51	9.53	2.51	6.82
Posttest							
Range		90-160	7-20	0-28	17-52	5-13	4-34
<i>M</i>		127.85	14.59	8.56	36.06	9.00	16.35
<i>SD</i>		16.40	4.53	5.25	8.31	1.84	6.04
Experimental	85						
Pretest							
Range		92-161	1-20	0-24	12-53	3-15	1-22
<i>M</i>		115.04	11.88	7.42	32.21	9.07	13.31
<i>SD</i>		15.06	3.82	4.79	8.01	2.09	4.43
Posttest							
Range		97-167	6-20	0-35	9-54	2-18	7-33
<i>M</i>		121.69	17.16	18.74	35.54	9.98	18.75
<i>SD</i>		14.68	3.26	7.24	7.79	2.15	4.33

Table 17

Experimental vs Control Groups: Mean Power Transformed Scores for Level A

	<i>n</i>	Word Attack	Digit Span	Spelling
Control	72			
Pretest				
Minimum		1.00	0.06	2.72
Maximum		2.49	0.26	9.76
<i>M</i>		1.90	0.12	5.91
<i>SD</i>		0.31	0.03	1.37
Posttest				
Minimum		1.00	0.07	2.37
Maximum		2.47	0.21	8.97
<i>M</i>		1.94	0.12	6.26
<i>SD</i>		0.30	0.02	1.33
Experimental	134			
Pretest				
Minimum		1.00	0.07	1.00
Maximum		2.53	0.35	8.47
<i>M</i>		1.83	0.12	5.40
<i>SD</i>		0.34	0.03	1.28
Posttest				
Minimum		1.00	0.06	3.35
Maximum		2.72	0.18	10.07
<i>M</i>		2.20	0.11	6.58
<i>SD</i>		0.30	0.02	1.12

Table 18:

Experimental vs Control Groups: Mean Raw Scores for Level B

	<i>n</i>	Age	TOPA	Word Attack	PNT	Digit Span	Spelling
Control	38						
Pretest							
Range		92-135	5-20	6-31	19-48	6-14	11-31
<i>M</i>		111.60	15.25	15.10	31.85	9.44	20.10
<i>SD</i>		11.16	4.38	6.60	7.64	1.96	5.02
Posttest							
Range		99-145	6-20	2-37	14-52	7-15	13-41
<i>M</i>		118.50	16.15	17.52	34.88	10.08	23.48
<i>SD</i>		11.32	4.30	8.20	8.51	1.75	5.98
Experimental	49						
Pretest							
Range		102-144	6-20	0-27	14-52	6-14	6-31
<i>M</i>		116.16	14.00	14.16	34.95	9.82	18.87
<i>SD</i>		9.13	3.43	5.97	7.30	2.00	5.76
Posttest							
Range		112-148	6-20	0-41	17-55	6-17	15-37
<i>M</i>		124.11	16.68	19.89	40.22	10.97	24.13
<i>SD</i>		8.01	3.60	9.41	7.94	2.37	5.43

Table 19:

Experimental vs Control Groups: Mean Power Transformed Scores for Level B

	<i>n</i>	Word Attack	Digit Span	Spelling
Control	38			
Pretest				
Minimum		1.68	0.08	4.44
Maximum		2.49	0.18	8.47
<i>M</i>		2.05	0.12	6.40
<i>SD</i>		0.23	0.02	1.07
Posttest				
Minimum		1.45	0.07	4.93
Maximum		2.47	0.15	8.97
<i>M</i>		2.09	0.11	6.87
<i>SD</i>		0.25	0.02	1.00
Experimental	49			
Pretest				
Minimum		1	0.08	3.05
Maximum		2.53	0.18	8.47
<i>M</i>		2.06	0.11	6.23
<i>SD</i>		0.25	0.02	1.14
Posttest				
Minimum		1	0.06	5.16
Maximum		2.72	0.18	10.07
<i>M</i>		2.23	0.10	7.33
<i>SD</i>		0.33	0.02	1.08

Effect Size Calculation by Program

The effect sizes of the individual programs are presented in Table 20 below. As can be readily observed, the pattern of results was similar to those noted earlier, though there were considerable differences in the effect sizes for the same variable from one program to the other. In particular, the results for Level A on literacy variables (TOPA, Word Attack, and Spelling) exceeded those for Level B.

Table 20

Effect Sizes (d) for Level A and Level B

	Level A		Level B	
	Experimental <i>n</i>	Control	Experimental	Control
	85	34	49	38
TOPA	1.70	0.11	0.87	0.24
Word Attack	1.96	0.12	0.80	0.35
PNT	0.40	0.11	0.79	0.43
Digit Span	0.46	-0.06	0.58	0.38
Spelling	1.26	0.15	0.95	0.62

The Outcomes For the Level A Program

The research questions for this section paralleled those addressed in the first section, but examined the two levels of the *Corrective Reading* program separately. Did participation in the *Corrective Reading* program Level A increase phonemic awareness, phonological recoding (word attack) skills, and other phonological processes (naming, working memory)? Did the *Corrective Reading* program Level A effects generalise to spelling?

Assumptions of normality and homogeneity of variance were tested for all data used in ANCOVA and ANOVA analyses, and data transformations were performed when necessary, as shown in Table 21 below. Power transformations were used for all experimental and control group, pretest and posttest data for Word Attack, Digit Span, and Spelling. These were effective in meeting normality and variance assumptions except for posttest spelling, for which the more stringent requirement for significance $\alpha = .01$ was adopted. Transformations were unhelpful for pretest Picture Naming experimental scores in obtaining normality (though importantly, assumptions of homogeneity of variance were met), and were unnecessary for the other Picture Naming scores. A similar situation occurred for Digit Span after transformation, in that the variance assumption was met. Transformations were unhelpful for TOPA, for which the more stringent requirement for significance $\alpha = .01$ was adopted.

Table 21

Tests of Normality and Homogeneity of Variance: Level A

Variable	Test	Group	Trans formation	Lilliefors' test of Normality	Levene's test of Homogeneity of Variance
Word Attack	Pretest	Control	Power	.08	
		Experimental	Power	.09	.33
	Posttest	Control	Power	> .2	
		Experimental	Power	.03	1.00
Digit Span	Pretest	Control	Power	.01	
		Experimental	Power	.00	.86
	Posttest	Control	Power	.00	
		Experimental	Power	.00	.28
Spelling	Pretest	Control	Power	> .2	
		Experimental	Power	> .2	.06
	Posttest	Control	Power	> .2	
		Experimental	Power	> .2	.01
PNT	Pretest	Control	No	> .2	
		Experimental	No	.02	.39
	Posttest	Control	No	> .2	
		Experimental	No	> .2	.64
TOPA	Pretest	Control	No	.01	
		Experimental	No	> .2	.03
	Posttest	Control	No	.10	
		Experimental	No	.00	.00

The Outcomes For the Level B Program

The questions asked about the Level A program were duplicated for Level B. Did participation in the *Corrective Reading* program Level B increase phonemic awareness, phonological recoding (word attack) skills, and other phonological processes (naming, working memory)? Did the *Corrective Reading* program Level B effects generalise to spelling?

Assumptions of normality and homogeneity of variance were tested for all data used in ANCOVA and ANOVA analyses, and data transformations were performed when necessary, as shown in Table 22 . Power transformations were used for all experimental and control group, pretest and posttest data for Word Attack, Digit Span, and Spelling. Transformations were unnecessary for Picture Naming, and were unhelpful for posttest TOPA scores in obtaining normality, although importantly, assumptions of homogeneity of variance were met. A similar situation occurred for Digit Span after transformation.

Table 22

Tests of Normality and Homogeneity of Variance: Level B

Variable	Test	Group	Trans formation	Lilliefors' test of Normality	Levene's test of Homogeneity of Variance
Word Attack	Pretest	Control	Power	> .2	.80
		Experimental	Power	> .2	
	Posttest	Control	Power	> .2	
		Experimental	Power	> .2	
Digit Span	Pretest	Control	Power	.00	.22
		Experimental	Power	.00	
	Posttest	Control	Power	.00	
		Experimental	Power	.01	
Spelling	Pretest	Control	Power	> .2	.78
		Experimental	Power	> .2	
	Posttest	Control	Power	> .2	
		Experimental	Power	> .2	
PNT	Pretest	Control	No	.15	.07
		Experimental	No	> .2	
	Posttest	Control	No	> .2	
		Experimental	No	> .2	
TOPA	Pretest	Control	No	.07	.11
		Experimental	No	> .2	
	Posttest	Control	No	.01	
		Experimental	No	.00	

The results of ANCOVA and ANOVA procedures and ANCOVA procedures for Levels A and B1 separately are presented below.

TOPA (Level A)

Results for TOPA were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 115) = 14.41, p < .001$ so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pretest scores covaried significantly with posttest scores for both the control, $F(1, 115) = 92.49, p < .001$, and experimental groups, $F(1, 115) = 46.46, p < .001$. With the pretest results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 115) = 35.77, p < .001$.

Results for TOPA were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 117) = .05, p = .828, \text{power} = .04$, but a significant main effect was found for time, $F(1, 117) = 204.01, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 117) = 64.18, p <$

.001, power = 1.00, which is illustrated in Figure 6. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, $F(1, 117) = 7.82, p = .006, \text{power} = .79, d = -0.65$, and at posttest, $F(1, 117) = 12.01, p = .001, \text{power} = .93, d = 0.81$. Further, no significant pre- to posttest difference was found for the control, $F(1, 117) = 0.75, p = .389, \text{power} = .17, d = 0.06$, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 117) = 267.44, p < .001, \text{power} = 1.00, d = 1.30$, and the magnitude of effect was large for the experimental group.

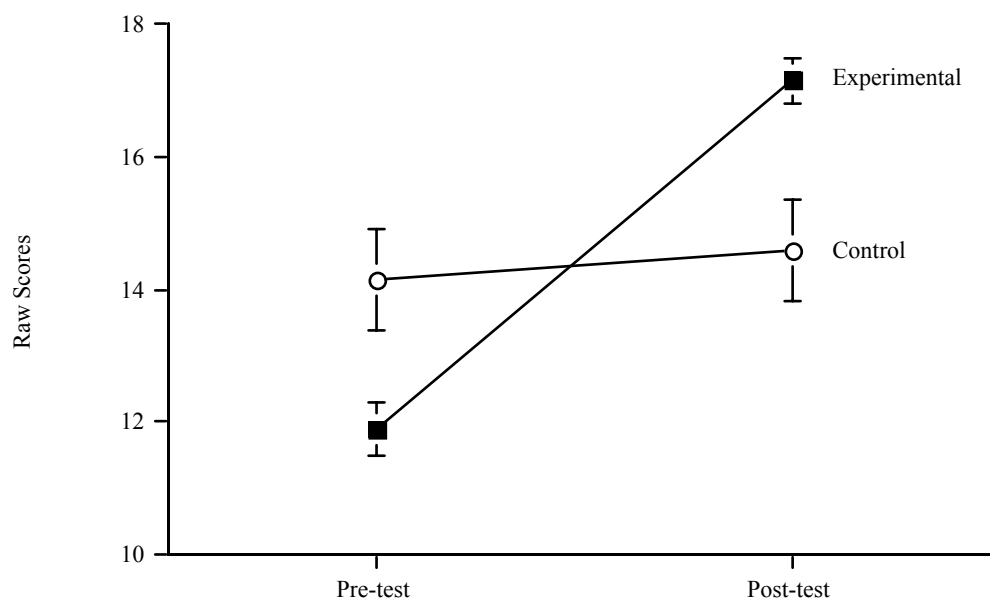


Figure 6. Interaction (+SE) between experimental and control group at pre and posttest for TOPA (Level A).

TOPA (Level B)

Results for TOPA were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 83) = 0.19, p = .667$. With the pretest results partialled out there was no significant overall difference between the experimental and control groups, $F(1, 84) = 3.79, p = .055$.

Results for TOPA were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 85) = 0.09, p = .761, \text{power} = .05$, but a significant main effect was found for time, $F(1, 85) =$

25.75, $p < .001$, power = 1.00, and the group-by-time interaction, $F(1, 85) = 3.99$, $p = .049$, power = .50, which is illustrated in Figure 7. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 85) = 0.25$, $p = .616$, power = .05, $d = -0.37$, and at posttest, $F(1, 85) = 1.15$, $p = .286$, power = .19, $d = 0.16$. Further, no significant pre- to posttest difference was found for the control, $F(1, 85) = 3.44$, $p = .067$, power = .45, $d = 0.24$, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 85) = 26.30$, $p < .001$, power = 1.00, $d = 0.87$, and the magnitude of effect was large for the experimental group.

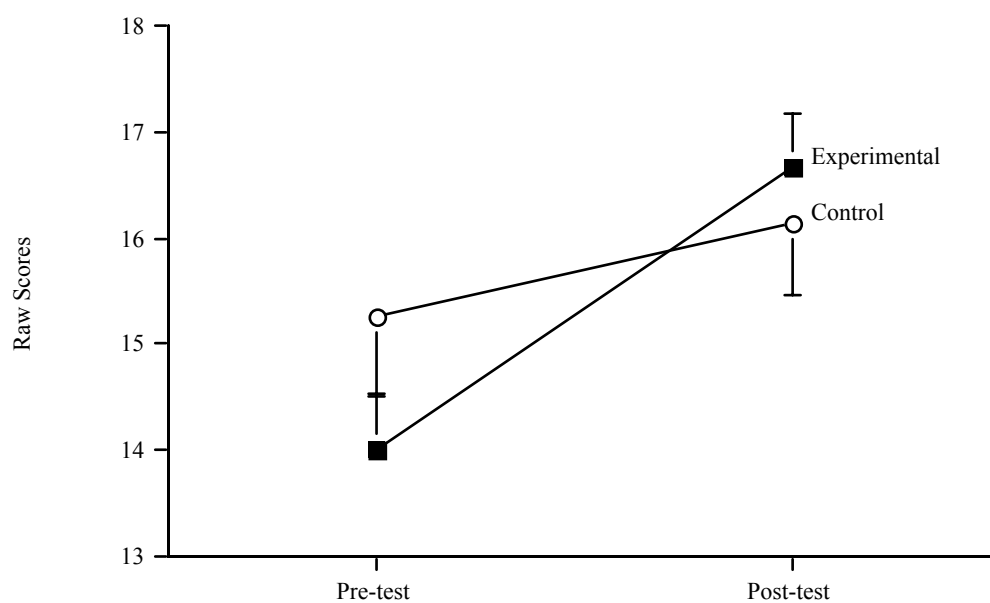


Figure 7. Interaction (+SE) between experimental and control group at pre and posttest for TOPA (Level B1).

Word Attack (Level A)

Results for Word Attack were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 115) = 3.79$, $p = .054$. With the pretest results partialled out for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 116) = 86.50$, $p < .001$.

Results for the power transformed Word Attack scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main

effect was found for group, $F(1, 117) = 11.05, p = .001, \text{power} = .91$, and for time, $F(1, 117) = 218.49, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 117) = 67.28, p < .001, \text{power} = 1.00$, which is illustrated in Figure 8. Follow-up testing of the interaction using simple main effects found a non significant difference between the experimental and control groups at pretest, $F(1, 117) = 0.32, p = .575, \text{power} = .04, d = -0.12$, but a significant difference at posttest, $F(1, 117) = 48.6, p < .001, \text{power} = 1.00, d = 1.64$. Further, no significant pre- to posttest difference was found for the control, $F(1, 117) = 0.94, p = .335, \text{power} = .17, d = 0.35$, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 204) = 284.84, p < .001, \text{power} = 1.00, d = 1.95$, however the magnitude of effect was large for the experimental group.

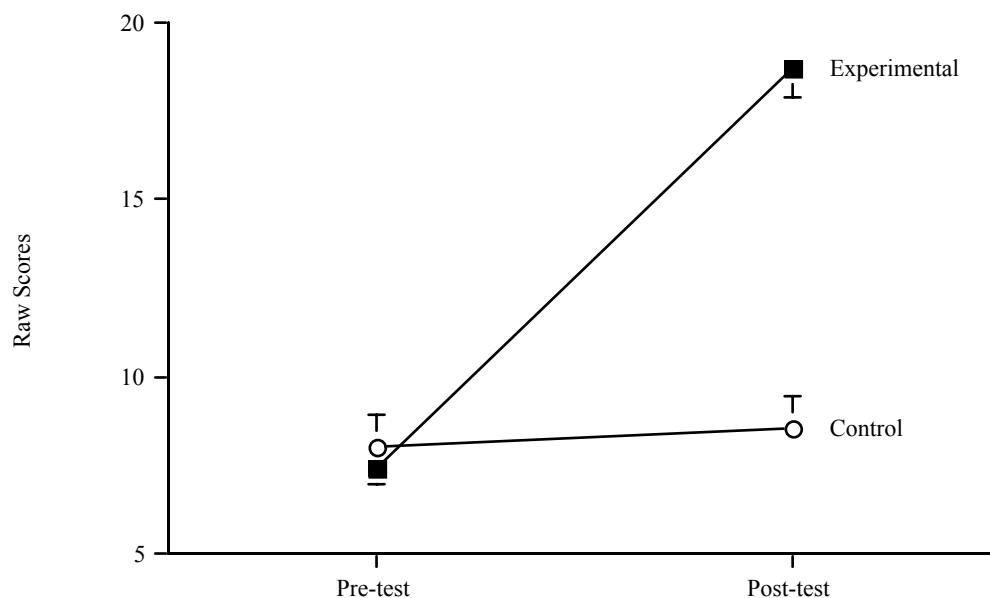


Figure 8. Interaction (+SE) between experimental and control group at pre and posttest for Word Attack (Level A).

Word Attack (Level B)

Results for Word Attack were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 83) = 1.92, p = .169$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 84) = 11.28, p = .001$.

Results for the power transformed Word Attack scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 85) = 1.94, p = .167, \text{power} = .28$, but a significant main effect was found for time, $F(1, 85) = 35.38, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 85) = 11.43, p < .001, \text{power} = .92$, which is illustrated in Figure 9. Follow-up testing of the interaction using simple main effects found a non significant difference between the experimental and control groups at pretest, $F(1, 85) = 0.06, p = .800, \text{power} = .04, d = -0.16$, but a significant difference at posttest, $F(1, 85) = 4.80, p = .031, \text{power} = .58, d = 0.29$. Further, no significant pre- to posttest difference was found for the control group, $F(1, 85) = 1.94, p = .167, \text{power} = .28, d = 0.35$, but a significant pre- to posttest difference was found for the experimental group, $F(1, 85) = 44.86, p < .001, \text{power} = 1.00, d = 0.8$. and the magnitude of effect was large for the experimental group.

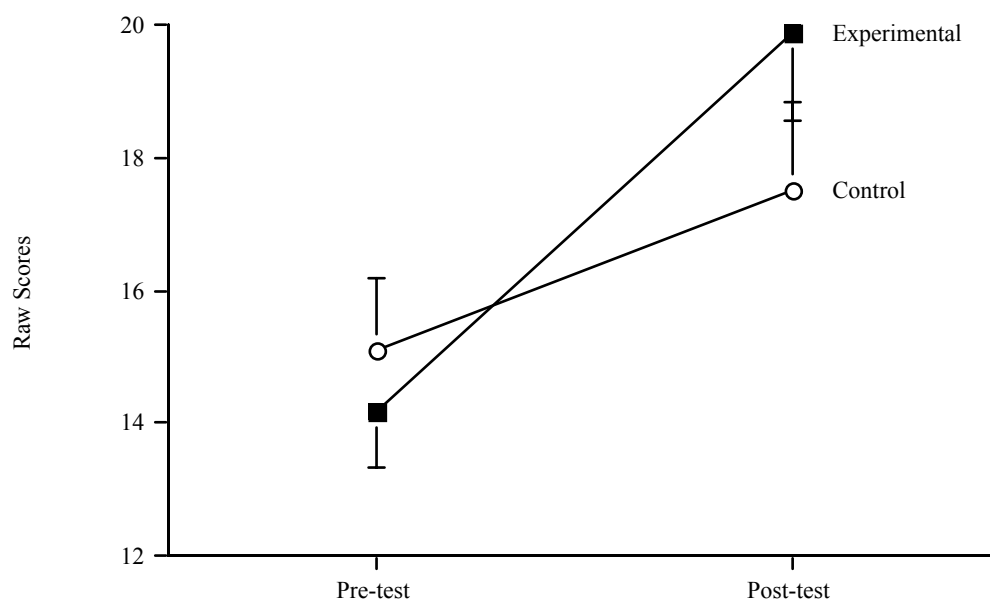


Figure 9. Interaction (+SE) between experimental and control group at pre and posttest for Word Attack (Level B1).

3. Picture Naming Test (Level A):

Results for Picture Naming Test were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed a non violation of the assumption of homogeneity of slopes, $F(1, 115) = 0.78, p = .378$. With the pretest results partialled out there

was no significant overall difference between the experimental and control groups, $F(1, 116) = 0.69, p = .408$.

Results for Picture Naming Test were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 117) = 1.56, p = .214$, power = 0.23, and for the group-by-time interaction, $F(1, 117) = 2.20, p = .141$, power = .31, but a significant main effect was found for time, $F(1, 117) = 13.12, p < .001$, power = .95, which is illustrated in Figure 10. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 117) = 3.10, p = .081$, power = .42, $d = -0.41$, and at posttest, $F(1, 117) = 0.27, p = .603$, power = .04, $d = -0.07$. Further, no significant pre- to posttest difference was found for the control group, $F(1, 117) = 0.47, p = .496$, power = .10, $d = 0.33$, but a significant pre- to posttest difference was found for the experimental group, $F(1, 117) = 14.85, p < .001$, power = .10, $d = 0.83$, and the magnitude of effect was large for the experimental group.

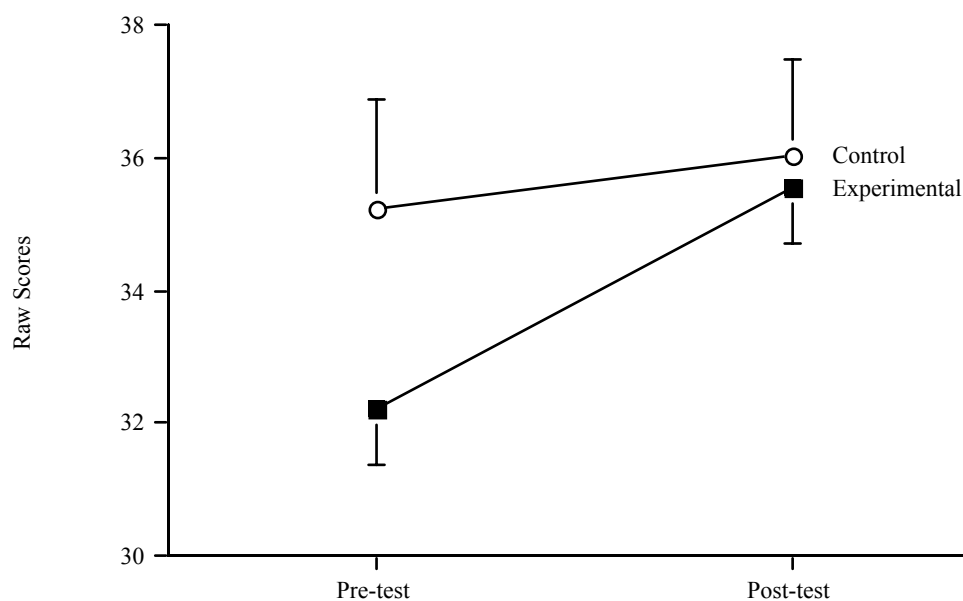


Figure 10. Interaction (+SE) between experimental and control group at pre and posttest for Picture Naming Test (Level A).

Picture Naming Test (Level B):

Results for Picture Naming Test were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest

scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 83) = 2.82, p = .097$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 84) = 22.46, p < .001$.

Results for Picture Naming Test were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, $F(1, 85) = 10.02, p = .002, \text{power} = .88$, and for time, $F(1, 85) = 55.04, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 85) = 18.13, p < .001, \text{power} = 1.00$, which is illustrated in Figure 11. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 85) = 2.62, p = .109, \text{power} = .36, d = 0.47$, but a significant difference at posttest, $F(1, 85) = 18.37, p < .001, \text{power} = 1.00, d = 0.74$. Further, no significant pre- to posttest difference was found for the control group, $F(1, 85) = 2.92, p = .091, \text{power} = .39, d = 0.43$, but a significant pre- to posttest difference was found for the experimental group, $F(1, 85) = 70.25, p < .001, \text{power} = 1.00, d = 0.79$, and the magnitude of effect was larger for the experimental group.

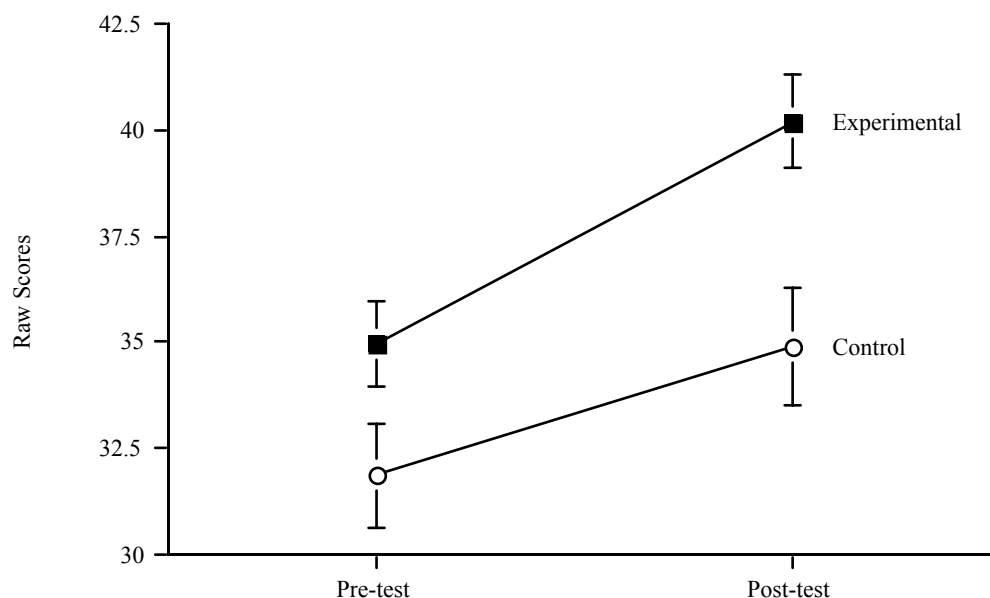


Figure 11. Interaction (+SE) between experimental and control group at pre and posttest for Picture Naming Test (Level B1).

4. Digit Span (Level A):

Results for Digit Span were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 115) = 0.41, p = .525$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 116) = 9.62, p = .005$.

Results for power transformed Digit Span scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for time, $F(1, 117) = 13.36, p < .001, \text{power} = .95$, but no significant main effect was found for group, $F(1, 117) = 1.61, p = .207, \text{power} = .24$, and the group-by-time interaction, $F(1, 117) = 3.72, p = .056, \text{power} = .48$, which is illustrated in Figure 12. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 117) = 0.00, p = .993, \text{power} = .03, d = -0.02$, but found a significant difference at posttest, $F(1, 117) = 7.90, p = .006, \text{power} = .79, d = 0.51$. Further, no significant pre- to posttest difference was found for the control, $F(1, 117) = 0.11, p = .746, \text{power} = .05, d = -0.06$, but a significant difference was found for the experimental groups, $F(1, 117) = 16.97, p < .001, \text{power} = .98, d = 0.46$, and the magnitude of effect was medium for the experimental group.

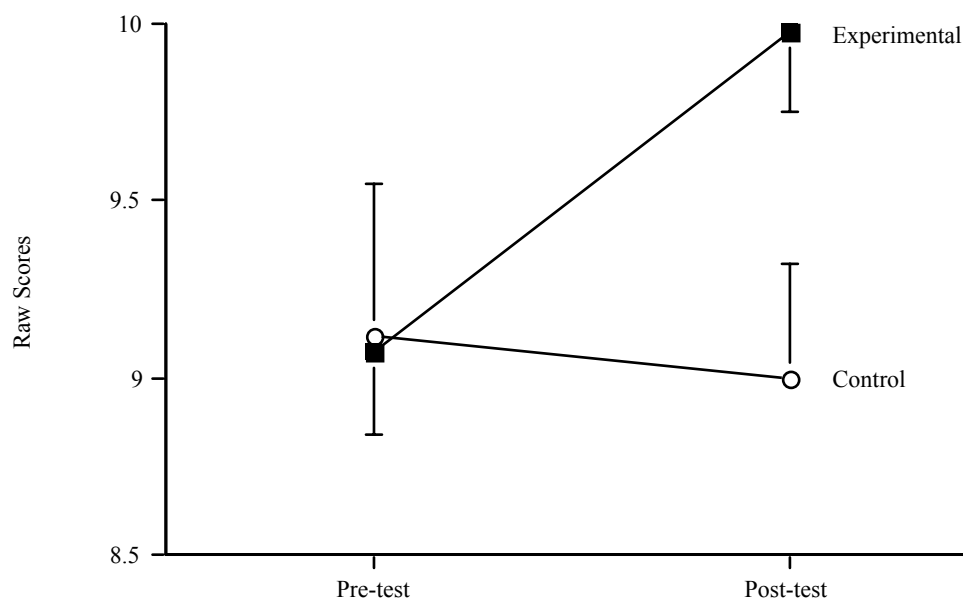


Figure 12. Interaction (+SE) between experimental and control group at pre and posttest for Digit Span (Level A).

Digit Span (Level B):

Results for Digit Span were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 83) = 3.33, p = .072$. With the pretest results partialled out there was no significant overall difference between the experimental and control groups, $F(1, 84) = 0.98, p = .326$.

Results for power transformed Digit Span scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 85) = 1.34, p = .250$, power = .21, and for the group-by-time interaction, $F(1, 85) = 0.21, p = .651$, power = .05, but a significant main effect was found for time, $F(1, 85) = 23.14, p < .001$, power = 1.00, which is illustrated in Figure 13. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 85) = .69, p = .407$, power = .17, $d = 0.21$, nor at posttest, $F(1, 85) = 1.69, p = .197$, power = .25, $d = 0.46$. Further, significant pre- to posttest differences were found for both the control, $F(1, 85) = 8.06, p = .006$, power = .80, $d = 0.38$, and the experimental groups, $F(1, 85) = 15.29, p < .001$, power = .97, $d = 0.58$, and the magnitude of effect was greater for the experimental group.

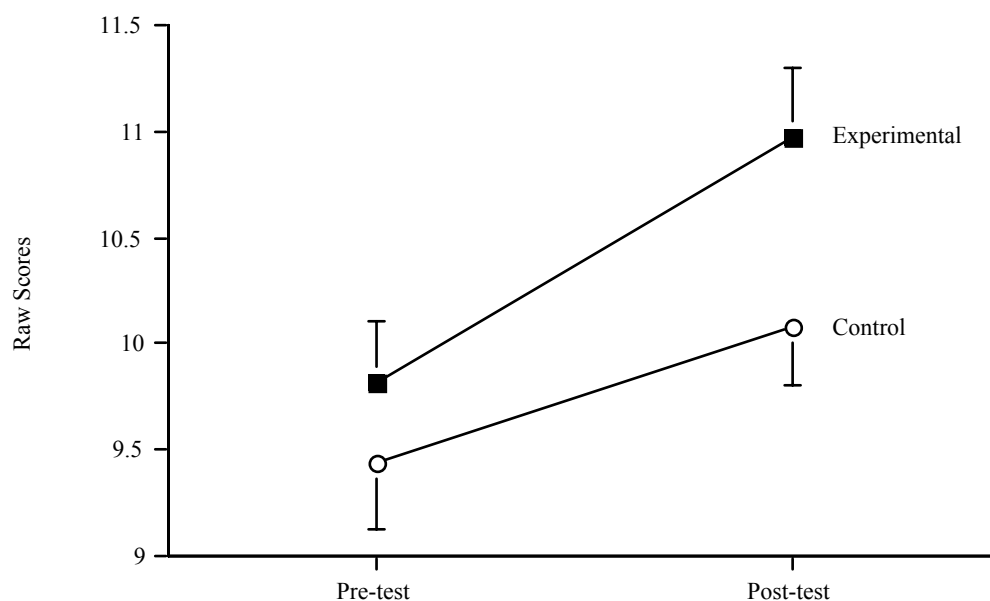


Figure 13. Interaction (+SE) between experimental and control group at pre and posttest for Digit Span (Level B1).

5. Brigance Spelling (Level A):

Results for Brigance Spelling were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 115) = 3.77, p = .055$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 116) = 21.73, p < .001$.

Results for the power transformed Spelling scores were analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 117) = 0.10, p = .750, \text{power} = .05$, but a significant main effect was found for time, $F(1, 117) = 105.61, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 117) = 23.94, p < .001, \text{power} = 1.00$, which is illustrated in Figure 14. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 117) = 3.03, p = .085, \text{power} = .41, d = -0.42$, but a significant difference at posttest, $F(1, 117) = 6.97, p = .009, \text{power} = .74, d = 0.51$. Further, no significant pre- to posttest difference was found for the control group, $F(1, 117) = 1.84, p = .117, \text{power} = .27, d = 0.15$, but a significant difference was found for the experimental group, $F(1, 117) = 127.71, p < .001, \text{power} = 1.00, d = 1.26$ and the magnitude of effect was large for the experimental group.

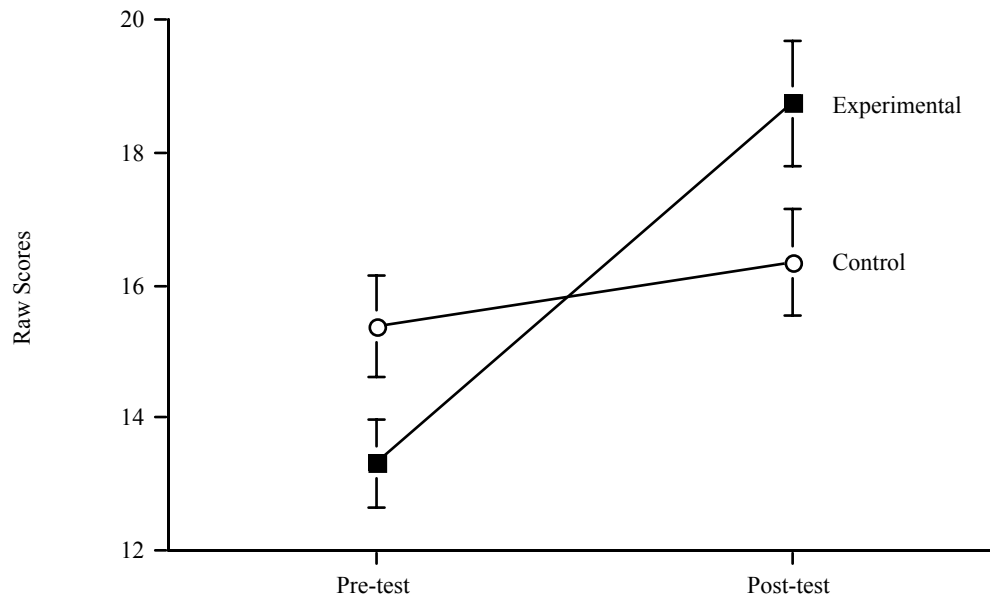


Figure 14. Interaction (+SE) between experimental and control group at pre and posttest for Brigance Spelling (Level A).

Brigance Spelling (Level B):

Results for Brigance Spelling were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 83) = 2.37, p = .128$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 84) = 12.90, p = .001$.

Results for the power transformed Spelling scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 85) = 0.44, p = .507, \text{power} = .09$, but a significant main effect was found for time, $F(1, 85) = 85.49, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 85) = 12.75, p = .001, \text{power} = .94$, which is illustrated in Figure 15. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 85) = 0.55, p = .460, \text{power} = .14, d = -0.23$, but a significant difference at posttest, $F(1, 85) = 4.23, p = .043, \text{power} = .53, d = 0.12$. Further, significant pre- to posttest differences were found for both the control, $F(1, 85) = 11.77, p = .001, \text{power} = .92, d = 0.62$, and the experimental groups, $F(1, 85) = 86.47, p <$

.001, power = 1.00, $d = 0.95$ and the magnitude of effect was large for the experimental group.

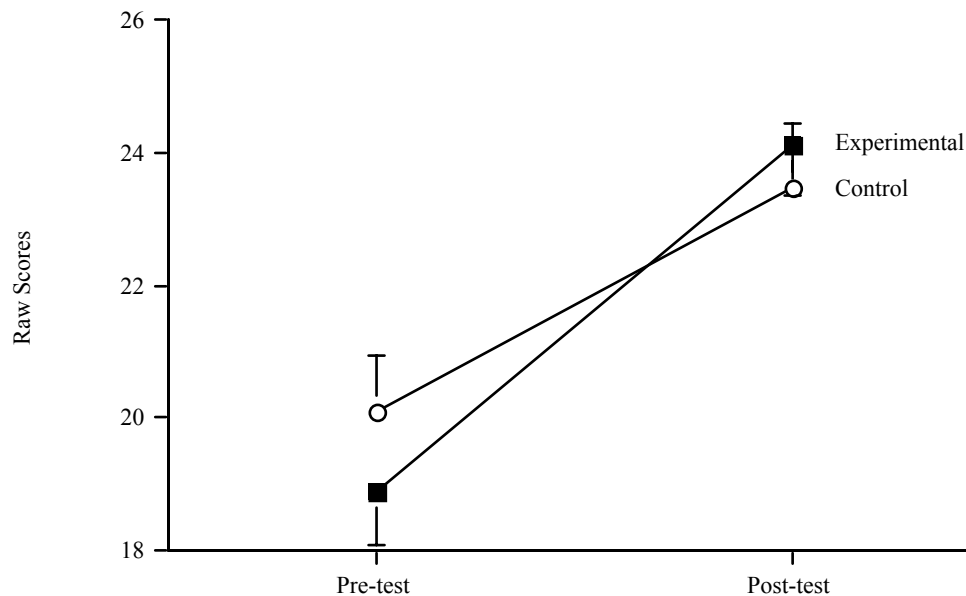


Figure 15. Interaction (+SE) between experimental and control group at pre and posttest for Brigance Spelling (Level B1).

Are the Effects of Educational Value? Examining Results in Relation to Test Norms.

An examination of pretest and posttest scores in relation to test norms provided an indication of the degree to which this sample of students with reading difficulties lacked normally developing phonological processing skills. Additionally, these figures provided information concerning the degree to which program effects “normalised” skills in phonological processes.

TOPA (Level A & B)

In Figure 16, raw scores for the total group (Levels A and B) are shown at pretest and posttest for the Test of Phonological Awareness. It was clear that the experimental group included a number of students now in the normal range, and a group-mean close to that range. By comparison, little change was evident for the wait-list group.

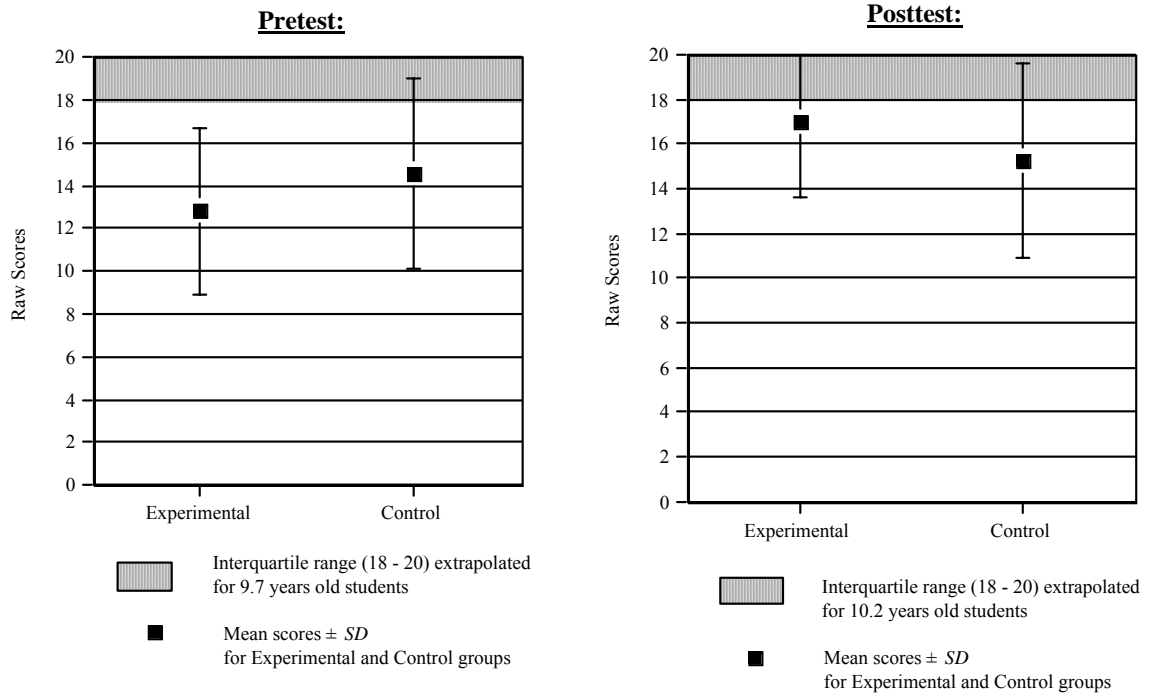


Figure 16. Mean scores for TOPA (A & B combined).

Word Attack (Level A & B)

In Figure 17, the program effects on Word Attack displayed a similar pattern to that for TOPA. The experimental group included a number of students now in the normal range, and a group-mean much closer to that range. By comparison, little change was evident for the wait-list group.

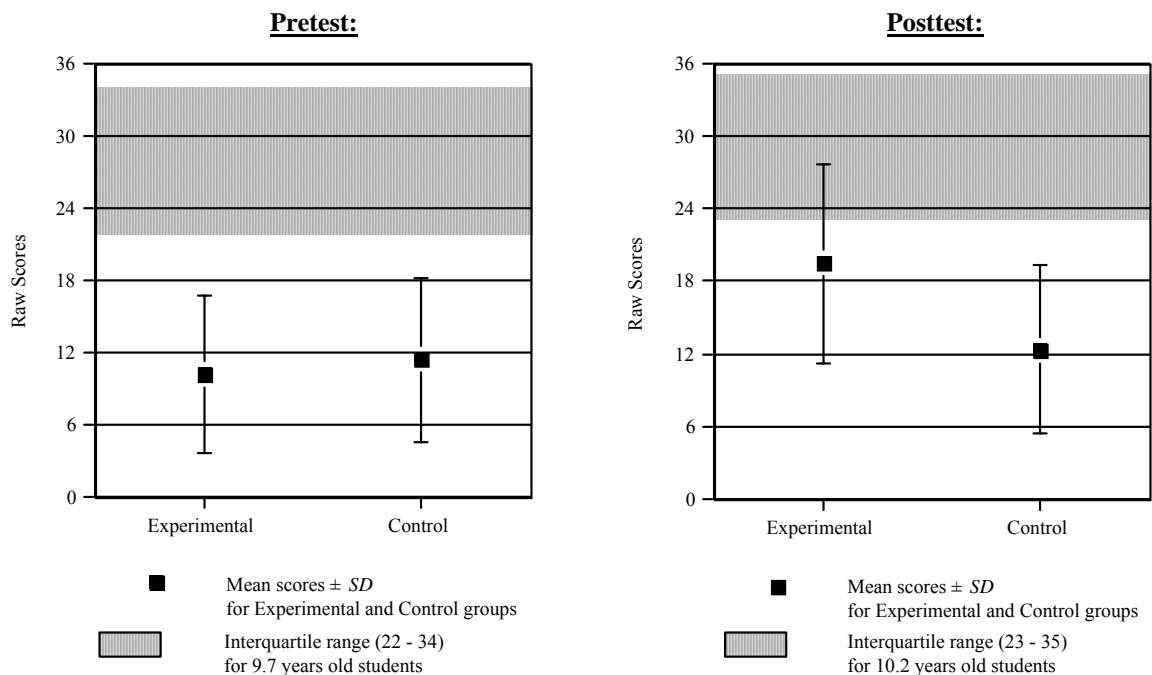


Figure 17. Mean scores for Word Attack (A & B combined).

Picture Naming Test (Level A & B)

Figure 18 displays the pretest-posttest scores for Picture Naming. As the Picture Naming Test was an experimental test even preliminary norms were unavailable. The figure displayed a moderate gain for the experimental group, and a smaller gain for the control group.

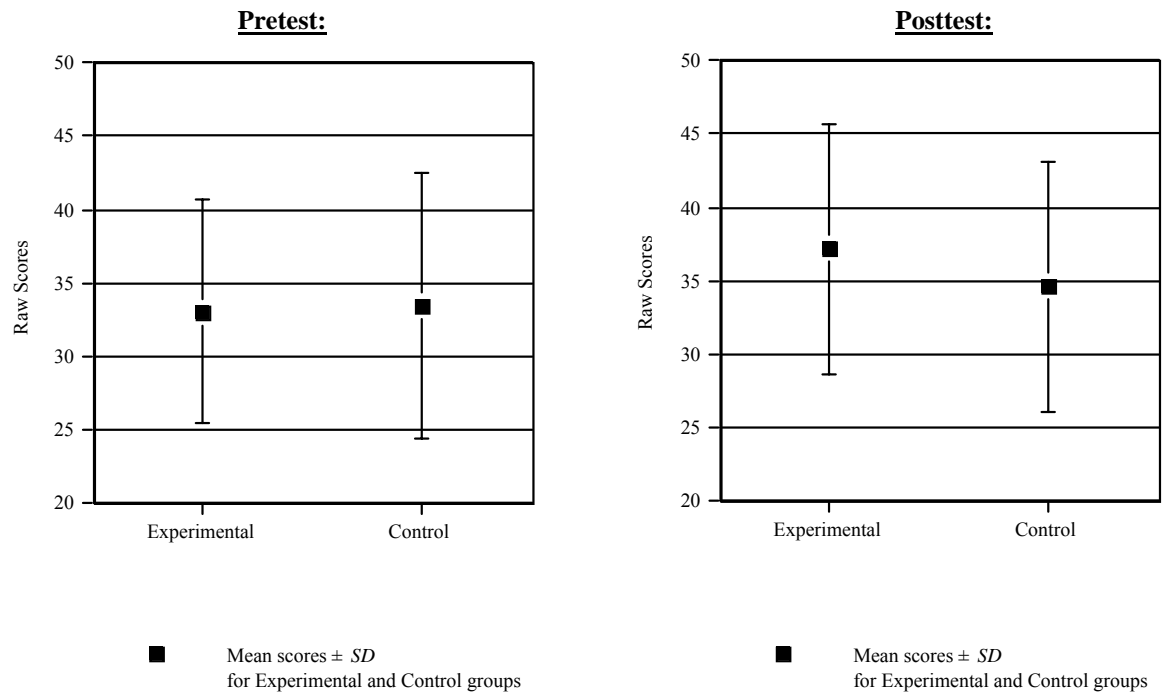


Figure 18. Mean scores for Picture Naming Test (A & B combined).

Digit Span (Level A & B)

Figure 19 shows the effects of the program on Digit Span in relation to the WISC-111 norms. The effect was moderate but elevated the experimental group mean close to the normal range, and an increase in the number of students within the normal range occurred. By contrast, little change was evident in the control group.

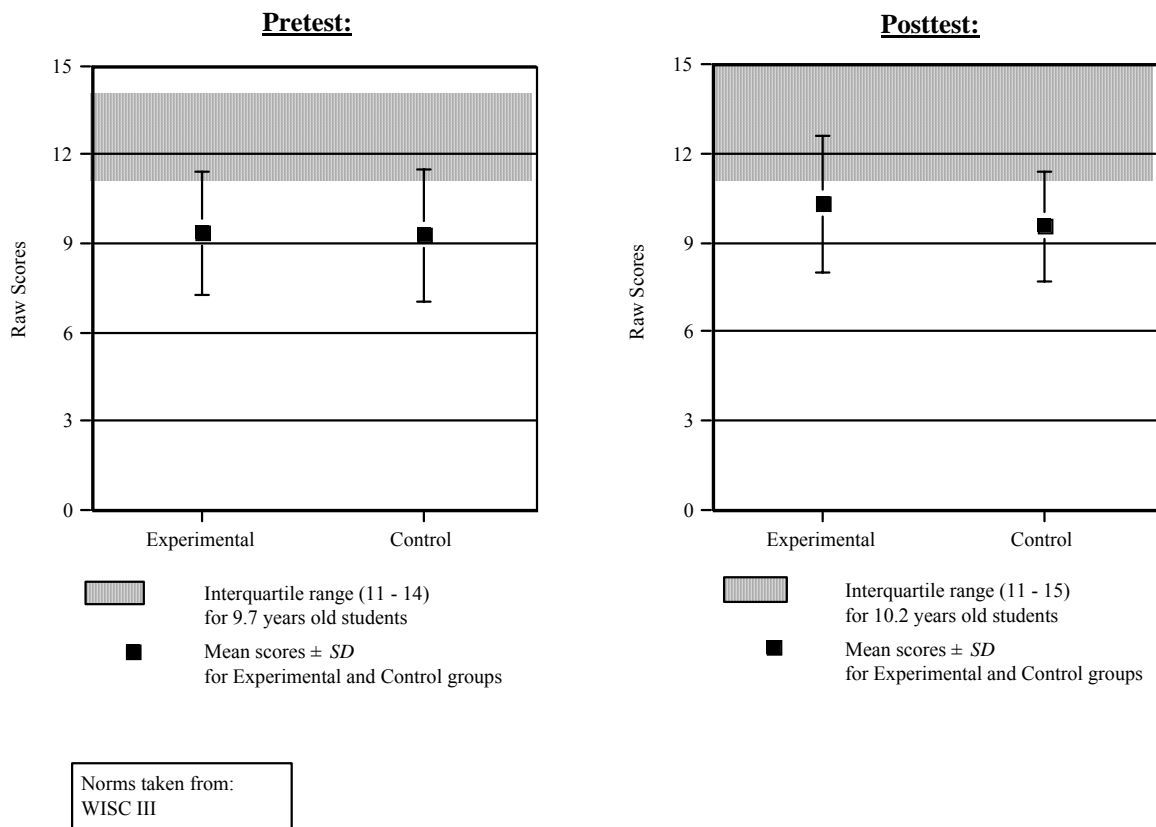


Figure 19. Mean scores for Digit Span (A & B combined).

Brigance Spelling (Level A & B)

Figure 20 displays the results for the Brigance Spelling test, and followed a similar pattern to the other tests. A feature of this graph was the extent to which all students were below the norm for this test, even after the program. The effects were strong; however, the students clearly remained in need of assistance.

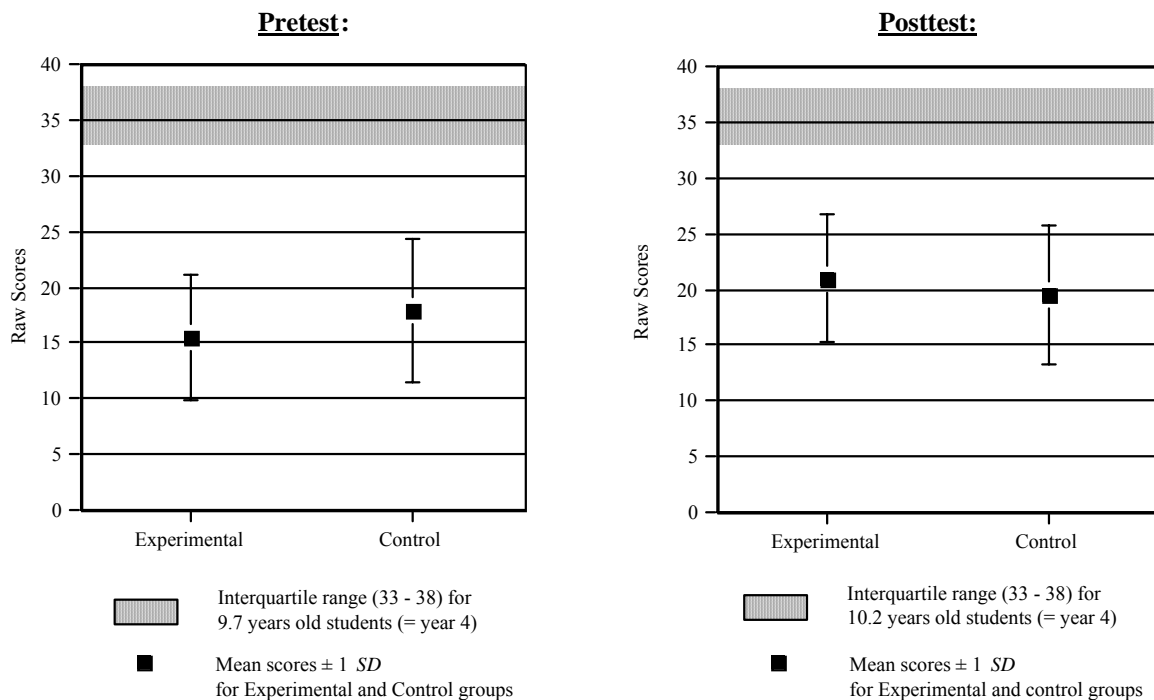
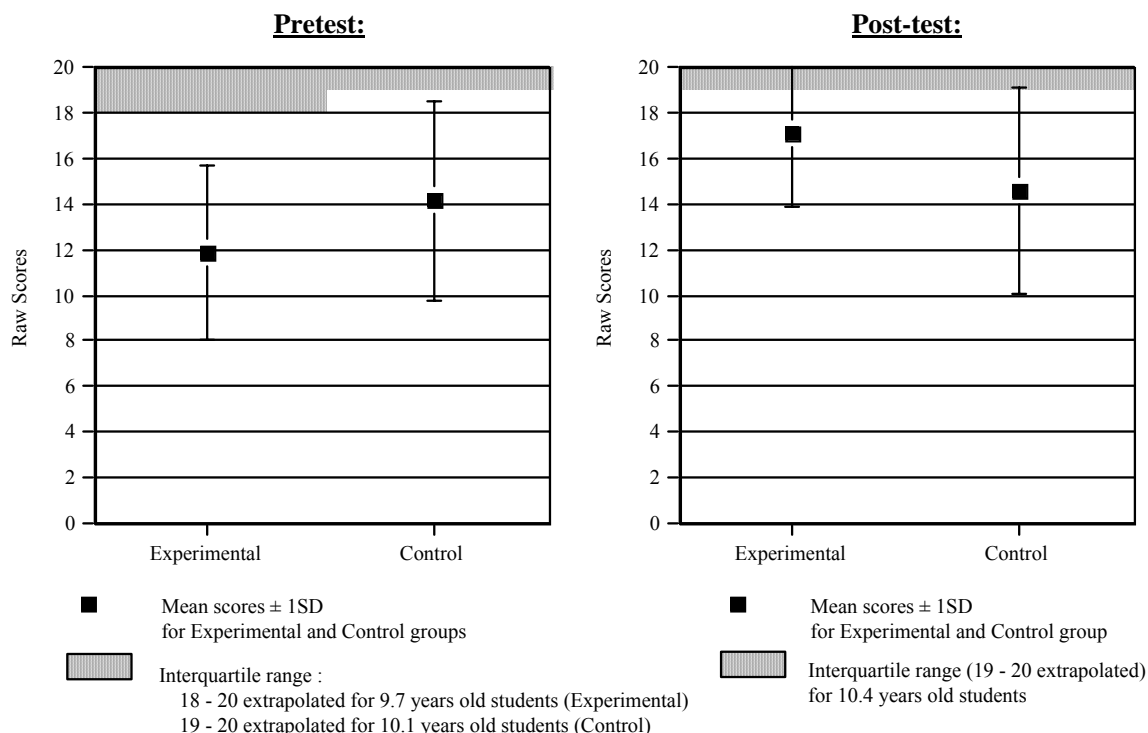


Figure 20. Mean scores for Brigance Spelling (A & B combined).

Normed Graphs for Level A and Level B Separately : TOPA (Level A)

In Figure 21, raw scores for the total group (Level A) are shown at pretest and posttest for the Test of Phonological Awareness. It was clear that the experimental group included a number of students now in the normal range, and a group-mean close to that range. By comparison, little change was evident for the wait-list group.



*Figure 21. Mean scores for TOPA (Level A).
TOPA (Level B)*

In Figure 22, raw scores for the total group (Level B) are shown at pretest and posttest for the Test of Phonological Awareness. It was clear that the experimental group included a number of students now in the normal range, and a group-mean close to that range. By comparison, little change was evident for the wait-list group.

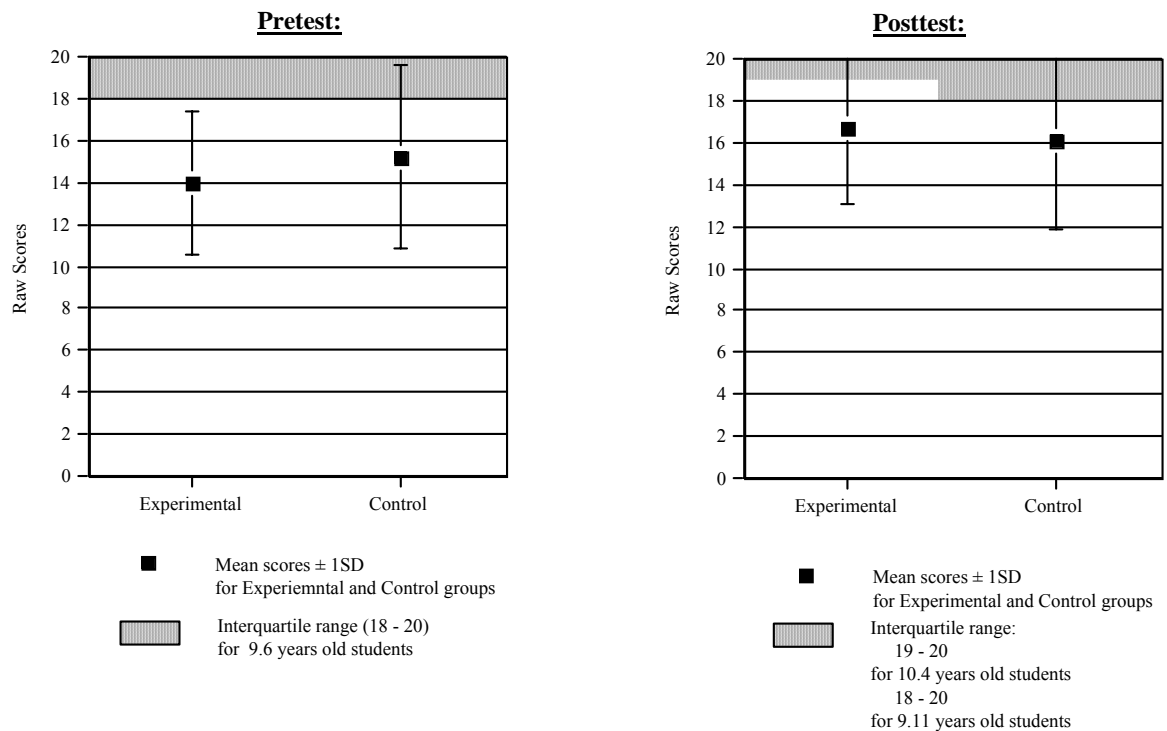


Figure 22. Mean scores for TOPA (Level B)

Word Attack (Level A)

In Figure 23, the program effects on Word Attack displayed a similar pattern to that for TOPA. The experimental group included some students now in the normal range, and a group-mean much closer to that range. By comparison, little change was evident for the wait-list group.

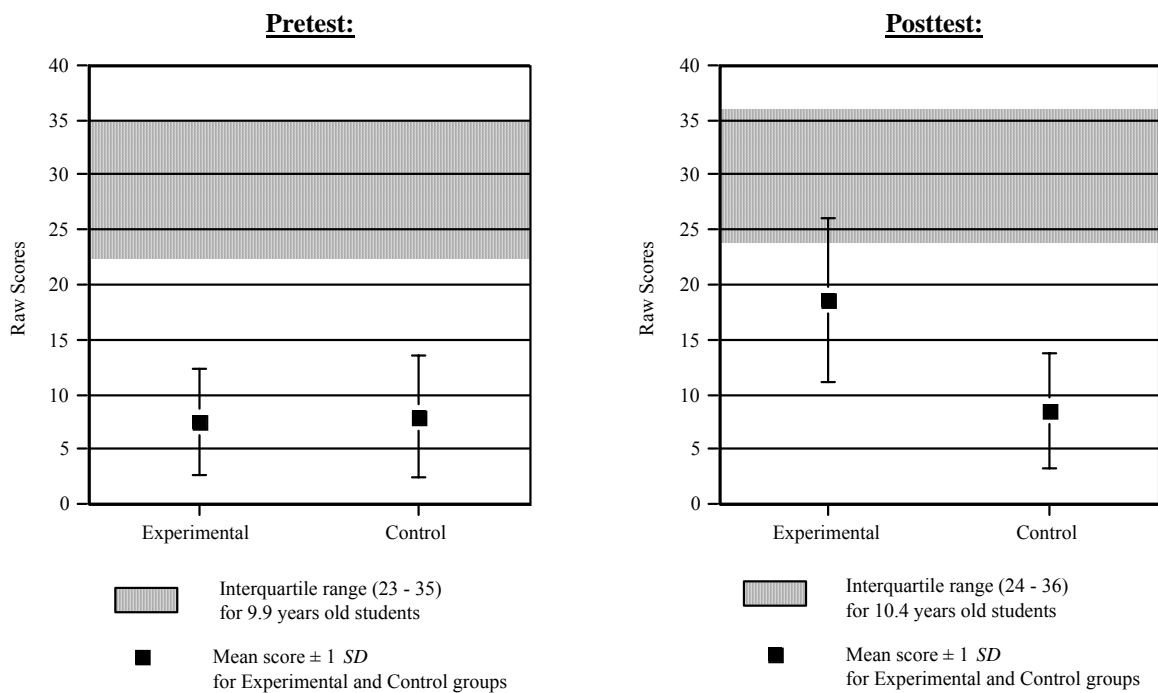


Figure 23. Mean scores for Word Attack (Level A).

Word Attack (Level B)

In Figure 24, the program effects on Word Attack displayed a similar pattern to that for TOPA. The experimental group included a number of students now in the normal range, and a group-mean much closer to that range. By comparison, little change was evident for the wait-list group.

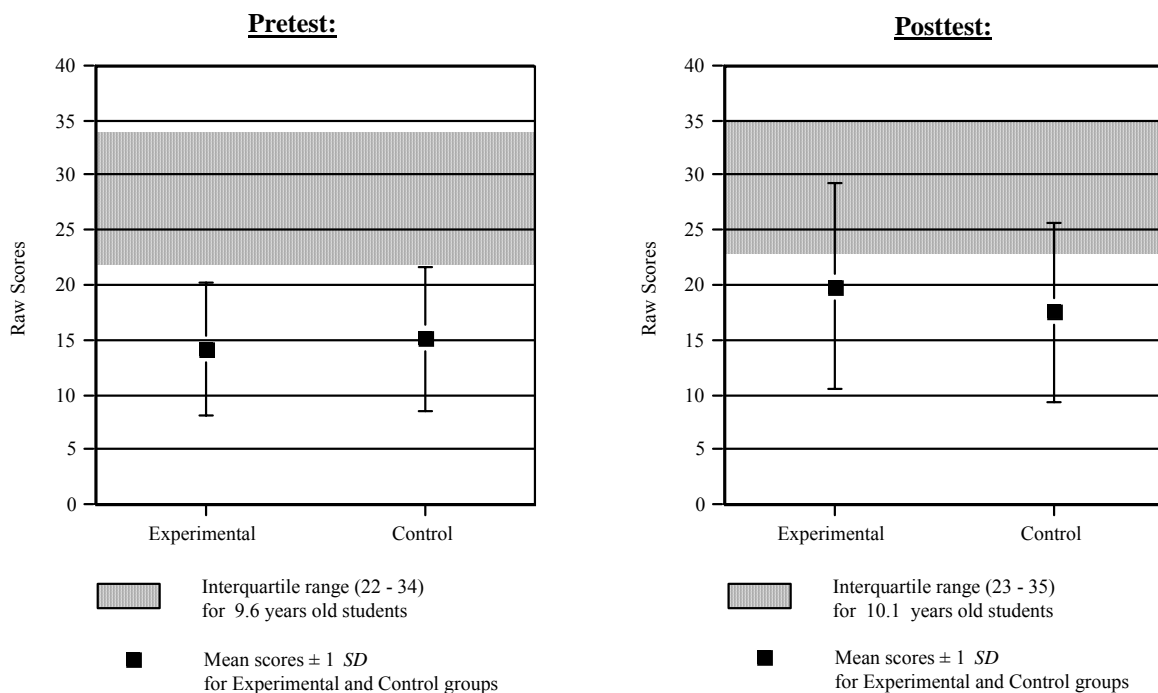


Figure 24. Mean scores for Word Attack (Level B).

Picture Naming Test (Level A)

Figure 25 displays the pretest-posttest scores for Picture Naming. As the Picture Naming Test was an experimental test no norms were available for it. The figure displays a moderate gain for the experimental group, and a smaller gain for the control group.

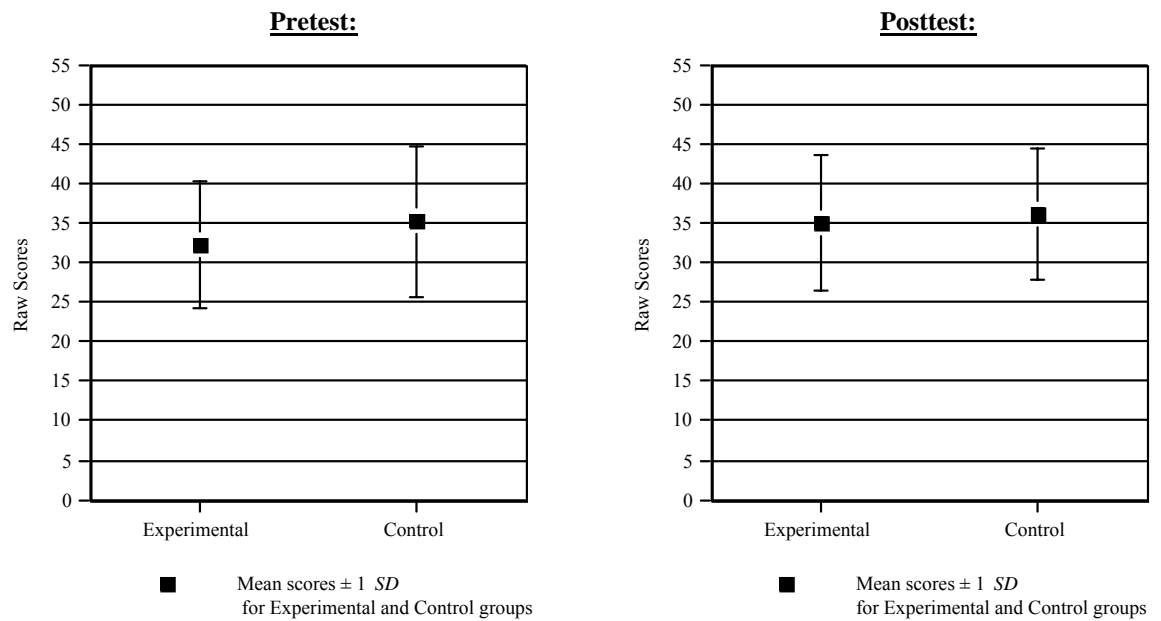


Figure 25. Mean scores for Picture Naming Test (Level A).

Picture Naming Test (Level B)

Figure 26 displays the pretest-posttest scores for Picture Naming. As the Picture Naming Test was an experimental test no norms were available for it. The figure displays a moderate gain for the experimental group, and a smaller gain for the control group.

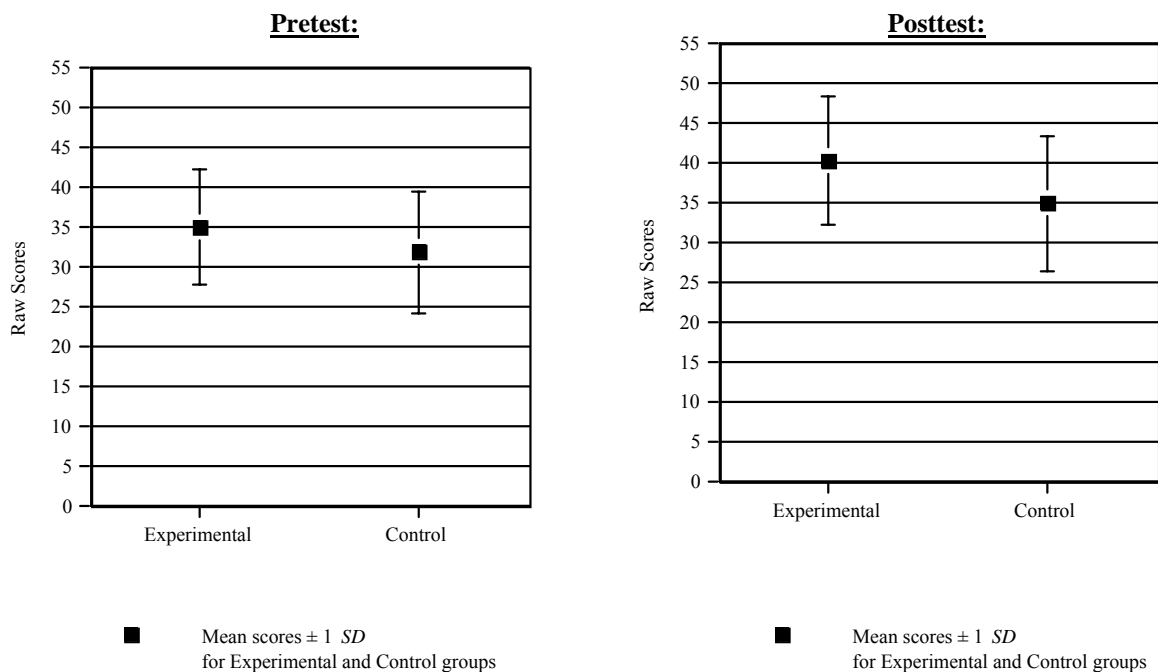


Figure 26. Mean scores for Picture Naming Test (Level B).

Digit Span (Level A)

Figure 27 shows the effects of the program on Digit Span in relation to the WISC-111 norms. The effect was moderate but elevated the experimental group mean close to the normal range, and an increase in the number of students within the normal range occurred. By contrast, little change was evident in the control group.

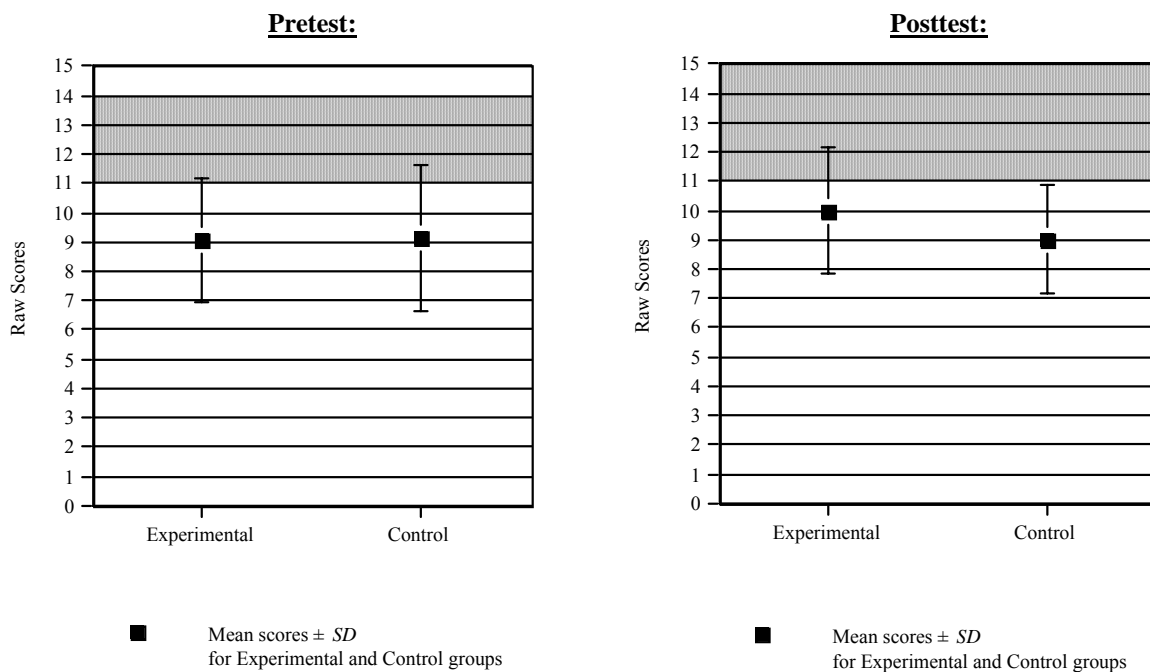


Figure 27. Mean scores for Digit Span (Level A)

Digit Span (Level B)

Figure 28 shows the effects of the program on Digit Span in relation to the WISC-111 norms. The effect was moderate but elevated the experimental group mean very close to the normal range, and an increase in the number of students within the normal range occurred. By contrast, a smaller change was evident in the control group.

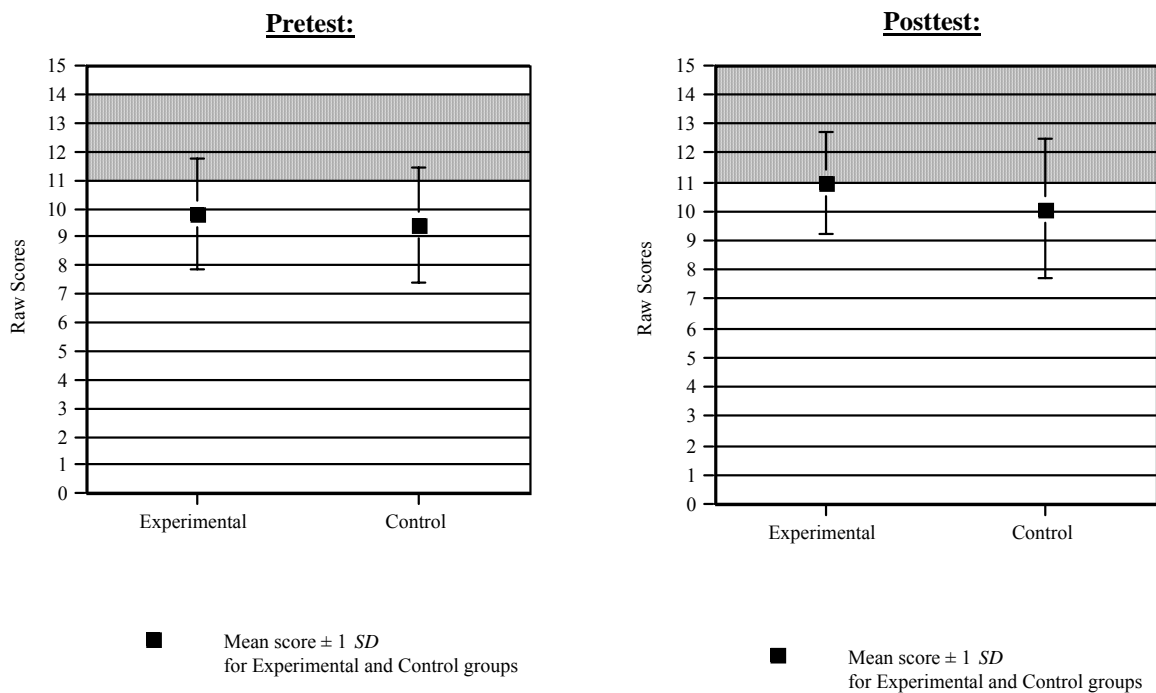


Figure 28. Mean scores for Digit Span (Level B)

Brigance Spelling (Level A)

Figure 29 displays the results for the Brigance Spelling test, and follows a similar pattern to the other tests. A feature of this graph is the extent to which all students were below the norm for this test, even after the program. The effects were strong; however, the students clearly remained in need of assistance.

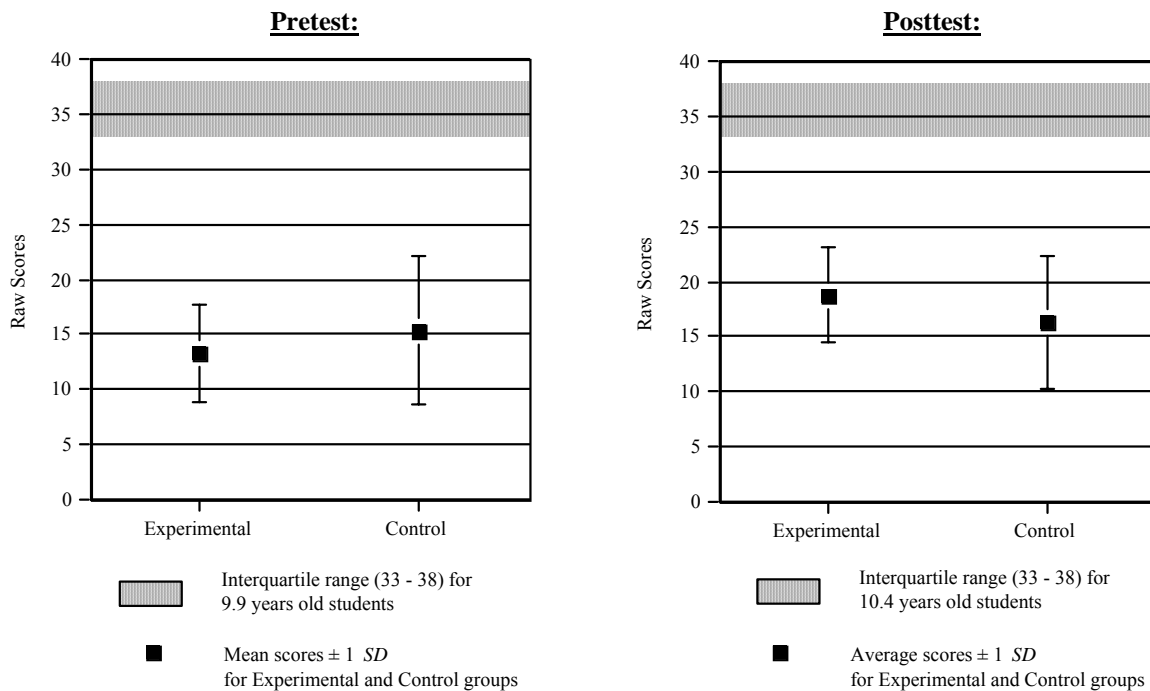


Figure 29. Mean scores for Brigance Spelling (Level A)

Brigance Spelling (Level B)

Figure 30 displays the results for the Brigance Spelling test, and follows a similar pattern to the other tests. A feature of this graph is the extent to which all students were below the norm for this test, even after the program. The effects were strong; however, the students clearly remained in need of assistance.

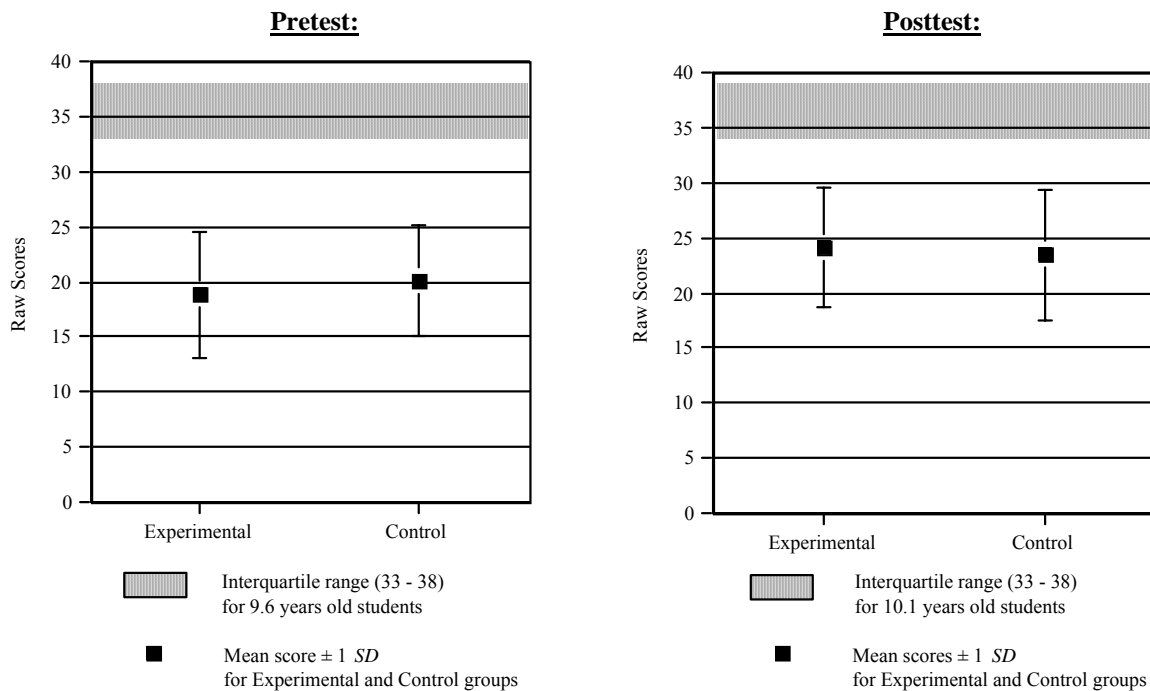


Figure 30. Mean scores for Brigance Spelling (Level B)

What is the Relationship Between the Measured Phonological Variables?

Correlations

As discussed in Chapter 4, there has been considerable interest in the makeup of phonological processes, and their relationships. In Table 23, the correlations between the pretest measures are displayed.

Table 23

Correlations Between Pretest scores (N = 206)

	1.	2.	3.	4.	5.
1. TOPA	-				
2. Word Attack	.4482 P<.001	-			
3. PNT	.2780 P<.001	.0952 P=.174	-		
4. Digit Span	.3432 P<.001	.2259 P=.001	.2826 P<.001	-	
5. Spelling	.3892 P<.001	.5251 P<.001	.1454 P=.037	.2192 P=.002	-

Note. Coefficient / 2-tailed Significance

". ." is printed if a coefficient cannot be computed

In comparison to other studies, as described in Chapter 4, the correlations reported here are generally lower. This is unsurprising given that only poor readers were represented - a restricted range usually under-estimates the strength of correlations. The strongest correlation was between Spelling and Word Attack. (In their 1996 study, Shankweiler, Lundquist, Dreyer, and Dickinson also found that Word Attack at pretest was the best predictor of spelling at pretest). The weakest correlation occurred between Picture Naming and Word Attack.

Regression Analyses

In the following tables, regression analyses provide a similar picture to that above. All the variables (Table 24) were significant predictors of TOPA pretest scores, though the squared partial correlations indicate only small unique contributions of each variable (Word Attack being the strongest).

Table 24

Summary of Simultaneous Regression Analysis for Pretest Variables Predicting TOPA

Pretest Score

Variable	<i>B</i>	\square	<i>sr2</i>	<i>p</i>
Word Attack	4.27	.34	.081	.000
PNT	0.10	.19	.033	.002
Digit Span	-19.17	-.15	.019	.021
Spelling	0.45	.14	.014	.044

In Table 25, only TOPA and Spelling made significant contributions towards the prediction of Word Attack, with Spelling the major contributor.

Table 25

Summary of Simultaneous Regression Analysis for Pretest Variables Predicting Word

Attack Pretest Score

Variable	<i>B</i>	\square	<i>sr2</i>	<i>p</i>
TOPA	0.025	.309	.074	.000
PNT	-0.002	-.061	.003	.307
Digit Span	-0.270	-.026	.001	.669
Spelling	0.105	.419	.147	.000

In Table 26, only TOPA and Word Attack made significant contributions toward the prediction of Spelling, with Word Attack providing the strongest unique contribution.

Table 26

Summary of Simultaneous Regression Analysis for Pretest Variables Predicting

Spelling Pretest Score

Variable	<i>B</i>	\square	<i>sr2</i>	<i>p</i>
TOPA	0.044	.140	.014	.044
Word Attack	1.801	.450	.158	.000
PNT	0.006	.039	.001	.530
Digit Span	-3.634	-.087	.007	.165

Principal Components Analyses

A series of exploratory factor analyses were carried out to investigate the nature and type of latent variables underlying the five main dependent measures. All analyses used principal components factor extraction followed by varimax rotation. The first analysis, which used the SPSS 6.1 (1995) default factor extraction procedure based on eigenvalues greater than 1, resulted in a two factor solution that accounted for 65.5% of variability. The rotated factor solution results are shown in Table 27. It was apparent that a Word

Attack/TOPA/Spelling factor was distinct from a working memory/naming factor (apart from a minor role for phonemic awareness in the second factor).

Table 27

Varimax Rotated Two Factor Solution for Total Sample at Pretest

	Factor 1	Factor 2
Word Attack	.86032	.04937
Spelling	.81769	.08500
TOPA	.62284	.45628
PNT	-.02003	.84027
Digit Span	.23287	.70538

For the second analysis, three factors were forced into the final solution, which accounted for 79.5% of total variability. The rotated factor solution results for this analysis are shown in Table 28. The factors were a reading/spelling/phonemic awareness factor, a working memory factor, and a naming factor. In this solution, working memory and naming were clearly delineated.

Table 28

Varimax Rotated Three Factor Solution for Total Sample at Pretest

	Factor 1	Factor 2	Factor 3
Word Attack	.85415	.11163	-.03669
Spelling	.83302	.02985	.07546
TOPA	.62079	.35180	.29513
Digit Span	.13352	.96467	.12575
PNT	.07025	.12978	.97476

Another set of exploratory factor analyses was performed separately on the experimental and control groups at pretest and posttest to investigate any changes in the latent variables underlying the five main dependent measures, consequential upon the intervention. All analyses used principal components factor extraction followed by varimax rotation. For the control group, three factors were forced into the final solution, which accounted for 81% at pretest and 80.4% at posttest of total variability. The rotated factor solution results are shown in Tables 29 and 30. Examination of the posttest factors for the control group indicated

little change in scores apart from TOPA in Factor 2, and similarly little alteration of the structure of the factors.

Control Group Alone

Table 29

Rotated Three Factor Solution for Control Group at Pretest

	Factor 1	Factor 2	Factor 3
Word Attack	0.86212	0.16390	-0.06830
Spelling	0.87734	0.08060	0.16461
TOPA	0.50519	0.50603	0.26704
Digit Span	0.10589	0.94395	0.13635
PNT	0.06722	0.18136	0.96588

Table 30

Rotated Three Factor Solution for Control Group at Posttest

	Factor 1	Factor 2	Factor 3
Word Attack	.79357	.24084	-.02114
TOPA	.78133	-.08919	.30525
Spelling	.73144	.48721	-.01829
Digit Span	.15084	.92369	.16214
PNT	.09745	.14669	.95661

Experimental Group Alone

Three factors were forced into the final solution for the experimental group, which accounted for 78.2% at pretest and 76.7% at posttest of total variability. The rotated factor solution results are shown in Tables 31 and 32.

Table 31

Varimax Rotated Three Factor Solution for Experimental Group at Pretest

	Factor 1	Factor 2	Factor 3
Word Attack	0.86175	0.03460	0.06002
Spelling	0.78342	-0.02902	0.06984
TOPA	0.63409	0.35161	0.28996
Digit Span	0.13639	0.11390	0.97517
PNT	0.04469	0.96934	0.10366

Table 32:

Varimax Rotated Three Factor Solution for Experimental Group at Posttest

	Factor 1	Factor 2	Factor 3
Word Attack	.87621	.16955	.04241
Spelling	.62153	-.05849	-.05849
Digit Span	-.08930	.80207	.35317
TOPA	.44291	.76161	-.10788
PNT	.05895	.18097	.85086

Examination of the posttest factors for the experimental group revealed greater change among the individual loadings, and a different factor structure. The phonemic awareness

score (TOPA) now loaded on Factor 2 rather than Factor 1, and the PNT at posttest now had a considerably higher loading on Factor 3. An implication of these results was that an alteration in the structure of the phonological processes followed the reading program.

Is Success in the *Corrective Reading Program* Predicted by Any of the Pretest Scores?

In addition to investigating the relationship among the phonological processes, another issue of interest was the potential of pretest scores to predict which students would make good progress, and which would not. The tables below report a series of regression analyses in which gain scores formed the dependent variables, and a number of variables (transformed when appropriate) were chosen as potential predictors.

Regression Analyses

In order to evaluate the contribution of overall pretest performance as predictors of gain, a multivariate multiple regression model was tested on the experimental group ($n = 134$). The dependent variable in this model comprised the five gain scores, and the predictor comprised the combined five pretest scores. The combined pretest scores were found to be a significant predictor of the combined gain scores, Wilks' $\lambda = .25$, $F(25, 462.14) = 8.32$, $p < .001$.

The next step was to examine the effect of this combined predictor score (or vector) on each of the dependent variables in turn. Table 33 indicates the predictive capacity on each gain score in turn of a vector comprising the pooled pretest scores. Each variable's gain was predicted significantly by this vector.

Table 33

Summary of Multivariate Multiple Regression for Pooled Pretest Scores Predicting Gains in the Experimental Group ($n = 134$)

Dependent Variable (Gain)	R^2	Adj. R^2	F	p
TOPA	.36	.34	14.70	.000
Word Attack	.10	.07	2.92	.016
PNT	.22	.19	7.12	.000
Digit Span	.22	.19	7.35	.000
Spelling	.26	.23	8.82	.000

Table 34 provides additional detail in that the relative contributions to the vector of each pretest variable were included. It is not surprising that, in general, the relevant variable's

pretest score provided the greatest contribution to the predictive capacity of the vector; whereas, few other scores reached significance. Even those additional variables that did reach significance were not at all strong in their predictive quality. In the case of TOPA, the variable with the strongest effect size, the relationship is negative, that is, high initial scores were predictive of lesser gains. It is likely that ceiling effects in the test provide the best account for this effect. Interestingly, no predictor reached significance for Word Attack, the variable in which Program provided the strongest effect.

Table 34

Details of Multivariate Multiple Regression for Pooled Pretest Scores Predicting Gains in the Experimental Group (n = 134)

Dependent Variable (Gain)	Predictors	<i>B</i>	β	<i>p</i>	β^2
TOPA	TOPA	- 0.56	-0.60	.000	.281
	PNT	- 0.04	-0.08	.301	.008
	Word Attack	0.85	0.08	.364	.006
	Digit Span	-17.48	-0.16	.038	.033
	Spelling	- 0.35	-0.12	.136	.017
Word Attack	TOPA	0.32	0.19	.057	.028
	PNT	- 0.11	-0.13	.146	.016
	Word Attack	- 3.09	-0.16	.127	.018
	Digit Span	24.82	0.13	.167	.015
	Spelling	- 0.80	-0.16	.108	.020
PNT	TOPA	0.13	0.07	.476	.004
	PNT	- 0.36	-0.38	.000	.143
	Word Attack	4.91	0.23	.020	.041
	Digit Span	7.09	0.03	.703	.001
	Spelling	0.53	0.09	.303	.008
Digit Span	TOPA	0.11	0.22	.020	.041
	PNT	- 0.01	-0.04	.648	.002
	Word Attack	0.19	0.04	.718	.001
	Digit Span	26.55	0.46	.000	.194
	Spelling	- 0.06	-0.04	.666	.001
Spelling	TOPA	0.16	0.14	.126	.018
	PNT	- 0.02	-0.03	.746	.001
	Word Attack	3.70	0.29	.004	.064
	Digit Span	- 4.46	-0.03	.688	.001
	Spelling	- 1.98	-0.57	.000	.243

Table 35 displays the contributions of each pretest variable in accounting for the gains in each variable consequent upon the program. In general, only the pretest variable made a reasonable contribution; including the other variables that reached significance added little predictive power.

Table 35

Summary of Simple Linear Regression for Each Pretest Score Separately Predicting Each Gain for Experimental Group (n = 134)

Predictors	Dependent Variable (Gain)	R^2	Adj. R^2	sig F	B	\square	p
TOPA	TOPA	0.333	0.328	.000	-0.535	-0.577	.000
	Word Attack	0.000	0.000	.979	-0.004	-0.002	.979
	PNT	0.010	0.003	.242	0.189	0.102	.242
	Digit Span	0.009	0.001	.284	0.045	0.093	.284
	Spelling	0.013	0.005	.195	0.128	0.113	.195
Word Attack	TOPA	0.067	0.060	.002	-2.736	-0.260	.002
	Word Attack	0.031	0.024	.042	-3.345	-0.176	.042
	PNT	0.071	0.064	.002	5.648	0.267	.002
	Digit Span	0.002	0.000	.594	0.254	0.046	.594
	Spelling	0.007	0.000	.339	1.078	0.083	.339
PNT	TOPA	0.040	0.033	.021	-0.094	-0.200	.021
	Word Attack	0.020	0.013	.102	-0.121	-0.142	.102
	PNT	0.116	0.109	.000	-0.323	-0.340	.000
	Digit Span	0.010	0.002	.260	-0.024	-0.098	.260
	Spelling	0.000	0.000	.949	-0.003	-0.006	.949
Digit Span	TOPA	0.001	0.000	.674	4.035	0.037	.674
	Word Attack	0.029	0.022	.049	33.828	0.171	.049
	PNT	0.003	0.000	.555	11.383	0.052	.555
	Digit Span	0.175	0.169	.000	23.941	0.419	.000
	Spelling	0.001	0.000	.787	3.183	0.024	.787
Spelling	TOPA	0.056	0.049	.006	-0.669	-0.237	.006
	Word Attack	0.048	0.041	.011	-1.119	-0.219	.011
	PNT	0.033	0.026	.035	1.039	0.183	.035
	Digit Span	0.005	0.000	.434	-0.100	-0.068	.434
	Spelling	0.146	0.139	.000	-1.327	-0.382	.000

In Table 36, the capacity of Program and the pretest scores to predict the Word Attack posttest scores is examined across the total sample. It is apparent that initial scores were strongly related to outcome scores (this is especially so for the control group); however, Program, whilst affecting only 134 students of the sample of 206), was also a very strong predictor.

Table 36

Summary of Stepwise Regression Analysis for Program and Pretest Scores Predicting Word Attack Posttest scores (N = 206)

Step	Variable	<i>R</i>	<i>R</i> ²	ΔR^2	<i>sig F</i>	Δ	<i>p</i>
1	Word Attack	.614	.376	.376	.694	.614	.000
2	Spelling	.629	.396	.020	-.077	.700	.000
3	Program	.767	.588	.192	.446	.694	.000
Constant in the final equation							.040
Variables not in the final equation							
	TOPA						.060
	PNT						.346
	Digit Span						.506

In prediction of gains in Word Attack for the experimental group, Table 37 indicates that Program membership was by far the strongest, whilst Word Attack and Spelling pretest scores were significant predictors, their combined contribution is less than 7% - small in comparison with that of Program (almost 30%).

Table 37

Summary of Stepwise Regression Analysis for Program and Pretest Scores Predicting Word Attack Gains Scores (N = 206)

Step	Variable	<i>R</i>	<i>R</i> ²	ΔR^2	<i>sig F</i>	Δ	<i>p</i>
1	Word Attack	.194	.038	.038	.005	-.094	.005
2	Spelling	.261	.068	.031	.001	-.096	.011
3	Program	.604	.365	.296	.000	.555	.000
Constant in the final equation							.040
Variables not in the final equation							
	TOPA						.060
	PNT						.346
	Digit Span						.506

CHAPTER NINE: ADDITIONAL STUDY

Introduction

Having noted the improvement in phonemic awareness and phonological recoding effected through the use of the Direct Instruction program, (*Corrective Reading*) with older remedial readers, interest arose in examining the effects of a Direct Instruction program specifically designed for beginning readers: *Teach Your Child to Read in 100 Easy Lessons* (Engelmann, Haddox, & Bruner, 1983). It is conventional wisdom that the earlier reading problems are addressed, the greater the likelihood of satisfactory and speedy resolution.

It was evident in the major study that significant improvement was possible in older poor readers over a period of months. However, many of these children were several years delayed in comparison with their age peers, and may require several years of additional support if they are to match their reading facile colleagues. Many of these children have experienced the debilitating sequence of interacting skill deficits described by Stanovich (1986) as the Matthew effect. For example, the early lack of phonemic awareness leads to a failure to master the alphabetic principle. This further entails slow, error-prone decoding, the overuse of contextual cues, and poor comprehension. This resultant laborious, unsatisfying reading style leads students to avoid text, with a consequential reduction in vocabulary growth, and a broadening of the skill deficit. The lack of practice means fewer words can be read by sight, thereby restricting automaticity. The continued expenditure of cognitive attention on decoding leaves few resources available for comprehension, and so the student's difficulties are compounded. The longer this set of circumstances prevails, the further delayed the student becomes, the more pervasive becomes the problem, and the more difficult the rescue operation. Hence, the concern for intervening earlier in this escalating chain.

If the operation is commenced earlier, when the primary deficit is restricted to phonemic awareness, and it is this deficit that is targeted, it is reasonable to anticipate a more efficacious process. If increased phonemic awareness and an early understanding of the alphabetic principle are the outcomes (thus precluding the by-products of early reading failure), the intervention at this stage should be more effective, efficient and socially just.

Although the content of the *Teach Your Child to Read in 100 Easy Lessons* (Engelmann, Haddox, & Bruner, 1983) was developed earlier than most of the research into phonemic awareness, it is now becoming more evident that the combination of letter-sound

instruction with phonemic awareness training (as evidenced in the *100 Lessons* program) is a potent one in stimulating early reading development (Byrne & Fielding-Barnsley, 1991, 1993, 1995; Ehri, 1987; Hatcher et al., 1994; Perfetti et al., 1987, Torgesen et al., 1994).

However, a wide range of phonemic awareness tasks have been incorporated into phonemic awareness programs, and a vital question (especially for at-risk students) is what combination of tasks is optimally related to accelerated reading development?

Torgesen, Morgan, and Davis (1992) tested two types of phonemic awareness training approaches - blending only, and a combination of blending and segmenting - and compared them to a language experience control group. The small groups trained three times per week for 20 minutes for a total of 7-8 weeks. The blending only group improved only on blending, their segmentation skills remaining similar to that of the controls. Similarly, their ability to learn in a reading analogue task did not significantly exceed that of the control group, indicating a lack of generalisation of this skill to this reading task. In contrast, the combination of blending and segmenting led to significant improvements in both skills, and evidence of transfer to the reading task. The authors acknowledge that the introduction of letter-sound training may have even further enhanced the transfer to reading tasks had they incorporated such strategy.

Davidson and Jenkins (1994) in a similar study included a segmentation-only training group, and while they noted some transfer to a reading analogue task for that group, they too argued against teaching only one type of phonemic awareness strategy, as generalisation of awareness is likely to be compromised.

O'Connor, Slocum, and Jenkins (1995) reported a study in which the combination of letter-sounds, blending and segmenting instruction led to educationally significant gains for at-risk beginning readers. The program intervention lasted a total of five hours (15 minutes twice weekly for 10 weeks). A second experimental group had a much greater range of phonemic awareness activities (in addition to segmentation and blending) but showed no increase in reading development over the first experimental group. The authors argue that both experimental groups were able to generalise the phonemic awareness skills they were taught, that is, they attained phonological insight, and were able to relate it to the reading process. Importantly, their findings suggest that the combination of blending and segmenting is sufficient to create this condition.

Lovett et al. (1994) used a 35 lesson training program developed from *Reading Mastery*, and *Corrective Reading* to teach word identification to dyslexic students for one

hour four times per week. They compared results to a control group taught a study skills program, and achieved highly significant posttest gains for the experimental group - effect sizes (d) of 0.76, 1.11, and 0.90 on the three training measures. The transfer to real words was impressive, and "was based on the successful training of what is considered the core deficit of developmental dyslexia: phonological processing and nonword reading skill" (p. 818). Further, they argue, "this training success rests on embedding letter-sound training in an intensive phonological training program" (p. 819).

Thus, there is evidence to support the use of a program that explicitly teaches letter-sound correspondence, and which simultaneously links this knowledge to two critical phonemic awareness skills, blending and segmenting. This should not surprise since segmenting and blending are the phonemic awareness processes most closely involved in reading, and letter-sound knowledge is both a prompt, and a necessary condition for this phonemic awareness knowledge to be useful in reading. The *100 Lessons* program meets these dual requirements of theoretically and empirically validated practice.

Method

The Participants

The participants were 13 students (eight boys, and five girls) from a number of northern suburbs primary schools who were attending a reading intervention unit for four one-hour sessions per week. All had been referred to the centre by their parents, or teachers, as being at-risk in their reading development. Their average age was eight years and seven months. A non-equivalent control group (eight boys and five girls, average age eight years and six months) was drawn from the wait list group used in the major study. This group was selected to match the sex ratio and average age of the experimental group, but their reading skills were assessed as above those in the *100 Lessons* group. The program was implemented by a qualified teacher who had had two years experience with Direct Instruction programs in that setting. Its duration was 7.7 months, while the control group pretest-posttest period was 6.6 months.

The Program

Teach Your Child to Read in 100 Easy Lessons (Engelmann, Haddox, & Bruner, 1983) was developed as a program for parents to use with preschool or struggling readers. It was from a school-based reading program. *Reading Mastery* Fast Cycle I/II Reading Program (Engelmann & Bruner, 1988).

The program follows the Direct Instruction principles of design, and the content emphasises the explicit teaching of phonemic awareness (rhyming, blending, segmenting) along with 44 letter sound correspondences. These selected correspondences allow for the decoding of 95% of the sounds in the students' typically available reading texts, and close approximations for 98% (Burmeister, 1975, cited in Grossen, 1995).

A specially developed orthography reduces the number of such correspondences to an attainable number (some programs had taught up to 200 such correspondences) and allows for the introduction (Lesson 13) of interesting sentences while still controlling the text for regularity (albeit artificially). This Distar orthography (Table 38) enables a range of interesting irregular words to be decoded using the segment/blend strategy, thus providing for students both practice and a developing assurance that the strategy is a successful one, worth persevering with until familiarity produces whole word recognition. This feature is very important as students can be overwhelmed by the number of irregular words in uncontrolled text - the result being an inability to appreciate the value of the recoding strategy, and a consequent failure to focus on developing the skill.

The orthography has several useful features that enable a variety of text, avoiding the "Nan can fan Dan" limitation of devising regular text when few sounds are known by students. Visual cues are provided to promote the distinction between long and short vowels, through the use of a macron over the relevant long vowel (See Table 38). Words with CVCe (consonant-vowel-consonant-e) sequences are regularised through the use of small letters that are not pronounced. Hence *lake* is written as *lake* (with macron), and can be decoded by blending the three sounds. The teaching of separate sounds for two letter blends, such as *er*, *wh*, *sh*, *th*, *ch*, and *qu*, similarly allows for the regularisation of troublesome words such as *she* and *where*. Not all words are made regular, as that would teach a misrule - that all words are regular in natural text - thus a few words are allowed to continue as irregulars (e.g., *to*, *was*, *said*). Learning such a misrule would make the subsequent transition to normal orthography difficult for students. The intention is to teach sufficient words in this manner to ensure that students are aware of exceptions, but not so many that the utility of mastering phonological recoding is jeopardised.

Table 38

*Distar Orthography***Pronunciation Guide**

Symbol	Pronounced	As in	Voiced or Unvoiced*	Introduced in Lesson
a	aaa	<u>and</u>	v	1
m	mmm	<u>ram</u>	v	1
s	sss	<u>bus</u>	uv	3
e	êêê	<u>eat</u>	v	4
r	rrr	<u>bar</u>	v	6
d	d	<u>mad</u>	v	9
f	fff	<u>stuff</u>	uv	11
i	iii	<u>if</u>	v	13
th	ththth	<u>this</u> and <u>bathe</u> (not thing)	v	15
t	t	<u>cat</u>	uv	17
n	nnn	<u>pan</u>	v	19
c	c	<u>tack</u>	uv	21
o	ooo	<u>ox</u>	v	23
ā	āāā	<u>ate</u>	v	26
h	h	<u>hat</u>	uv	28
u	uuu	<u>under</u>	v	30
g	g	<u>tag</u>	v	33
l	lll	<u>pal</u>	v	35
w	www	<u>wow</u>	v	37
sh	shshsh	<u>wish</u>	uv	39
I	(the word I)		v	43
k	k	<u>tack</u>	uv	45
ō	ōōō	<u>over</u>	v	48
v	vuv	<u>love</u>	v	50
p	p	<u>sap</u>	uv	54
ch	ch	<u>touch</u>	uv	57
e	eee	<u>end</u>	v	59
b	b	<u>grab</u>	v	61
ing	iiing	<u>sing</u>	v	62
I	īīī	<u>ice</u>	v	64
y	yyy	<u>yard</u>	v	66
er	urrr	<u>brother</u>	v	67
x	ksss	<u>ox</u>	uv	70
oo	oooooo	<u>moon</u> (not look)	v	71
J	j	<u>judge</u>	v	72
ȳ	īīī	<u>my</u>	v	75
wh	www or wh	<u>why</u>	v or uv	76
qu	kwww (or koo)	<u>qu^lck</u>	v	77
z	zzz	<u>buzz</u>	v	78
ū	ūūū	<u>use</u>	v	80

Reading Mastery Fast Cycle Teacher's Guide (Engelmann & Bruner, 1984)

The correspondences are introduced in a sequence different to that in the alphabet, to reduce the ambiguity associated with similar shapes or sounds being introduced at nearly the same time. For example, /d/ is introduced in Lesson 12, whereas /b/ is taught in Lesson 54. An additional distinguishing prompt sees the "ball" on the /d/ assigned a stretched (almost elliptical) shape (until Lesson 74) to separate it further from its mirror image /b/. This pair of letters often presents problems of interference (reversals) to young readers, who are sometimes accused of neurological deficits to account for a largely instructional problem.

Another rationale for the atypical sequence of letter introduction is to enlarge the range of words that can be created from the earliest stages of the program.

Words are first introduced in Lesson 3, and considerable attention is paid to oral reading practice with immediate corrective feedback. Research support for the Distar programs (later revamped as the *Reading Mastery* series) has been strong. See Chapter 2 for Follow Through results, and recently a meta analysis by Adams and Englemann (1996) has reported an effect size (*d*) of 0.68 for 44 acceptable comparisons involving *Reading Mastery* and other beginning reading programs.

The *100 Lessons* is very carefully constructed. Apart from the controlled vocabulary, the program prescribes the tasks to be presented, the examples chosen, and how often they occur. Even the teacher's wording is specified through the use of a script. This high level of control is based on the principles of faultless communication discussed in Chapter 5

The program emphasises letter sounds rather than letter names because of the functionality of the former in beginning reading, and to avoid the opportunity for unnecessary confusion entailed by teaching both sounds and names simultaneously. Names are introduced in Lesson 73, and capital letters in Lesson 81. The phonemic awareness skills of blending and segmenting are taught orally initially, because there are fewer elements in the oral than the written task, and hence less likelihood of error. Blending is taught as a simultaneous rather than discrete-sound format - “mmmaa” rather than “mmm-aaa-t” because the stimulus sequence of sounds is really a stretched form of the word “mat”, rather than a broken form in which the elements are completely separated. The authors argue that the mastery of continuous blending is a worthwhile objective because it provides more salient clues to the pronunciation of words. The oral blending activities proceed from large intra-word clusters to single phoneme blends.

“Let’s play say-it-fast.

My turn: *motor* (pause) *boat*.

(Pause) Say it fast. “*Motorboat*”.

From: *Teach Your Child to Read in 100 Easy Lessons* (Engelmann, Haddox, & Bruner, 1983, p.31).

To assist the mastery of simple two phoneme blends an additional step is included in the model-test sequence. The sequence becomes model-lead-test, thus providing an additional prompt.

“First I’ll say *am* slowly. Listen: *aaamm*.”

Now it's your turn to say the word slowly with me. Take a deep breath and we'll say *aaammm*. Get ready. *aaammm*

Your turn to say the word slowly by yourself. Say *aaammm*. Get ready. "Aaammm."

From: *Teach Your Child to Read in 100 Easy Lessons* (Engelmann, Haddox & Bruner, 1983, p.31).

Blending activities begin in the first lesson, and segmenting written words into constituent phonemes in Lesson 9. This latter process is assisted by the use of marks under the word that prompt the sounds one by one at the required pace. See Table 39 for an example of a blending sequence.

Table 39

Script for Blending Activity

Task 9 WORD READING

1. (Point to sat.) You're going to touch under the sounds as you sound out this word and say it fast. (Touch under s.) What's the first sound you're going to say? "sss." (Touch under a.) What's the next sound you're going to say? "aaa." (Touch under t.) What's the next sound you're going to say? "t."

2. Touch the first ball of the arrow. Take a deep breath and say the sounds as you touch under them. Get ready. Go. (Child touches under s, a, and t and says "sssaaat." (Repeat until firm.)

3. Say it fast. "sat." Yes, what word? "sat." You read the word *sat*. Good reading.



From: *Teach Your Child to Read in 100 Easy Lessons* (Engelmann, Haddox & Bruner, 1983, p. 53).

Other activities include: rhyming to promote a sensitivity to word families based on common endings (or rimes); *sounds-writing* because it prompts attention to the letter shape, and helps forge the association between shape and sound; *story reading* (from Lesson 13)

involving successive segmenting and blending; and, *picture comprehension* (from Lesson 13). Pictures are provided after the story is finished to assist comprehension, but to avoid the picture cues being used in place of print cues in the decoding task. *Sight words* (from Lesson 13). Words that have been practised sufficiently often (using the segment-blend procedure) for them to begin to become familiar are “read the fast way”, that is, the child slides his finger under the letters to prompt a thorough viewing, but does not sound out the word, rather he reads it orthographically.

Supporting this cumulative skill acquisition and skill synthesis model are clear scripted correction procedures. There are two basic principles - the first is that correction be applied immediately following the error, rather than delayed until the end of a sentence, or waiting for self-correction. The purpose of the program is to teach accurate decoding of words based on information provided by the print, rather than relying on contextual cues to prompt a word's pronunciation. Hence, the correction redirects the child's attention to the source of the information - the word. The second principle specifies the basic correction structure - the child is notified of the error, given the correct response, allowed to practise this response, and finally tested on the original task before moving on. Additionally, a delayed test presented later in the lesson is often recommended.

The change from Distar orthography to normal print occurs over a three lesson period (Lessons 74-76), and after that time all print is conventional. By this stage the child is reading stories of about 200 words orthographically, and answering comprehension questions. According to the program designers the child should be reading at around a Year Two level at the completion of the program. The shift from letter by letter decoding to orthographic whole word recognition occurs in students who are able to analyse fully the structure of words (Stanovich, 1991), and have had many opportunities for practice of such words in isolation, and in connected text - particularly with words containing high frequency spelling patterns (Ehri, 1992).

According to Ehri's (1992) work, the most effective way for beginning readers to store sight words in memory is to analyse fully the sounds in the spoken word and to match those sounds to the letters in the printed form of words. To do this, readers must know how to segment pronunciations of words into their smallest sounds, and they must know which letters typically symbolise those sounds. (Gaskins, Ehri, Cress, O'Hara, & Donnelly, 1996, p. 315)

Given that the content of the *100 Lessons* program focuses on the skill areas currently accepted as critical, that the style of teaching employs empirically supported effective-teaching principles, and that the instructional design principles ensure ample massed and spaced practice - it is reasonable to anticipate that the authors' claims of decoding instruction leads to eventual skilled whole-word recognition are reasonable claims.

The selection of the parent-based program over the *Reading Mastery* series was based on cost. Few schools are prepared to invest the relatively large sum of money in a program for a few at-risk beginning readers. The *Reading Mastery* series was written as a basal series designed for general classroom beginning reading instruction, but is not generally attractive to schools for that purpose. The *100 Lessons* program, however, is cheap and in the author's experience, effective if presented faithfully, either by parent or teacher. As the program is designed for one-to-one teaching, there are some modifications required for group instruction. As the teacher involved was skilled in presenting the *Corrective Reading* program, it was not difficult to incorporate the group-signalling, correction, and choral/individual turn-taking strategies from one program to the other. The most evident changes involve: the use of the blackboard to reproduce the graphics presented in the book; using the finger-slide signal at the board rather than on the page; providing roneo sheets containing the words and sounds for that lesson to allow the students to use the finger-slide prompt; and, using a hand-drop signal for the orally-presented tasks to ensure simultaneous choral responding.

Results

Table 40

Experimental vs Control groups: Mean Raw Scores for 100 Lessons

	<i>n</i>	Age	TOPA	Word Attack	PNT	Digit Span	Spelling
Control	13						
Pretest							
Range		97-113	5-19	1-17	17-47	6-12	5-30
<i>M</i>		104.31	11.77	10.15	30.00	8.69	15.62
<i>SD</i>		5.62	4.71	4.10	10.26	1.89	7.12
Posttest							
Range		103-119	6-16	3-21	20-52	7-12	13-31
<i>M</i>		111.62	11.54	11.62	32.77	8.92	18.77
<i>SD</i>		5.52	3.71	4.81	8.82	1.66	4.97
Experimental 13							
Pretest							
Range		83-121	3-17	0-9	19-39	3-10	0-12
<i>M</i>		103.23	9.23	2.00	28.46	6.15	5.54
<i>SD</i>		13.38	4.48	2.80	6.44	2.19	3.60
Posttest							
Range		93-128	7-20	0-29	23-46	5-12	0-24
<i>M</i>		111.00	14.38	12.38	33.38	8.92	11.62
<i>SD</i>		13.06	4.56	8.61	7.07	1.80	6.95

Table 41
Experimental vs Control Groups: Mean Power Transformed Scores for 100 Lessons

	<i>n</i>	TOPA	Word Attack	Spelling
Control	13			
Pretest				
Minimum		0.78	1.00	2.24
Maximum		1.30	4.12	5.48
<i>M</i>		1.08	3.10	3.86
<i>SD</i>		0.17	0.77	0.89
Posttest				
Minimum		0.85	1.73	3.61
Maximum		1.23	4.58	5.57
<i>M</i>		1.08	3.33	4.30
<i>SD</i>		0.13	0.77	0.56
Experimental	13			
Pretest				
Minimum		0.60	0.00	0.00
Maximum		1.26	3.00	3.46
<i>M</i>		0.97	1.01	2.12
<i>SD</i>		0.20	1.03	1.06
Posttest				
Minimum		0.90	0.00	0.00
Maximum		1.32	5.39	4.90
<i>M</i>		1.17	3.24	3.17
<i>SD</i>		0.14	1.42	1.30

Assumptions of normality and homogeneity of variance were tested for all data used in ANCOVA and ANOVA analyses, and data transformations were performed when necessary, as shown in Table 42. Square root transformations were used for experimental and control group, pretest and posttest data for Word Attack and Spelling. Transformations were unnecessary for Picture Naming and Digit Span, and Log transformation were beneficial for TOPA posttest scores but unhelpful for TOPA pretest scores (so the more stringent requirement for significance $\alpha = .01$ was adopted).

Table 42

Tests of Normality and Homogeneity of Variance for 100 Lessons

Variable			Trans- formation	Lilliefors' test of Normality	Levene's test of Homogeneity of Variance
Word Attack	Pretest	Control	Square root	> .2	.24
		Experimental	Square root	.08	
	Posttest	Control	Square root	> .2	.09
		Experimental	Square root	> .2	
Digit Span	Pretest	Control	No	.14	.51
		Experimental	No	> .2	
	Posttest	Control	No	> .2	1.00
		Experimental	No	> .2	
Spelling	Pretest	Control	Square root	> .2	.80
		Experimental	Square root	> .2	
	Posttest	Control	Square root	> .2	.06
		Experimental	Square root	> .2	
PNT	Pretest	Control	No	> .2	.05
		Experimental	No	> .2	
	Posttest	Control	No	.07	.51
		Experimental	No	> .2	
TOPA	Pretest	Control	Log 10	> .2	.02
		Experimental	Log 10	> .2	
	Posttest	Control	Log 10	> .2	.61
		Experimental	Log 10	> .2	

TOPA

Results for TOPA were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 22) = 1.27, p = .273$. With the pretest results partialled out there was a significant overall difference between the experimental and control groups, $F(1, 23) = 53.90, p < .001$.

Results for log transformed scores for TOPA were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 24) = 0.03, p = .865, \text{power} = .04$, but a significant main effect was found for time, $F(1, 24) = 42.80, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 24) = 40.41, p < .001, \text{power} = 1.00$, which is illustrated in Figure 31. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 24) = 2.18, p = .153, \text{power} = .29, d = -0.63$, and at posttest, $F(1, 24) = 2.52, p = .125, \text{power} = .33, d = 0.78$. Further, no significant pre- to posttest difference was found for the control, $F(1, 24) = 0.02, p = .897, \text{power} = .04, d = -0.06$, but a significant pre- to posttest difference was found for the experimental groups, $F(1,$

24) = 83.20, $p < .001$, power = 1.00, $d = 1.30$, and the magnitude of effect was large for the experimental group.

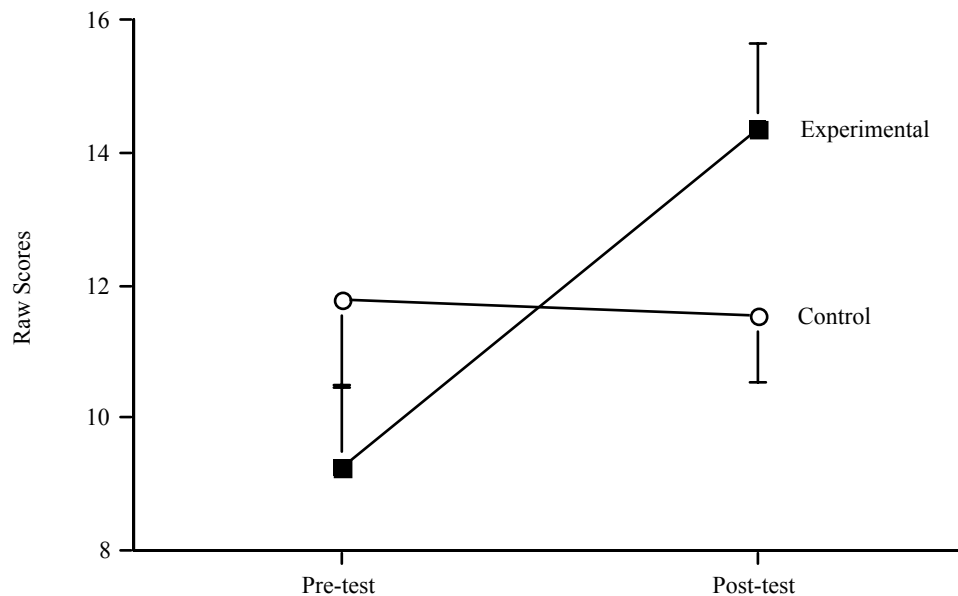


Figure 31. Interaction (+ SE) between experimental and control group at pre and posttest for TOPA.

Word Attack

Results for Word Attack were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 22) = .01$, $p = .941$. With the pretest results partialled out there was no significant overall difference between the experimental and control groups, $F(1, 23) = 2.46$, $p = .130$.

Results for the square root transformed Word Attack scores were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, $F(1, 24) = 10.19$, $p = .004$, power = .86, and for time, $F(1, 24) = 31.22$, $p < .001$, power = 1.00, and the group-by-time interaction, $F(1, 24) = 20.66$, $p < .001$, power = .99, which is illustrated in Figure 32. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, $F(1, 24) = 34.18$, $p < .001$, power = 1.00, $d = -2.53$, but a non significant difference at posttest, $F(1, 24) = 0.04$, $p = .852$, power = .05, $d = 0.12$. Further, no significant pre- to posttest difference was found for the control, $F(1, 24) = 0.54$, $p = .468$, power = .13, d

= 0.35, but a significant pre- to posttest difference was found for the experimental groups, $F(1, 24) = 51.33, p < .001, \text{power} = 1.00, d = 1.95$ and the magnitude of effect was large for the experimental group.

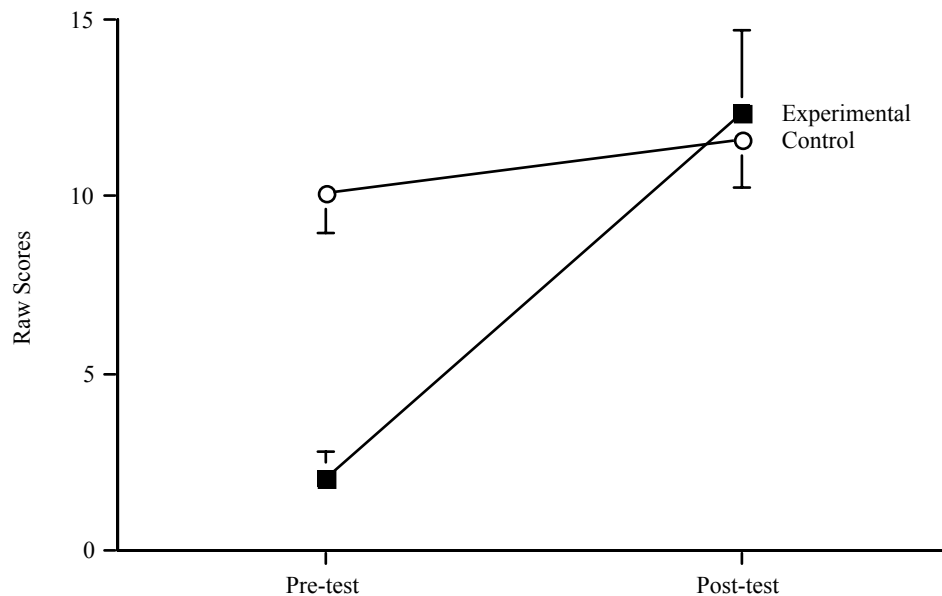


Figure 32. Interaction (+ SE) between experimental and control group at pre and posttest for Word Attack.

Picture Naming Test

Results for Picture Naming Test were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 22) = 0.39, p = .538$. With the pretest results partialled out there was no significant overall difference between the experimental and control groups, $F(1, 23) = 0.71, p = .409$.

Results for Picture Naming Test were analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 24) = 0.02, p = .881, \text{power} = .04$, and the group-by-time interaction, $F(1, 24) = 0.95, p = .339, \text{power} = .16$, but a significant main effect was found for time, $F(1, 24) = 12.16, p = .002, \text{power} = .92$, which is illustrated in Figure 33. Follow-up testing of the interaction using simple main effects found no significant difference between the experimental and control groups at pretest, $F(1, 24) = 0.21, p = .651, \text{power} = .06, d = -0.21$, and at posttest, $F(1, 24) = 0.04, p = .846, \text{power} = .05, d = 0.09$. Further, no significant pre- to posttest differences was

found for the control group, $F(1, 24) = 3.15$, $p = .089$, power = .40, $d = 0.33$, but a significant difference was found for the experimental group, $F(1, 24) = 9.96$, $p = .004$, power = .86, $d = 0.83$ and the magnitude of effect was large for the experimental group.

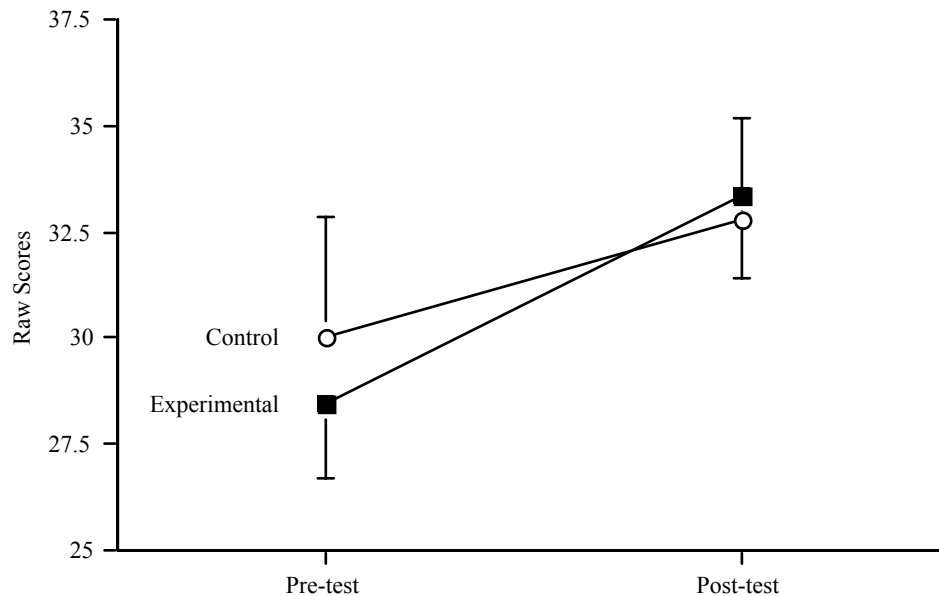


Figure 33. Interaction (+ SE) between experimental and control group at pre and posttest for Picture Naming Test .

Digit Span

Results for Digit Span were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with pretest scores serving as the covariate and posttest scores as the dependent variable. An initial test revealed no violation of the assumption of homogeneity of slopes, $F(1, 22) = 0.20, p = .658$. With the pretest results partialled out there was no significant overall difference between the experimental and control groups, $F(1, 23) = 2.61, p = .120$.

Results for Digit Span were also analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). No significant main effect was found for group, $F(1, 23) = 3.84, p = .062, \text{power} = .47$, but a significant main effect was found for time, $F(1, 24) = 16.99, p < .001, \text{power} = .98$, and the group-by-time interaction, $F(1, 24) = 12.17, p = .002, \text{power} = .92$, which is illustrated in Figure 34. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest, $F(1, 24) = 10.01, p = .004, \text{power} = .86, d = -1.24$, but no significant difference at posttest, $F(1, 24) = 0.00, p = 1.000, \text{power} = .05, d = 0.00$. Further, no significant pre- to posttest difference was found for the control, $F(1, 24) = 0.20, p = .658, \text{power} = .06, d = 0.14$, but a significant difference was found for the experimental groups,

$F(1, 24) = 28.96, p < .001, \text{power} = 1.00, d = 1.50$, and the magnitude of effect was large for the experimental group.

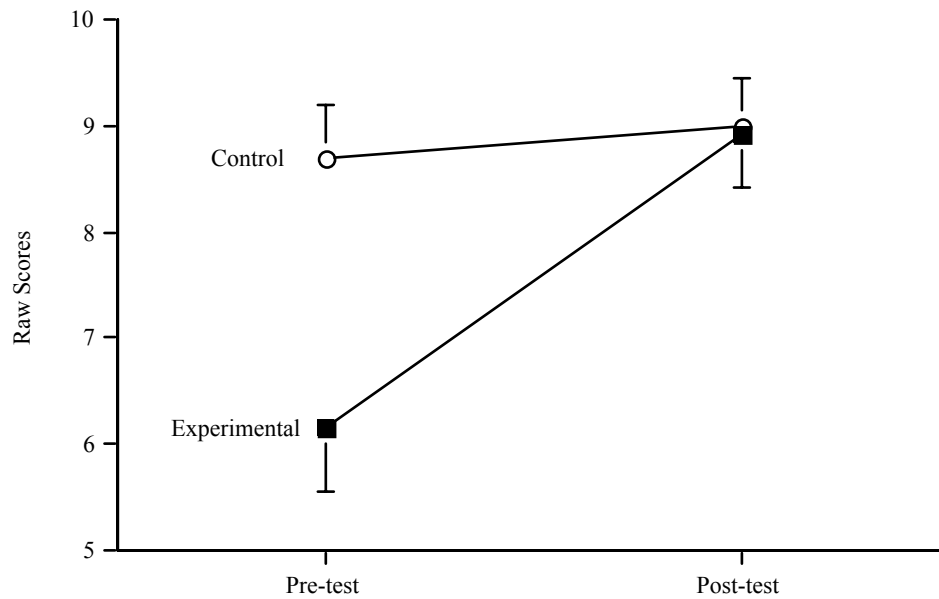


Figure 34. Interaction (+ SE) between experimental and control group at pre and posttest for Digit Span.

Brigance Spelling

Results for Brigance Spelling were analysed using a single-factor between-subject analysis of covariance (ANCOVA), with transformed pretest scores serving as the covariate and transformed posttest scores as the dependent variable. An initial test revealed a violation of the assumption of homogeneity of slopes, $F(1, 22) = 6.49, p = .018$, so subsequent analysis required fitting separate slopes for each level of the experimental group factor. This analysis revealed that pretest scores covaried significantly with posttest scores for both the control, $F(1, 22) = 5.26, p = .032$, and experimental groups, $F(1, 22) = 44.59, p < .001$. With the pretest results partialled out separately for the two groups, there was a significant overall difference between the experimental and control groups, $F(1, 202) = 4.28, p = .050$.

Results for the power transformed Spelling scores were analysed using a two-way mixed analysis of variance (ANOVA). The within-subjects factor was time (pre vs. post); the between-subjects factor was group (experimental vs. control). A significant main effect was found for group, $F(1, 24) = 15.31, p = .001, \text{power} = .96$, and for time, $F(1, 24) = 33.17, p < .001, \text{power} = 1.00$, and the group-by-time interaction, $F(1, 24) = 5.51, p = .027, \text{power} = .61$, which is illustrated in Figure 35. Follow-up testing of the interaction using simple main effects found a significant difference between the experimental and control groups at pretest,

$F(1, 24) = 20.32, p < .001, \text{power} = .99, d = -1.88$, and at posttest, $F(1, 24) = 8.30, p = .008, \text{power} = .79, d = -1.24$. Further, significant pre- to posttest differences were found for both the control, $F(1, 24) = 5.82, p = .024, \text{power} = .64, d = 0.53$, and experimental groups, $F(1, 24) = 32.87, p < .001, \text{power} = 1.00, d = 1.17$, and the magnitude of effect was greater for the experimental group.

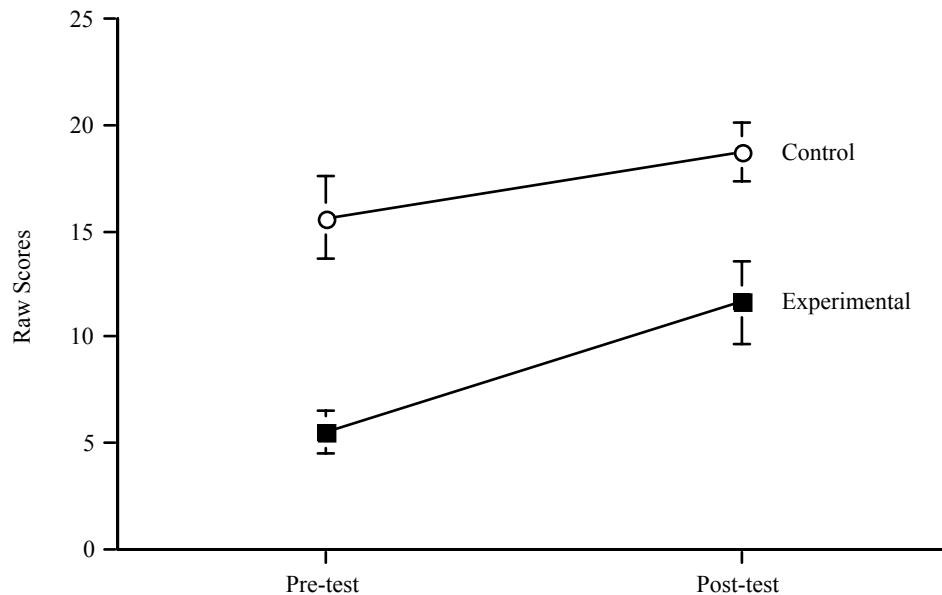


Figure 35. Interaction (+ SE) between experimental and control group at pre and posttest for Brigance Spelling.

Effect Size Calculation for Program

Table 43 contains a summary of the effect sizes (calculated in the same manner as in the major study) for the experimental and control groups. It reveals large effect sizes for the experimental group on all the variables following the reading intervention.

Table 43

Effect Sizes (d) for 100 Lessons Group

	Control n = 13	Experimental n = 13
TOPA	-0.06	1.30
Word Attack	0.35	1.95
PNT	0.33	0.83
Digit Span	0.14	1.50
Spelling	0.53	1.17

Are the Effects of Educational Value? Examining Results in Relation to Test Norms

An examination of pretest and posttest scores in relation to test norms provides an indication of the degree to which this sample of students with reading difficulties lack normally developing phonological processing skills. Additionally, these figures provide information concerning the degree to which program effects “normalise” skills in phonological processes.

TOPA

In Figure 36, raw scores for the total group are shown at pretest and posttest for the Test of Phonological Awareness. It is clear that the experimental group includes some students now in the normal range, and a group-mean close to that range. By comparison, little change is evident for the wait-list group, other than some students improving and some declining (indicated by the standard deviation).

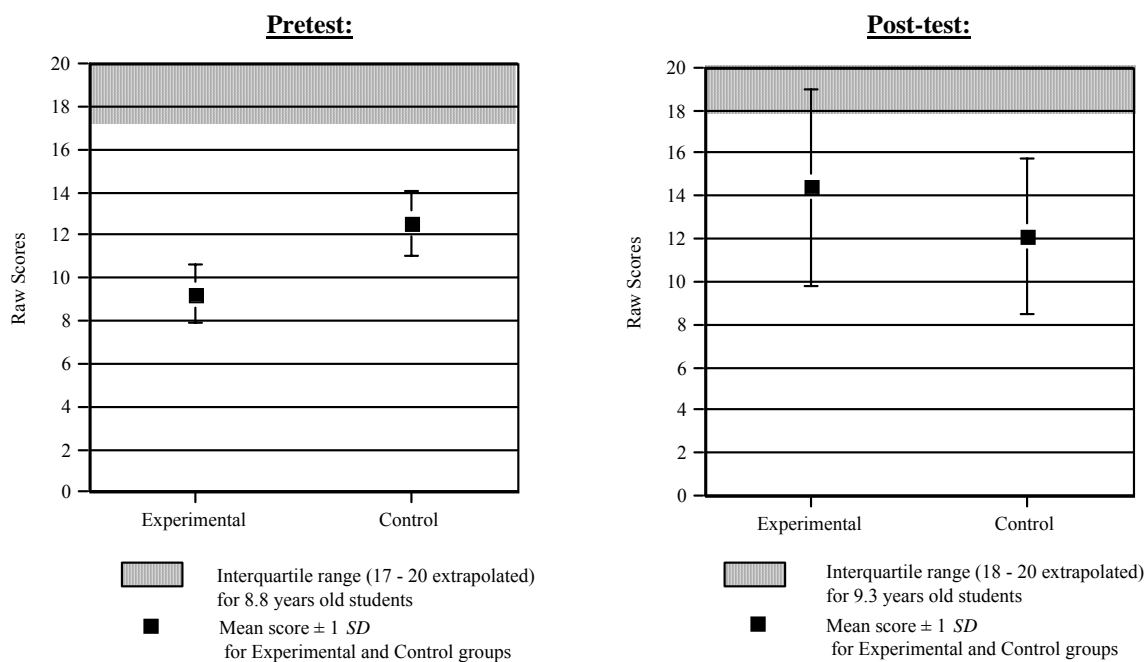


Figure 36. Mean TOPA scores for the 100 Lessons.

Word Attack

In Figure 37, the program effects on Word Attack display a similar pattern to that for TOPA. The experimental group includes a number of students now close to the normal range, and a group-mean closer to that range. By comparison, little change is evident for the wait-list group.

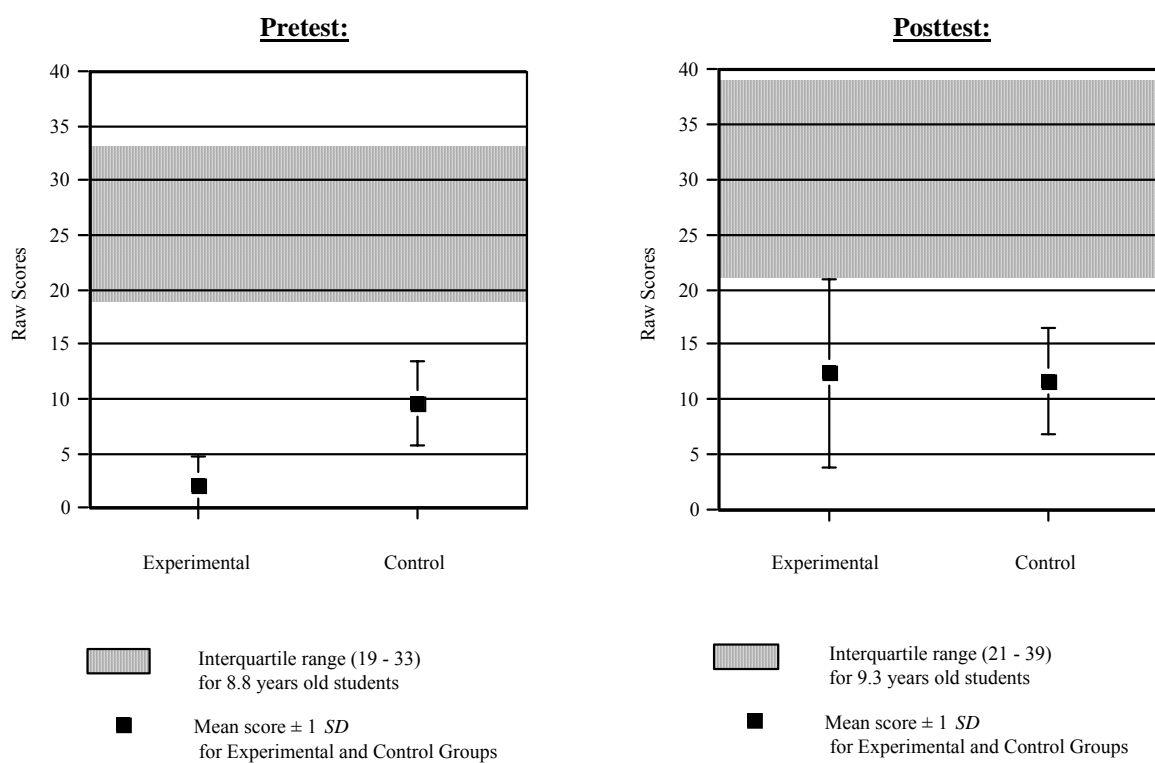


Figure 37. Mean Word Attack scores for the 100 Lessons.

Picture Naming Test

Figure 38 displays the pretest-posttest scores for Picture Naming. As the Picture Naming Test is an experimental test, no norms are available for it. The figure displays a moderate gain for the experimental group, and a smaller gain for the control group.

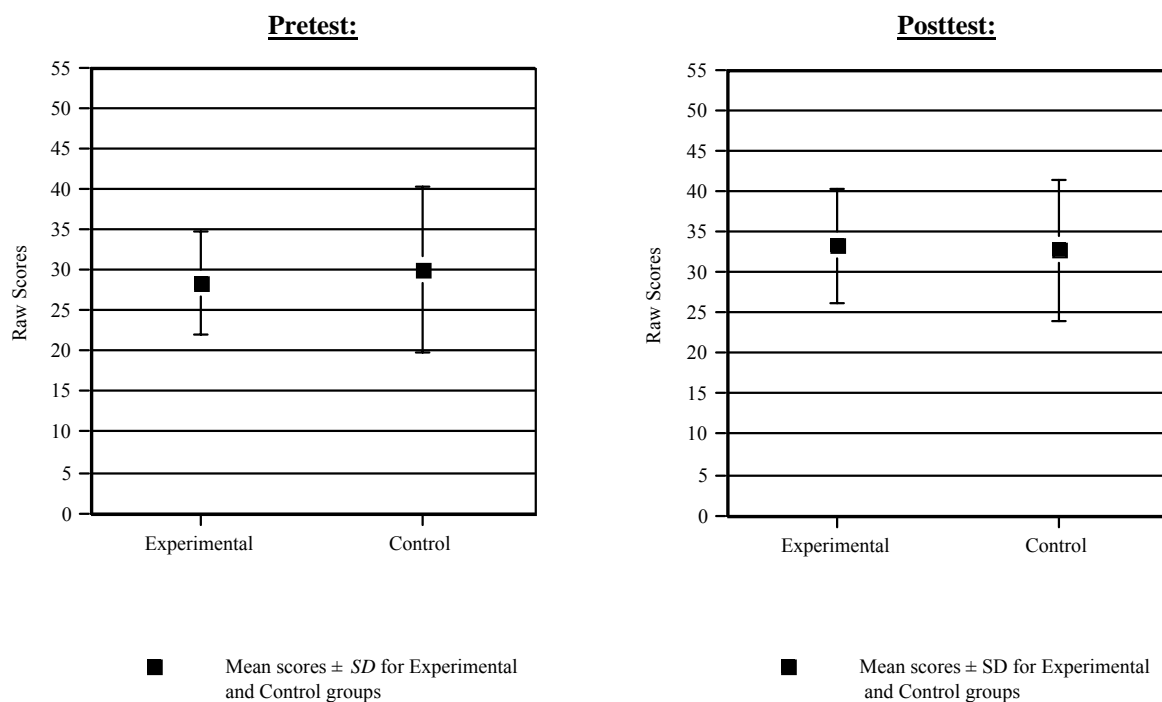


Figure 38. Mean Picture Naming Test scores for the 100 Lessons

Digit Span

Figure 39 shows the effects of the program on Digit Span in relation to the WISC-111 norms. The effect has been marked, and elevates the experimental group mean closer to the normal range. By contrast, little change is evident in the control group.

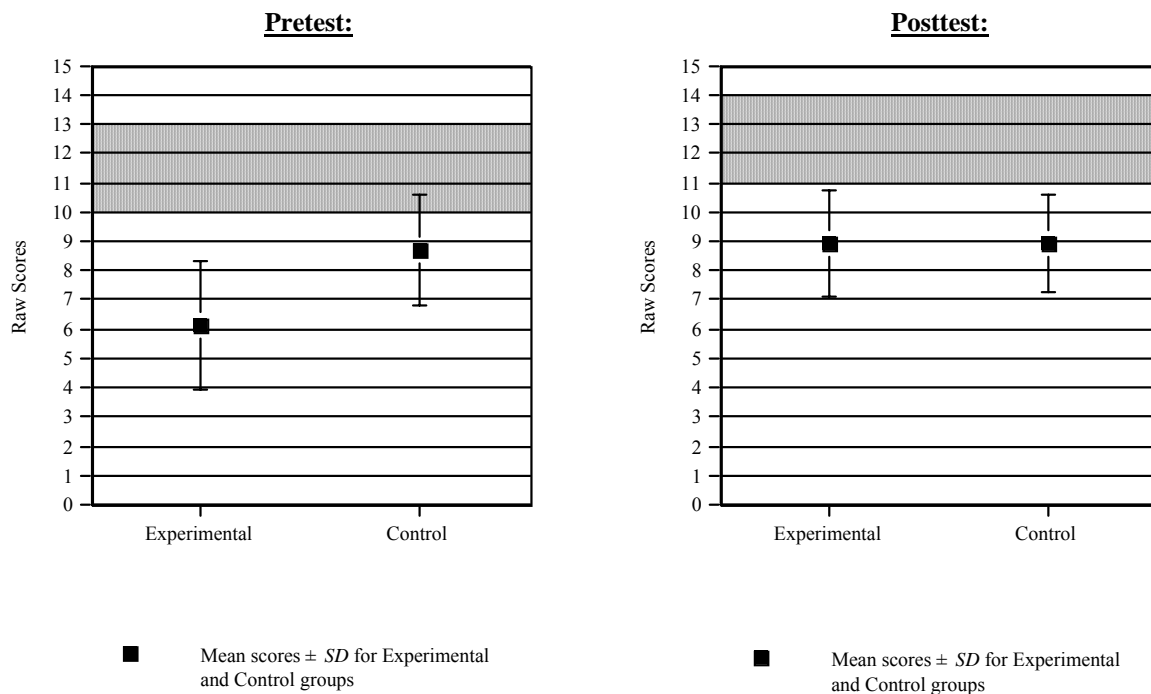


Figure 39. Mean Digit Span scores for the 100 Lessons.

Brigance Spelling

Figure 40 displays the results for the Brigance Spelling test, and follows a similar pattern to the other tests. A feature of this graph is the extent to which all students are below the norm for this test, even after the program. The effects are strong; however, the students clearly remain in need of assistance.

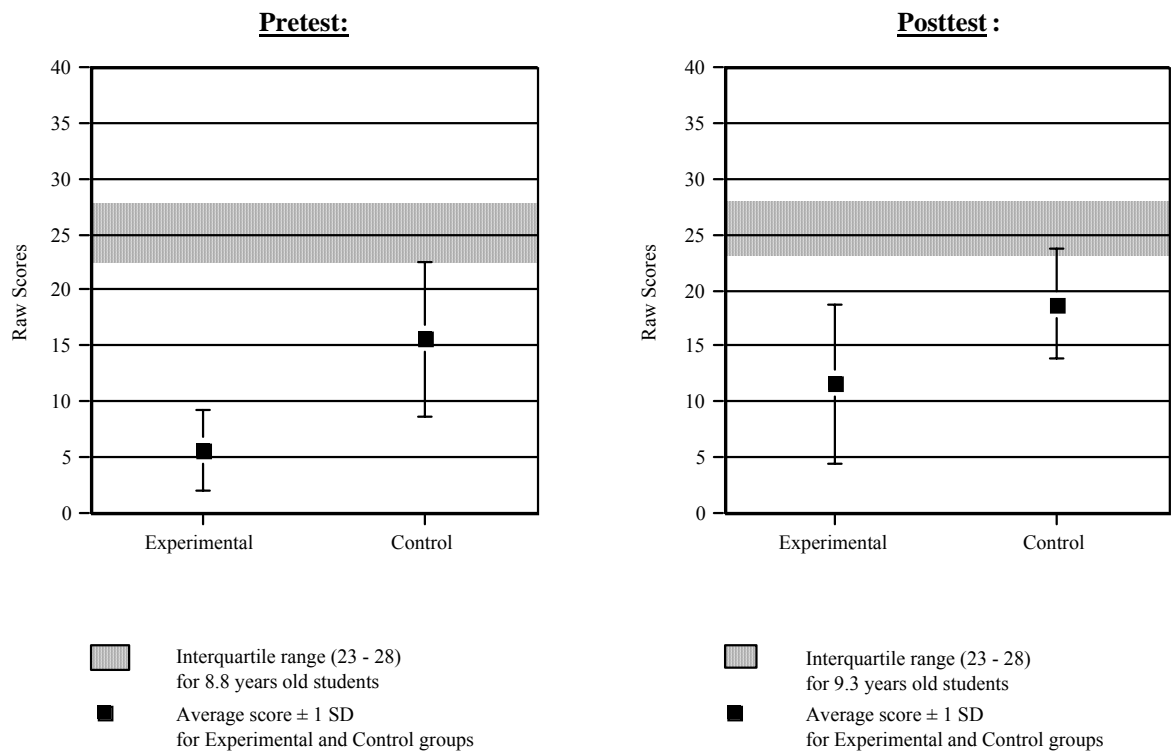


Figure 40. Mean Spelling scores for the *100 Lessons*.

Results Summary

The results achieved with this slightly younger, beginning stage reading group are quite similar (even a little stronger) compared with those achieved with the *Corrective Reading* program. Despite the small size of the sample, the results are readily apparent. The emphases in both programs are similar - a strong focus on phonemic awareness, letter-sound correspondence, corrective feedback, and ample practice. It provides further support for the growing research consensus that herein lies the core of effective reading instruction. However, the *Corrective Reading* program is designed as a group program; whereas, the *100 Lessons* program is written for 1:1 teaching. Thus, the strong effects with this group are reliant upon the capacity of the teacher to develop group management skills, or to translate skills obtained from teaching the *Corrective Reading* program.

CHAPTER TEN: DISCUSSION

This chapter is organised under five main headings: Summary of results, theoretical implications, implications for practice, methodological considerations, and further research. The questions addressed in the research are summarised below.

Does participation in the *Corrective Reading* program increase phonemic awareness, phonological recoding (word attack) skills, other phonological processes (naming, working memory), spelling? Are the effects of educational value? How widespread are the effects? Are there differences in success between the two program levels: A, B1? Does the *Teach Your Child to Read in 100 Easy Lessons* program produce similar results? What is the relationship between the phonological variables measured, and what theoretical implications flow from this.? For example, is there a single or multiple phonological processes? Is success in the *Corrective Reading* program predicted by any of the pretest scores? What are the theoretical implications that follow? For example, are there implications for the phonological representation theory; or, for the reciprocal causation model, or for the best approach to older disabled readers?

Summary of Results

In this study of 206 disabled readers from Year 3 to Year 6 in a number of Melbourne primary schools, the *Corrective Reading* program was instituted for 134 students, while 72 students on a wait-list provided a control. The program has a systematic, explicit phonics emphasis, with attention to letter-sound correspondences, and to the phonemic awareness skills of segmenting and blending.

Pretest and posttest of phonological processes, word attack, and spelling indicated statistically significant and educationally important changes in all variables for the experimental group. The pattern of effects was similar regardless of age, sex, and school, with some variations in magnitude.

All the students in this thesis had received reading instruction in their schools prior to participating in the *Corrective Reading* program. Their failure to make adequate progress can be construed as arising from individual weaknesses, or from a failure of the schools' reading programs to elicit appropriate progress, or from some combination of the two.

The general model of reading assumed in this thesis places word-level processes at the centre of reading disability, and phonological processes as the major underlying abilities

causal to reading development. This model has been neatly described by Ehri (1995) and discussed in Chapter 1. An examination of the normed graphs presented in Chapters 8 and 9 indicates the extent of the phonological skill deficit in this disabled reader population.

The outcomes of the study indicate that these skills can be developed, even in students who have had prior opportunity but been unable to do so in the context of earlier instruction. That these phonological processes develop simultaneously with advances in word attack suggests that such skills remain important even for older students. That the developmentally earlier (phonetic decoding) stage cannot be by-passed has been emphasised in recent times by Share (1995), Share and Stanovich (1995), and by Shankweiler, Lundquist, Dreyer, & Dickinson (1996). This finding conflicts with a view often expressed that any phonic emphasis should be discontinued before Year 3, corresponding to a new emphasis on orthographic processing.

The results of the interventions in this study indicate that discernible and educationally significant change in word attack becomes evident within a relatively short period of time (approximately 50 hours over 7 months for the *Corrective Reading* program or the *100 Lessons* program. These changes in word attack do not appear to be reliant on high levels of pre-existing phonological skills. For example, low picture naming speed at entry was not predictive of poor progress. It is argued that the environmental contribution of carefully structured phonics program is sufficiently powerful to overcome any possible resistance to progress produced by low initial naming speed. Perfetti et al. (1987) noted that, when structured code emphasis teaching was not provided, then initial levels of variables such as naming speed were predictive of reading progress. They also noted that, when effective, phonically-based teaching occurred, the former levels of such variables were no longer predictive of progress.

In fact, the effects of the programs used in this thesis were to increase the level of phonological skills in the areas of naming speed and phonological recoding in working memory in addition to that of phonemic awareness. These findings are consistent with both the reciprocal causation view and the pre-eminence of phonological representation.

The pattern of effects was similar regardless of sex, school, or program, and improvement was evident in a high proportion of the participants. When the two program levels were analysed separately, it was evident that Level A students were more severely reading disabled than the slightly younger Level B students; their scores were lower on all assessed variables at pretest. The effect sizes indicated greater gains for the Level A than the

Level B experimental groups in the literacy variables (TOPA, Word Attack, and Spelling), though all three effects were large. For the other phonological variables (Digit Span and Picture Naming), the effects were greater for Level B (moderate to large) than for Level A (small to moderate).

In an additional study designed to test the effects of the beginning reading program *Teach Your Child to Read in 100 Easy Lessons*, similar analyses were performed. The students were less than a year younger (mean age: 8.8 years) than the Level A and Level B groups but their pretest scores were markedly lower on all variables. As with the other groups there were large effect sizes for the literacy variables (TOPA, Word Attack, and Spelling); however, the effect sizes for Digit Span and Picture Naming were also large rather than moderate as they were for the *Corrective Reading* program groups..

The pattern of results for the three levels of intervention (*100 Lessons*, Level A, Level B) suggests that the overall effects are larger for the more disabled readers. This outcome is suggestive of a period of rapid development as the alphabetic principle is first discovered, followed by real but less dramatic progress later - a notion of *diminishing returns*. Alternatively, the pattern may be explained by regression effects as the more disabled readers show increased tendency towards the mean. The continued large gains of those students who participated in consecutive programs (A, B) suggest, however, that each of the above interpretations is open to challenge.

In examining the relationship between the pretest variables, the strongest correlations were found between the literacy variables (TOPA, Word Attack, and Spelling), and similarly reflected in the regression analyses of pretest variables. Correlations were generally lower than those in other studies, probably reflecting the restricted range of reading ability in the sample. Principal component analyses indicated support for both a two factor solution (TOPA, Word Attack, and Spelling vs. Digit Span and Picture Naming), and a three factor solution in which Digit Span and Picture Naming provide separate factors. A consideration of the experimental and control groups separately revealed that the control group provided a similar three factor solution at pretest and posttest; however, at posttest a different factor structure resulted for the experimental group. Factor One now comprised Word Attack and Spelling; Factor Two - Digit Span and TOPA; Factor Three - Picture Naming. Thus, the effect of the program was to alter the factor structure of the scores obtained by the experimental group at posttest.

Regression analyses were performed on the total group and the experimental group at posttest to add information about the relationship between the variables, and to consider whether pretest variables were predictive of outcome for the experimental group. In analysing Word Attack gains, it was clear that the presence or absence of program was the most powerful predictor by far, although program and initial scores were strong predictors for Spelling and TOPA gains.

Theoretical Considerations

Phonological Representations

As noted in the results, all the students in this thesis were below the average on tests of phonological processes and spelling. Perfetti (1991, 1992) argued that the typically low scores on tests of phonological processing are indicative of problems with the quality of word representation in the lexicon. When representations of words are unstable (or stable but incorrect), matching a stimulus word with the correct phonemically stored counterpart will be slow and error prone, as the child rejects competing phonemically similar but semantically impossible responses. These written word representations are acquired through phonemic mappings to letters but are dependent also on some degree of awareness that words are constructed of manipulable, meaningless speech segments. An alternative explanation - that poor performance on phonological tasks is caused by inadequate auditory discrimination of speech sounds - has not been supported by recent studies (Cornelissen, Hansen, Bradley, & Stein, 1996; Gibbs, 1996).

If these phonological representations are imprecise then tasks such as phonological recoding in lexical access (as measured by Picture Naming speed) and phonological recoding in working memory (as measured by Digit Span) may also present problems for such students, and there is ample evidence that they do (Rubin et al., 1991). For example, if the phonological representation of “dog” is unreliable then the association between the name of the animal and its meaning will be vague. A picture of a dog may quickly evoke its meaning but the phonologically assembled label is slowed because other similar labels (e.g., god, dock, bog) may need to be rejected. Scrolling through a range of possibilities requires more time than accessing a clear uniquely described form.

Similarly, tasks involving short term auditory memory may be difficult because the orally presented stimuli (numbers in this case) are not effortlessly and instantly encoded as unique phonological forms - the process of storage and retrieval is inefficient, reflected in

lower performance. Whereas continuous rehearsal may partly compensate in digit span forward, digits reversed prevents the use of this strategy, and (it was thought) may better reflect the deleterious effects of phonologically inadequate representations. Lindamood described “comparator function” as a critical variable in reading skill, one in which (as for example, in blending) a stimulus or sequence must be retained in working memory whilst part of it is manipulated. Phoneme deletion (one of the most complex of phonemic awareness tasks) requires just this capacity. Analysis of Digit Span Forwards and Backwards in this thesis did not add to the information available from Digit Span Total, and was not included among the presented analyses.

The relatively effortless, automatic, rapid response to text that is the hallmark of skilled reading requires an orthographic lexicon at once comprehensive, and instantly and accurately accessible. It has been argued that the development of the orthographic lexicon in reading has its basis in phonological representations rather than in a visual store of whole words (Perfetti, 1991, 1992).

The connections between word spellings and these representations are a necessary element in orthographic knowledge development, hence it is unsurprising that spelling has been used as a means of assessing the quality of these representations (Perfetti, 1992). The gain in spelling in addition to that in other phonological processes is consistent with the view that the quality of underlying representations has improved in the experimental group.

Landerl, Frith, and Wimmer (1996) noted that in normal readers coactivation of orthographic knowledge occurs in phonological tasks (that is, knowledge of a word’s spelling is used to make judgements about the sounds in a word) whereas for dyslexic readers this coactivation is much less evident. They argue that there is only a weak link between the phonological and orthographic representations in dyslexic readers such that hearing a word does not evoke its spelling, and seeing a word fails to bring forth its sound segments. An inability to establish such reliable links has dire consequences for skilled reading and spelling, and may be due to the imprecision with which sounds are encoded in the phonological representation store.

Elbro et al. (1994) suggest that inadequate phonological representations impede the development of phonological awareness and further that it is at the individual phoneme level that this failure of differentiation may occur. Perhaps the most refractory to phonemic awareness training and to phonics instruction are those to whom Elbro et al. refer. If that is so, some argue, then specialised and intensive *phoneme* awareness may be required. For

example, in the Lindamood (1969) program considerable emphasis is devoted to kinaesthetic (in addition to auditory) cues to assist the recognition of and discriminability between phonemes. Hence, children are taught lip and tongue positions and how the breath is used in order to increase the salience of the sonic differentiation.

It should be noted, however, that low initial scores on phonological processing skills did not preclude progress in this thesis. There may be students who require such specialised intervention, although as yet there is doubt as to how to identify them. Parsimony suggests that, for students of this age, programs such as the *Corrective Reading* program should first be attempted, with the caveat that close monitoring of progress occurs.

Snowling, Goulandis, and Defty (1996) also argue that slowness in reading development of dyslexic students is due to delayed development of clear phonological representations at the beginning reading stage. Others (e.g., Bruck, 1990, 1992; Shankweiler et al., 1996) have noted that delay may be an inappropriate description, as untreated, such problems remain in evidence through to adulthood. In the self-teaching hypothesis described by Share (1995) rapid, whole word reading (enabled through direct lexical access) develops through the effects of practice, effects accumulating each time the phonological coding of words occurs. This sequence (of reliable phonological representations allowing phonological decoding, a skill further promoting direct lexical access) provides both an explanation and an intervention focus to overcome the limits placed on children's reading development by problems at the level of phonology. It is salient that the *Corrective Reading* program places heavy emphasis on precisely that practice. It is also important that development can be stimulated in older readers (as noted in this thesis).

In summary, the theory of phonological representation implies that phonological processes are dependent upon the clarity or accessibility of such representations. If the phonological processes improve during the program, is it because of better clarity of representations? Several studies have noted improvements in phonological processes when phonemic awareness development approaches are adopted.

Similar Studies

The results of this thesis are in concert with those of Lovett et al. (1994) that noted improved phonological processing skills (both speech and print based) in dyslexic children following a program adapted from those used here. The improvements were noted in measures of blending, segmenting, reading and spelling. Foorman et al. (1997) reported a study that compared a Direct Instruction model to both an embedded phonics, and a Whole

Language approach. The students in the Direct Instruction group demonstrated significantly greater gains in word reading, phonological processing and spelling than the other two groups.

Torgesen et al. (1994) studied 244 students from kindergarten through to the second grade and noted that there were reciprocal effects of pre-reading (letter knowledge) on the subsequent phonological development of their students. Although their study began earlier in the students' career and was of longer duration, their students were similar to those in this thesis. The authors noted the strongest effect of such knowledge on phonemic awareness, moderate effects on rapid naming and no discernible effects for phonological memory.

The most common interpretation of such findings is that emphasis on the structure of words increases the quality or accessibility of phonological representations, and such change is represented in improved performance on the variables assessed in this thesis. If, as they relate to reading, naming and working memory are reflective of an underlying variable (representation), there may be little value in attempting to influence these two variables through direct training of them. This is discussed further later in the chapter.

If these two phonological processes are simply marker variables for representation, their usefulness is not diminished as they may have an important function as early predictors of students at-risk (Badian, 1994; Hurford et al., 1994). As discussed in Chapter 4, combinations of tests emphasising phonological processes, given prior to reading instruction, have been very successful in predicting reading progress.

Reading and Phonological Awareness: Reciprocal Relationship?

Earlier discussion (Chapter 4) highlighted the relationship between phonemic awareness and reading development. Whereas, some degree of phonemic awareness is both predictive and causal in such development, the relationship is generally considered reciprocal, in that more sophisticated levels of phonemic awareness develop only after exposure to reading. However, Wagner, Torgesen, and Rashotte (1994) did not find a causal relationship from decoding (word attack) to phonemic awareness. They also found the phonological processing abilities to be very stable, and expressed concern that they may be resistant to intervention. Results of this thesis help alleviate that concern, and are consistent with the results reported by Morais et al., 1987; and Perfetti, Beck, Bell, and Hughes, 1987.

It may be that the stability of phonological processes over time found by Wagner et al. (1994) was indicative of the lack of emphasis on language structure provided in their students' reading education. It has been noted previously that the discovery of intra-word

structure does not arise naturally for many children, and that only when their attention is systematically drawn to it do many students perceive the value of the activity. In this thesis, there was a very strong emphasis on the alphabetic principle.

Several features of reading may contribute to this development of phonemic awareness. Phonemic awareness has no clear function prior to print involvement - it may be enjoyable as a game but there is no other motivation for its continued development. However, for those children who become aware of its utility in aiding decoding, both increased motivation and increased opportunities for use can occur. In such circumstances, more rapid development (due to practice) is unsurprising.

Letters also provide an additional aid to phonemic awareness as a representative of phonemes - their association with phonemes increasing the salience of phonemes. Further, letters are permanent reminders of phonemes whose acoustic properties have hitherto rendered them ephemeral, thus less easy to grasp fully.

For children attuned to the alphabetic principle, reading opportunities provide a qualitatively different experience than for the phonologically naive. For the former, experience with print directs attention to the relationship between the grapheme sequence and the pronunciation. Successful decoding cements this relationship while simultaneously promoting the attitude (important for further reading and spelling) that each letter, and its position in the word, provide important information about the word. They are phonologically alert.

In contrast, the phonologically naive reader may view the word as a visual gestalt, a letter landscape with peaks and troughs in which the contributions of letters are entirely visual. Reading practice (though struggling readers are not renowned for their enthusiasm for reading practice) is likely to entrench attention to the visual features on the periphery of the written word. If the routine use of this strategy is not interrupted, heavy loads on visual memory are likely to limit the rate of acquisition of reading vocabulary to a level similar to that of users of largely logographic languages. Additionally, such readers will have little capacity to independently read and incorporate (via self-teaching) new words.

In this thesis, the content of the programs makes definite conclusions about the reciprocal effects difficult to draw. Particularly in Level A and *100 Lessons*, there are emphases in letter-sound relationships, blending and segmenting (see Chapter 9). Thus, it is likely that phonemic awareness improvement can be parsimoniously explained by the phonemic awareness instruction, rather than because of reading development. The effect sizes

for TOPA in those programs were larger than for Level B. In Level B, there is less emphasis on such phonological skills, but ceiling effects were a greater concern for this group. Despite this, the effect was still large for Level B, a result at least consistent with the view that reading development enhances phonemic awareness. Perhaps tests involving more sophisticated levels of phonemic awareness might have been of value in shedding light on the reciprocity issue.

The findings of this thesis are also supportive of the proposition by Torgesen, Wagner, Rashotte, Alexander, and Conway (1997) that remedial phonics programs for older students with a reasonable degree of letter-sound mastery and phonemic awareness (as were most students) may not require such intensive, dedicated phonemic awareness programs as those for phonological novices. The programs used in this thesis had elements emphasising phonemic awareness, but should be considered primarily as a phonics emphasis approach to reading.

Further support arises from a Foorman et al. (1997) study that noted that Direct Instruction in Year 1 and 2 (preceded by a normal developmentally-appropriate preparatory grade program) produced significantly superior results to that of a Whole Language program that had been preceded by a dedicated phonemic awareness program in the first year of school. Their findings add weight to the argument that phonemic awareness alone does not guarantee reading success, but that phonemic awareness activities embedded within a systematic, explicit phonics program may be sufficient to induce the alphabetic principle in all but the most phonemically-resistant students. This latter aspect has the potential for educational cost savings - in that specialised phonemic awareness programs may be applied more sparingly (and thus more efficiently over a wider population) if one can identify those students unlikely to progress with a less intensive approach. For the others, exposure to a well constructed phonics program may be sufficient to stimulate adequate phonemic awareness, and assist students to progress towards reading independence. In this thesis, the programs were unfunded by other than normal school operating grants; they are very cost-efficient compared to one-to-one tutoring programs.

The Nature of Phonological Processes and the Program Effects

Earlier discussion described the three constructs phonemic awareness, phonological recoding in lexical access, and phonological recoding in working memory. Each of these constructs has been related separately to prediction of reading success (Wagner, Torgesen, et al., 1993). It is unclear however whether they represent independent constructs, or related

constructs. When correlations are found (as they typically are) between measured variables representing these constructs one cannot rule out the possibility that they merely share a common task demand (e.g., careful listening), but in reality are independent constructs. It is also possible that the three measured variables are imperfectly correlated due to differential task demands, but do in fact represent one construct. If that is so, then a study could be designed using multiple measures to produce latent variables. It should evoke increasingly the higher correlations as the number of appropriate measures increases and progressively lowers the extraneous component of any single measure.

Wagner, Torgesen, et al. (1994) make the point that in the prediction of reading numerous reading studies have failed to include a reading measure at pretest (often because in many studies the participants were yet to reach school). In this thesis, word attack scores were obtained at pretest. This is important because Word Attack and TOPA were correlated ($r = .45$, $p < .001$) at pretest, hence the correlation between TOPA at pretest and Word Attack at posttest ($r = .27$, $p < .001$) is subsumed under the correlation between Word Attack at pretest and at posttest ($r = .61$, $p < .001$). It is tempting to suggest that correlations between phonological processes at pretest and reading measures at posttest are at least consistent with causation; however, unless Word Attack is partialled out at pretest, the correlations may be spurious.

In an attempt to add to the understanding of the relationship between the phonological processing variables, the pretest scores were subjected to principal component factor analysis. As shown in Table 27 two factors emerged: a Word Attack/TOPA/Spelling factor, and a Picture Naming/Digit Span factor. Vellutino et al. (1994) proposed that in operationalising phonological coding ability two components are relevant. The first they label analytical phonology; it entails the capacity to deconstruct words into their component sounds, and is evidenced in phonemic awareness and pseudo-word decoding tasks (and also in spelling, Perfetti (1992) argued). The second component, non-analytical phonology, comprised the ability to encode and recall the names of letters and letter groupings. This dual capacity has been named phonological recoding in working memory (assessed by Digit Span) in combination with phonological recoding in lexical access (assessed by Picture Naming). The two factors extracted from the data lend support to this interpretation of Vellutino et al. (1994).

An alternative view forwarded by Wagner, Torgesen, et al. (1993) suggests three independent factors (non-analytical phonology being divided into its two constituents). In this

thesis, when a three factor solution was forced, as shown in Table 28, the factors did divide in that manner. The three factor model is also consistent with other exploratory factor analyses (Mann & Vitunno, in press; Penington, Van Orden, Kurson, & Haith, 1991; both cited in Wagner, Torgesen, et al.) that also find separate factors for phonological recoding in working memory and phonological recoding in lexical access.

Thus, the results can be construed as supportive of either interpretation, and further research needs to be performed before a decision can be made regarding which interpretation is more consistent with the data. Interestingly the factor structure for the experimental group altered following participation in the Program; whereas, the control group factor structure remained the same at pretest and posttest. For the experimental group, Picture Naming remained a separate factor in a three factor solution, but TOPA joined Digit Span producing a Digit Span/TOPA factor, and leaving a Word Attack/Spelling factor (changed from its former TOPA/Digit Span/Spelling factor. The significance of this change is unclear, but it does suggest that the effects of the program are real enough.

Word Attack:

As discussed in Chapter 7, there are excellent theoretical reasons for using Word Attack as a focus for improvement among poor readers. It is a testimony to the program effectiveness that, among all the phonological variables assessed, it was Word Attack that demonstrated the greatest improvement. This was true for the combined experimental group (Group AB), for the Level A group, and for the *100 Lessons* group. For the Level B group the effect was large but marginally less than the effect for phonemic awareness (TOPA).

It is of interest that the Level A and *100 Lessons* groups displayed the largest effect sizes for Word Attack ($d = 1.96$, and $d = 1.95$ respectively), and for phonemic awareness (TOPA: $d = 1.70$, and $d = 1.30$ respectively). In these two programs the explicit instruction in blending and segmenting may play a part in accounting for the greater effects than for Level B students (in which structural analysis receives correspondingly greater emphasis). Segmenting and blending were two factors highlighted as central to phonological processing by Wagner, Torgesen, et al. (1993). With a group of phonemically naïve Prep Year students, O'Connor et al. (1995) noted significant improvements in phonological processing following training in segmenting and blending, conjoined with training in letter-sound correspondences. There are strong emphases on precisely those skills in this program, and the parallel findings of this thesis extend those of the O'Connor et al. study to older students.

From Tables 7 and 8 it is apparent that there was a sex difference overall in the outcomes for Word Attack. Although effect sizes were large for each sex, they were larger for the males, and sex was a predictor of gains in Word Attack. It is unclear why this would be so. Initial scores were similar; thus, regression to the mean is an unlikely explanation. It may be that as teachers appear less sensitive to the presence of reading difficulty in girls (see later in this chapter), they may also be less sensitive to their progress (or otherwise) in the program. Perhaps the lesson style (public reading and responding) is more threatening for females, and their progress is compromised by less open participation.

The question is often asked by teachers concerned about older poor readers, whether instruction should return to the beginnings of reading development, or whether providing increased opportunities for practice will suffice. Although a comparative longitudinal study would help to resolve this question, it appears likely that, for most poor readers, the alphabetic principle will only be induced when there is a clearly explicated program incorporating letter-sound knowledge, segmenting, and blending. Hoping that increasing the volume of reading *alone* (as important as that objective is), or attempting to teach a survival vocabulary of sight words, are not likely to lead to the self-generative stage of reading described by Share (1995) as the point from which self-teaching can replace formal teaching.

Picture Naming

The present findings regarding the moderate effect of the program on naming is consistent with results with third grade students by Rubin et al. (1991) in which phonological analysis training was provided, and an improvement in naming was observed. Also of interest is the Cantwell and Rubin (1992) finding that object naming deficiencies are also evident in adult poor readers. The implications are that maturation may not resolve such difficulties, but also that phonologically-based interventions may be a worthwhile intervention for older students, and even for adult poor readers. In an alternative view posited by Bowey (1996), naming speed is only viewed as important in the beginning stage of reading, as its impact declines when general processing speed increases due to age effects.

The non-significant correlation between Picture Naming and Word Attack is consistent with several findings (Brady, 1994; McGuinness et al., 1995; Wolf, 1991) suggesting that naming speed is related to word identification (through orthographic imaging) rather than to decoding. Orthographic imaging is more likely to have a role in spelling tasks, and a low but significant correlation between Picture Naming and Spelling ($r = .15, p < .05$) is consistent with this assertion. McGuinness et al. have argued that the skill involved in this

type of spelling test may relate to word finding efficiency. Such capacity may enhance the retrieval of the orthographic image of a word, or of constructing such an image by analysing phonological information contained in the orally presented word.

Examination of the correlation and regression analyses similarly indicate non-significant contributions to the prediction of Word Attack at pretest or posttest, or of gains in Word Attack (for the experimental group). Another possible explanation for such findings is that the picture naming task may not be the best task to employ in relation to reading, given that it is not as obviously reading-like as are tasks such as letter naming. The correlation between Picture Naming and TOPA was $r = .28, p < .001$. The results together are consistent with a latent phonological processing variable in decoding, expressed most prominently in phonemic awareness. Further support derives from the low but significant correlation between Digit Span and Picture Naming ($r = .29, p < .01$), and together they form a factor in the two factor solution for the pretest results.

Digit Span

In the phoneme oddity task assessed with the TOPA, memory load is reduced through the provision of pictures to remind students of each of the four words presented. Nevertheless, in order to note which two words (in the end-sound-same subtest) or three words (end-sound-different subtest) share the same final phoneme they must be able to keep the representations active in working memory for sufficient time to note and compare the final phonemes. Hence, it seems likely that phonological working memory plays some part in successfully completing the TOPA, and additionally, in the tasks of sequencing and blending important in decoding unfamiliar words, or pseudo-words (Troia, Roth, & Yeni-Komshien, 1996). Swanson and Alexander (1997) in their study of learning disabled readers noted that working memory contributed only 4% to pseudo-word decoding. In this present thesis, the correlation figure ($r = .23, p = .001$) provided a similar picture.

Brady (1991) pondered whether there is a threshold phonological working memory capacity necessary for success at such tasks. For children who struggle with phonemic awareness, blending and sequencing, and who also perform poorly on short term memory tasks, the question arises as to the optimum foci for intervention. If phonological working memory underpins the other tasks, perhaps it should be an intervention target in its own right. During the 1960's and 1970's the approach known as the ability training model espoused training memory (along with other presumed underlying processes such as visual perception and motor skills). Despite much research energy expended in this field, results were

unsatisfactory (Arter & Jenkins, 1979). Whilst performance on those specifically taught tasks may have improved, there was little or no transfer to the reading task. On the other hand, the literature is replete with examples in which training in phoneme awareness subsequently aided reading and spelling. Gillam and Van Kleeck (1996) reported a study in which pre-school aged children with speech and language disorders improved both in phonemic awareness and phonological working memory following a phonemic awareness training program. Further, they noted that children with poor initial phonological working memory were as responsive to the intervention as were those with better phonological working memory.

What is clear is that the emphasis on sounds in the phonic based reading programs has had a significant impact on students' phonemic awareness and their phonological working memory. These findings provide support for the notion that a better understanding of the structure of words (perhaps leads to improved representational clarity) has a positive impact across the range of phonological processes.

Ehri (1994) suggests part of the mechanism in her Amalgamation theory. When alphabetic readers practise reading specific words by phonologically recoding the words, they form access routes for those words into memory. Readers build these access routes by using their knowledge of grapheme-phoneme correspondences to amalgamate letters-in-spellings to phonemes-in-pronunciations of the words. The letters are processed as visual symbols for the phonemes and the sequence of letters is retained in memory as an alphabetic, phonological representation of the word.

The effects of the program on Digit Span were moderate for the Level A and Level B, and large for the *100 Lessons* group. These results are consistent with the view that there is common variation in Word Attack shared by TOPA and Digit Span. A latent phonological processing ability represented in Word Attack, TOPA and Digit Span was postulated by Bowey (1996) to account for similar findings. Given that the contribution of phonological recoding of working memory is relatively small compared with that of phonemic awareness, then instructional emphasis on directly stimulating phonemic awareness may present a more productive target than that on working memory.

Spelling

Snowling and Hulme (1991) argued that in the normally developing reader the knowledge of word structure gathered during reading activities will transfer to spelling. Treiman (1993) extended the argument in claiming that phonemic analysis training will

positively impact spelling performance even without any instruction in conventional spellings. However, the effect may not be dramatic if gain is measured only by an increase in the number of words spelled conventionally. The relationship between spelling and reading has been compared to that between recall and recognition, in that we are often able to recognise what we cannot recall. Reading may be achieved with only partial acknowledgment of all the letters in a word, whereas spelling requires a complete orthographic representation. Hence, there may be words we recognise on the basis of partial cues, but our cursory attention to the detail of the word does not enable correct reconstruction. Word attack skills alone can certainly aid in producing regularly spelled words, but there are numerous potential phonetically correct spellings for many words, blurring the ready transferability of reading to spelling. Markedly irregular words of course are not constructible from individual phonemic-graphic conversions, even though irregular words usually have predictable letter patterns. It is for this reason that some studies have incorporated a style of spelling assessment that enables the identification of improvement - for example, in phonetic precision. This issue was discussed in more detail in Chapter 7.

In this thesis, there was a significant benefit to spelling from participating in the reading program (Levels AB effect size: $d = 0.98$). It should be noted, however, how delayed was the spelling prior to the program. Even after the program, the students were still markedly disadvantaged with respect to their peers. Viewing the change more optimistically, however, allows the hypothesis that students may have begun to perceive some logical structure in spelling, rather than viewing it as a system completely arbitrary and capricious. It may be that the emphasis on word structure, especially the importance of each letter and its position in a word, may lead to a process analogous to Share's (1995) assertion of a self-teaching mechanism in reading. Davidson and Jenkins (1994) view the relationship of phonemic awareness and spelling as bi-directional, and these results are supportive of at least one of these directions. Burt and Butterworth (1996) assert a direct effect from phonological skills to spelling through the mnemonic enhancement of working memory, and an indirect effect through the benefits to spelling of enforced attention to letter sequence. It may also be that improved segmenting (a result of clearer phonological representations?) allows for more accurate conversion to spellings of the sounds in words. Such an interpretation would be supported if future studies indicated that most improvement occurred in regular words. That possibility was not considered in this study.

In terms of the progress made by the students in this study, an examination of the normed graphs in Chapters 8 and 9 suggests that they may also benefit from a dedicated spelling program. If they are to make the accelerated progress necessary to overtake the ever-increasing average spelling expectations, it may be advantageous to include a dedicated spelling program with a similar emphasis on word level understanding, and with similar design characteristics to those Direct Instruction reading programs described earlier.

How Phonologically Disabled Were These Students?

An obvious finding from the study is the poor performance of the students in each of the measures adopted. At each level of reading program the mean score for phonemic awareness, pseudo-word decoding, picture naming, digit span, and spelling was markedly below the standardisation samples employed in the tests. These findings are in accord with a great deal of research supporting the proposition that phonological coding deficits are present in most struggling readers, and are the predominant cause of their reading difficulties (Vellutino, Scanlon, & Spearing, 1995).

These deficits in phonological coding affect a student's capacity to establish enduring linkages between a printed word (as an entity) and its name; and additionally, limit the capacity to establish enduring linkages between the printed word's letters and the sounds represented. Thus, the effects are evident in under-developed sight word recognition and word attack skills. Sight word reading was not assessed in this study, but increased word recognition skill is certainly a program objective. It is important to recognise that the two access methods are related as "mature forms of sight word learning are alphabetical and phonological at root" (Ehri, 1995, p. 117).

In comparing the pretest means on the various phonological processing measures of the *100 Lessons* group, the A group, and the B group it is evident that the reading rate and accuracy that led to assigning students to their respective groups is paralleled by their scores on the various phonological processing tests.

In this study the students were not beginners, and in most cases already possessed some level of phonemic awareness as assessed by the TOPA, though clearly not at age appropriate levels. Given the causal role of early levels of phonemic awareness in reading progress one may surmise that, for most of these students, their phonemic awareness was not well developed at the time of beginning reading instruction. If one accepts that higher levels of phonemic awareness are dependent on reading progress, and the students were all delayed in reading, it is unsurprising that their phonemic awareness is currently less well developed

than their better reading peers. It is also possible that problems in reading in those students for whom English is a second language are not primarily related to an initial lack of phonemic awareness, but simply a problem in coming to terms with English as a new language per se.

Implications for Practice

Program Effectiveness and Individual Differences

Torgesen et al. (1997) describing a study involving combined explicit phonemic awareness and phonics instruction noted dramatic gains in alphabetic reading skills (almost two standard deviations). This did not mean however that all students responded equally to the intervention; in fact, change in standard scores ranged from 3 to 49 points on the Word Attack subtest. The authors anticipate future analysis detecting individual student differences in behavioural, cognitive and background factors. For example, their preliminary results suggest that initial levels of phonemic awareness, and teachers' ratings of attention were predictive of growth. Other studies (Felton & Wood, 1989) have suggested that naming speed deficits are predictive of slow progress in intervention programs.

The results of this study were remarkably consistent across a range of participants, phonological pre-skills, teachers, and settings. Regardless of age, sex, school, teacher, and SES status, positive and strongly beneficial student outcomes were observed for most students. Using the same criterion as Barker and Torgesen (1995), 87% of students demonstrated improvement in Word Attack.

Is failure to progress primarily due to factors intrinsic to the child, or to factors related to program implementation? In this study, there were no pretest variables that strongly predicted Word Attack outcome other than program membership. A consideration of those students who did not make significant progress failed to reveal any common features - they were not necessarily characterised by low scores on any test. It is not argued that within-subject predictors do not exist, only that they were not among the variables considered in this study.

Treatment Resistance

It is possible that varying proportions of students are able to achieve the state of self-sufficiency in any given program depending on what they bring to the task, in terms of pre-existing skills and diathesis (a constitutional predisposition, in this case, for phonological development). A variant of this position was discussed recently by Vellutino et al. (1996) a study which attempted to identify cognitive profiles that would predict treatment-resistors.

They considered, as had Berninger and Abbott (1994), that those students who make least progress in a validated program may be those with some form of cognitive deficit, whilst the successful students are those whose former lack of reading progress can best be explained by an experiential deficit. This interaction of student capacity and teaching has been described as involving an epigenetic conception of the process of reading development. In this view, genetic signals provide organisms with different propensities to attend to, differentiate, and generalise from environmental stimuli, and different environments provide different levels of exposure to these stimuli (Leonard et al., 1996).

Yopp and Singer (1994) noted in their study that children who were poor at the oral task of phoneme segmentation found it very difficult to learn sounding-out and blending as reading strategies. They also found that appropriate sound-and-blend instructional assistance improved the outcome for beginning students at all levels of phonemic segmentation ability. They view the instructional contribution as influencing the interaction between text and reader resources. At the initial stages of reading development the relative contribution of the teacher should be at its highest, reducing as students become increasingly self-reliant. Thus, an important role for teachers involves the close monitoring of progress in order to enable the choreography of task difficulty with levels of teacher scaffolding. This continuous assessment process is necessary to provide the sufficiently supportive conditions for individual student resources to increase.

Byrne, Fielding-Barnsley, and Ashley (1996) suggest a similar focus in order for instruction to assist students in the complex task of appreciating the alphabetic principle. They also suggest that there is an identifiable category of at-risk students who, in studies thus far, have tended to be resistant to the compensatory techniques adopted. The authors further assert the value of direct instruction in addressing the early needs of these at-risk students, providing increased instructional input to compensate for the limited capacities brought to the task by students.

This epigenetic model implies that progress (P) can be viewed as a product of the student contribution (S) and the environmental contribution (E), thus $P = S \times E$. In turn, S may be conceived of as a combination of capacities (some at least partly inherited) such as intelligence and phonological ability; and E as comprising pre-school history and school-based instruction. Another important variable, that of student motivation, could be construed as subsumed under E, either as a product of history (as in a self-motivated student), or as a product of instruction (as in a motivation system integrated into a reading program). The

value of this conception lies in the extent to which it directs attention to the environment as a critical contributor to progress for some students, and at some stages.

Although there were no apparent predictors of success in this study, perhaps the future will allow for a treatment-student interaction individually optimised to provide for greater or lesser instructional input (intensity, duration), based on pretested student qualities. Byrne et al. (1996) suggest that family history may be such a gross predictor, while others (Badian, 1994; Majsterek & Ellenwood, 1995; Stuart, 1995) suggest that pre-school assessment of relevant variables will allow for a more accurate screening, especially as regards avoiding the ubiquitous and troublesomely high incidence of false positives. Torgesen et al. (1997) point to motivational/attentional factors as possibilities; however, instructional variables should not be ignored.

What are the Limits of Instructional Influence on Progress?

An instructional emphasis does not preclude within-subject causes of failure, but allows for the possibility of resolving problems by manipulating instruction regardless of the source of the difficulty. There are a number of elements within the Direct Instruction programs that may have the effect of mellowing student resistance. For example, the within-program attention to student responses allows for the identification of difficulties at the time they occur, rather than at the program's conclusion.

In particular, the program requirements for repeating tasks until mastery is achieved, of monitoring each student's responses and their daily rate and accuracy checks - should be examined in considering a student's failure to progress as assessed by the phonological processing measures. The mastery tests either within (Level A), or additional (Level B) to the program also provide a safeguard against a student's failure remaining unobserved throughout the program. Even motivational/attentional variations are addressable through the incentive program included as integral to the *Corrective Reading* program.

It may be that there are treatment resisters in most groups, and their identification is dependent upon teacher monitoring skills, and teachers' preparedness to follow the program's guidelines in this regard. It is possible that variation in these teacher/program interactions may be an important focus for future research in reducing the problem of student failure to progress.

There are several safeguards against failure addressed by the program. One involves information provided to teachers on how best to react to any incorrect student responses detected during the lesson. There are clear scripted correction procedures specific to different

tasks, designed to redirect students to the appropriate response. It typically involves an instantaneous correction sequence in which the teacher models the correct response, leads the student through the correct response, and finally tests the student for the correct response.

Teachers are exhorted at the conclusion of most teaching routines to *repeat until firm*. This is designed to provide additional practice when errors are noted, the practice intended to reduce error incidence in the future. If errors are continually made by the same one or two students, the teacher is faced with a dilemma - to slow the pace of the lesson, provide more practice of each task for the entire class, or, to continue at the pace comfortable to most of the class, and hope that the stragglers at least derive some benefit.

A more humane, though resource expensive option is to coopt an aide or parent volunteer to preteach each lesson prior to the regular group lesson. This allows for individually appropriate pacing, tailored to the student's need, and allows the student to continue a rate of progress in concert with his peers during the group session. Usually this double-teaming has the effect of supporting the student in the critical early stages of foundation skill development, improving the student's adaptation to the program structure, and increasing the student's confidence to respond with the group. In the author's experience, and in the outcomes for several students in this study, a short burst of this added assistance allows for successful return to reliance on the group instruction alone.

Another instructional decision point occurs when most of the group makes an incorrect response. In this case, the teacher should examine instructional variables - faulty (perhaps ambiguous) presentation, overly rapid lesson pacing, and, the presence or absence of preskills necessary for correct responding during the current task.

The major issue arising from the foregoing discussion is the emphasis on instructional considerations in any attempts to increase the breadth of a program's success. Both the early detection of problems (monitoring), and the planned response to detected problems should be critical foci in such attempts. As the *Corrective Reading* program was carefully designed to allow continuous monitoring of student progress, a failure to present the curriculum in the prescribed manner (if the deviations are deleterious) should become readily apparent. Some of the deviations noted by the author in schools merely comprise unnecessarily verbose explanations, or interesting but largely irrelevant excursions into other topics. These minor deviations may detract from the elegance of the design, thus reducing efficiency, but they are unlikely to jeopardise outcomes for students.

Other departures from the prescribed program such as omitting some elements, for example, timed reading checkouts, individual turn-taking, or specific tasks, may have a significant effect on the average group progress (if the departures are severe). Alternatively, the modifications may interfere with the progress of some (probably the most vulnerable) students, for it is the most vulnerable students who adapt least easily to ambiguous or incomplete instructional sequences. The early detection of difficulties in any given student is critical to the achievement of broad-band success.

The program designers argue that the *Corrective Reading* program is an individual program, but presented in a group format. For this efficiency to succeed, the teacher must observe each student's responses by first ensuring that choral responding is precise, thus enabling the detection and teacher correction of incorrect responses. The teacher also requires well developed powers of observation to systematically attend to each response of each student. The extent to which teachers can do this depends upon physiological factors (such as, hearing), ability and determination to ensure their students achieve truly choral responding, and, the group size. The Teachers' manual recommends group sizes of 12 or less for Level A, an 15 or less for Level B. In this study, the recommendation to inexperienced teachers was to reduce the number further until the teachers became more skilled; hence, all groups were below 10 in number. The vigilance provided by teachers regarding student response is a major defence against any student's failure in the program. Given that there were students (admittedly a small minority) who did not progress as hoped, this may be an area in which additional training and monitoring of teachers should be a priority.

Thus, several elements of program fidelity appear critical. In a cumulative curriculum, it is essential that all tasks are mastered if students (especially the vulnerable) are to progress. Continuous progress evaluation is needed to detect quickly individual or group difficulty at any point. It is through these program features that problems of progress resistance can be addressed, and hence students spared the fate of participating in an ineffectual educational process.

In the long term, it may be that individual programming, enabling appropriate and immediate response to student difficulty, can more precisely be delivered through the use of computer-based interactive videodisc in conjunction with voice recognition software. In such a scheme, variations in student learning rates can be effectively and efficiently compensated for through differential presentation rates, error correction, and massed and spaced practice.

Student responses could then determine the lesson structure that would, in turn, be capable of adjustment as the needs of the student alter.

Learning Styles

Some research has suggested that reading intervention should focus on a student's strengths rather than weaknesses (Vellutino & Scanlon, 1986); or on students' preferred learning styles (Carbo, 1992; Dunn, Beaudry, & Klavas, 1989). It has also been suggested that even intensive synthetic phonics intervention as applied in this study may be ineffective for students with phonological processing difficulties (Brown & Felton, 1990). In this view, processing difficulties are the immediate and irremediable cause of the reading problem, and alternative (e.g., visual) strategies should be employed to bypass (rather than address) the impediment. Others (Share, 1995) have argued that the development of phonological skills is a necessary element in successful reading and cannot be bypassed. The results of this study indicate that, for the students studied, the problems are not immutable. Whether intervention is enhanced when addressed to a student's assessed learning style remains contentious (Snider, 1992; Stahl & Kuhn, 1995); however, the argument for using preferred modality was reviewed earlier in this chapter, and there has been little experimental support for it. Particularly in reading, the need for students to develop phonemic awareness (because of its causal relationship with reading) is pressing, and should not be dismissed even if a student appears to learn readily in the early stages of a visual approach.

Other Program Characteristics and Effectiveness

There is a consensus that the earlier the intervention for at-risk learners the more rapid and widespread is the success; however, in this study, the students had already experienced some years of reading failure, and were practised at using ineffective strategies for reading. The effects of resistance born of failure can form obstacles to progress at least as difficult to overcome as the original source of the reading difficulty. For this reason, the *Corrective Reading* program includes a motivational system based on assigning points for maintaining speed and error limits. Teachers' comments suggest that this element of the program should not be underestimated in making judgements about the program's effective elements. Numerous positive comments have been made about the student enjoyment and increased on-task behaviour attributable to the points system. Additionally, the system has helped to *capture* the cooperation of many students initially negative about being involved in the program.

An issue relating to program length and intensity is the extent to which the program elevates children onto a plane that allows them to engage in self-teaching (Share, 1995), and thus continue to progress after program discontinuation. Reading programs that include some emphasis on phonemic awareness, such as those by Clay (Reading Recovery), and an adaptation by Hatcher (Sound Linkage), have demonstrated reasonably strong effects over the short term, but share a reported washing-out of the effect in the year or two following the intervention (Hatcher et al., 1994; Glynn, Bethune, Crooks, Ballard, & Smith, 1992). An interesting observation is that Clay's program is intensive (daily), but of relatively short duration (6-20 weeks), whereas Hatcher's is not so intensive (twice weekly) over a 20 week period. The inconsistent results in which some students do well, while others do not, when considered with the washing-out effect over time, force consideration of the issues of optimum program content, intensity, and duration.

One difficulty evident in much of the reading research involves ensuring students transfer their newly developed knowledge and skills to the task of everyday reading. For this to occur, the students need to notice that the new strategies are superior to the old (context cues and initial letters, for example). If the program uses uncontrolled text it is likely to be more difficult for students to effectively use their knowledge, and they may not appreciate the long term benefits of careful word analysis. The *Corrective Reading* program stories used in daily reading are carefully constructed to be predictable by using decoding strategies (though not from context cues), and by using the individually taught sight-words. This provides students with a supportive reading environment that allows for success when the decoding strategies are used, and practice, so that the decoded words are gradually able to be recognised as wholes.

An element contributing to the impressive gains no doubt involves the time and intensity of the intervention. Longer interventions allow for greater content coverage and adequate practice, though of course, there is no guarantee that all intervention designs specifically incorporate such effective teaching characteristics.

Program intensity involves a combination of lesson length, lesson density, and lesson frequency. Lesson length for the *Corrective Reading* programs was about 50-60 minutes, and for the *100 Lessons*, about 30 minutes (designed for younger students). This period allows for a reasonable content coverage in each session and for the integration of new knowledge into the existing structure. As the programs involve a cumulative subskills approach to reading - the introduction of new skills, the practice of recently acquired skills and the amalgamation of

these with the already-established core - requires careful lesson planning and sufficient time for this amalgamation to occur. Program density involves the extent to which students are actively engaged in learning during the lesson time. Various concepts such as time on task, academic engaged time, and academic learning time have been employed to address the issue of student engagement. An observational study by Allington, Stuetzel, Shake, and Lamarche (1986) noted that only about one third of the time allocated to remedial reading instruction was actually spent in direct reading activities, the rest consumed by management issues, waiting, transition, and absence from the room.

One way of promoting student engagement is to plan for overt responses. When students are producing overt responses it is apparent that students are participating, and their learning can be monitored. The additional advantage of overt responses involves the opportunity to provide corrective feedback.

Another element of lesson density involves the proportion of correct to incorrect responses. Students who struggle with reading require high rates of success if they are to adopt new strategies, transfer new skills across tasks, and persevere with the new strategies. Teachers in this study have commented on the high success rates achieved daily through careful lesson design, and student placement at the appropriate program level. The author once counted 300 responses from a student in a 10 minute word attack segment of a *Corrective Reading* program lesson. This represents a very high intensity of participation; additionally, the success rate was very high, above 90%.

Lesson frequency appears to be important, perhaps because of the need for spaced practice of newly mastered skills. It has been noted that students, particularly those at-risk, readily forget what they have learned when lesson frequency is too low. If this occurs additional time is spent in relearning rather than in incorporation activities. Frustration and disengagement are the possible negative outcome of under-scheduling. The program guidelines recommend five lessons per week, although this was not achieved by any of the schools. All schools allowed for four or five sessions per week, but inevitably other priorities intruded over the program period. These involved activities such as school swimming programs and other sports, visiting guests and excursions. In all cases a period of school holidays (either 2 or 6 weeks) interrupted the lesson sequence. The effect of variable frequency was unclear, none of the schools indicated serious problems arising from it, though possibly it may have led to reduced gains for some students. Overall, the average frequency was between three and four lessons per week.

The total contact hours are also relevant. The *Corrective Reading* and the *100 Lessons* programs each entailed about 50 hours of instruction, despite the differences in the numbers of lessons. In the *100 Lessons* program the actual lesson length may vary, depending on the arrangements made to enable hearing every child in the group read during each lesson.

Program Fidelity

Teacher Training

The Direct Instruction model as explicated for the Follow Through experiment paid significant attention to the issue of fidelity of implementation. The designers' examination of implementation research had found moderate to high correlations between student outcome and degree of adherence to prescribed procedures (Engelmann, Becker, Carnine, & Gersten, 1988). The training program for their teachers involved several elements: presenting the rationale, demonstrating technique, providing practice and feedback in response to teacher performance, and, observing real classes - weekly for the first four months, then fortnightly. The process may take a year overall, with the level of complexity of the skills to be introduced increasing over that period. It is evident that the model of teacher training adopted by the designers involves the same direct instruction principles as underlie the student skill development programs.

In the design of the delivery system, the focus was on those teacher behaviours that resulted in optimum student achievement. This concern for detail mirrored the designers' approach to field testing instructional routines also. In that process, theoretical principles of instructional design drove the initial development of content, but it was multiple-setting field testing that determined the final design. For example, the *Corrective Reading* program (*Level B Decoding*) underwent nine revisions before publication (Hanner & Engelmann, 1984).

Engelmann (1988) argues that the average teacher would need to practise an exercise in a reading program at least a dozen times before the fluent orchestration of component presentation and correction skills is attained. These skills involve comfortable and facile use of the specified teacher wording, using lesson pacing appropriate to the example and to the student group, using signals in an unambiguous and natural manner, and providing adequate (but not excessive) reinforcement. In his view, this practice and associated feedback should not take place in the classroom, but in less complex settings such as "dummy" runs with colleagues, etc. Such practice is considered important to aid transfer of training to the real world of the classroom. Engelmann's experience has been that, without safeguards, less than

30% of the skills practised (outside the classroom) will be evident subsequently in classrooms.

The provision of in vivo coaching was found to be especially important for the acquisition of skill. This is unsurprising given the increased salience of observing a model performance in one's own classroom. Glang and Gersten (1987) commented on the value for teachers in seeing how their own students responded to the expert instructional techniques presented by the visiting supervisor.

Within Program Controls

In this study it was not possible to provide the intensity and duration of teacher-training recommended by the authors. It has been noted in other studies that program fidelity can be a major issue in the success of an intervention. Schneider, Kuspert, Ruth, Vise, and Marx (in press) found that differences in focus and duration (time allotted daily and overall program length) had a significant effect on outcome. Additionally, the degree of pre-program and within-program teacher training was found to influence outcomes for students.

A difference in this current study compared to most experimenter-developed curricula involves the extent of control of curriculum and delivery prescribed within the program. The programs used in this study are very prescriptive - the teacher making few judgements about curriculum issues. The content and delivery are scripted, and the teachers' role is relatively transparent. The teachers' skill revolves around classroom management, task presentation, and response monitoring (making decisions about the degree of repetition needed, or the need for error correction).

By contrast, some less prescriptive approaches allow for significant variation among teachers, whose expertise in teaching is assumed. This assumption may not be justified as studies by Lindamood (1993) and Moats (1994b) have shown. A significant proportion of teachers do not themselves have good phonemic awareness, and hence may be expected to have difficulty in both teaching phonological skills, and monitoring their development. In addition, teacher training institutions have been criticised for under emphasising the importance of language structure, failing to provide a good knowledge base in this area for their graduates. The call for renewed emphasis on phonics in initial reading instruction may well fall on "deaf" ears!

Thus, one source of variation in "loose" programs may involve under-developed teaching abilities. Another source in programs that provide only general lesson plans (or even less structured, topic areas), is the variation in the manner in which different teachers may

choose to present the curriculum - the degree of teacher directed vs. self directed learning, the amount of massed and spaced practice, the error correction opportunities, for example. Such-variables are known to impact on student outcomes, and variation at this level can be confounded with the effects of program content.

The level of prescription in the Direct Instruction programs is valuable in reducing, though not necessarily eliminating, teacher differences. There was reasonable consistency of results across different schools in the sense that effect sizes (for word attack and phonemic awareness) were large for all schools. This suggests that the designers' intent of reducing the impact of teacher differences has been achieved to some extent. This is a non-trivial finding as the requirement of training in some programs has been a significant added cost to be considered in conjunction with program effectiveness. For example, in the Foorman et al. (1997) studies, teacher training involved between 30 and 90 hours initially, and subsequent twice monthly lesson observation.

The current study did not systematically attempt to ascertain differential student outcomes based on teacher training levels. The results obtained by the Orana staff (who teach Direct Instruction solely, and train others) were generally superior to those of the other teachers, and the students referred to Orana were more severely reading disabled than the other students (evidenced by lower scores on most pretests). This superiority may relate to the teachers' experience, to their greater commitment to program fidelity, their powers of observation of student responses, their use of feedback and correction strategies, or other teacher effectiveness variables. Alternatively, or additionally, there may be some element of regression to the mean contributing to the results.

It is possible that an increased level of initial training and subsequent monitoring of teacher behaviour would have increased the student achievement levels across the study. It is also possible that as teachers become more experienced their effectiveness will increase. However, the improvements evoked by the teachers who were inexperienced in the program are educationally and educationally significant at low levels of support, an important finding in the real world of inadequate funding for addressing the high prevalence of reading failure. Pressley and Beard El-Dinary (1997) make the point that designers cannot afford to be too precious when their excellent results are not replicated when schools fail to exactly duplicate their procedures. An important research question for any offered program is the degree to which it is robust to changes in its content or delivery across a range of settings.

Where To for These Students?

An issue for schools is whether to continue upon completion of one program level for any cohort. Some schools consider that all needy children should have an opportunity to participate; whereas, others prefer to follow the same cohort through several levels. The issue is a vexed one when resources are insufficient to meet the longer term needs of all the students. Felton (1993) made the point that, for disabled learners, several years of Direct Instruction may be necessary before they are able to make adequate progress without requiring additional educational assistance.

One measure which may assist schools in determining which students should be in the *continuers* group involves consideration of reading volume. The students who participate in the program are likely to have done much less reading than their more facile peers, and no evidence was available from this study as to any increase in the volume of reading of the students. Stanovich (1986) pointed to the effect of volume of reading on reading progress, and it may be that a mediating variable between program conclusion and the need for further intervention is in the amount of reading subsequently performed. The likelihood of students reverting to poor reading strategies is unknown, but a hazard when a study does not include a longitudinal component. It may be possible for students to develop word attack strategies and to make progress in phonological processes, but for such a skill to have little or no impact on day to day reading, or to lose its impact after program completion.

It is for this reason that the continuous within-program tests of rate and accuracy should be important elements in the overall evaluation of program success. There are clear behavioural objectives to be achieved by the end of the program. For example, by the end of Level A students are expected to be reading the daily stories and regular mastery tests at a rate of 60 words per minute at a specified error rate, and for Level B1, 90 words per minute. It is not possible to meet those speed and accuracy criteria if the reader adopts contextual cues, partial word cues, or word shape analysis strategies. Thus, the program prompts the practice of effective reading strategies. A useful further study could assess the extent of additional reading engaged in by students in the program, and the impact of this additional practice on subsequent progress. Such study may lead to within school and home-based programs designed to promote and monitor increased reading volume in the post-program period. Regular subsequent assessment could be used to ascertain the degree to which student progress in reading can be achieved independently for any given child. Some students may have reached the independence level (self-teaching) described by Share (1995); whereas, the progress of other students may stall, indicating the need for a further program level.

An additional question relates to the effects of that additional practice using controlled text compared with the effects using uncontrolled text.

Considering Results in Wider Contexts

Empirically Validated Treatment (EVT)

In 1993, the APA's Division of Clinical Psychology (Division 12) produced a report of a task force on Promotion and Dissemination of Psychology Procedures (American Psychological Association, 1993). The intention was to identify empirically validated treatments in clinical psychology. By doing so it was hoped that the cause of science as the driving force in psychology would be advanced; that practitioners would be influenced in the direction of such treatments; and, that training institutions would place emphasis on such effective treatments. It was further anticipated that health funding services would restrict funding to verifiable treatments, and that the public be informed as consumers of empirically validated practices.

Hudson (1997) extended the concept arguing that similar principles can be applied to education practice. The criteria for well established treatments do appear to be met by the intervention chosen in this study. For example, one criterion involves *the demonstration of efficacy through two controlled clinical outcomes studies, or a large series of controlled single case design studies*. The Follow Through studies described in Chapter 2, and the meta-analyses by White (1988), and Adams and Engelmann (1996) constitute ample evidence at this level. Further criteria included *treatment manuals to enhance treatment fidelity and reliability*. The teachers' texts used in this study are designed for precisely this purpose. As they are scripted to standardise both teacher word and action, it requires only teacher acquiescence to the content and method to ensure that the designated program is actually presented to the experimental group. There are, however, teacher skills in classroom management that may lead to some variability in results.

The criterion of *clearly specified client characteristics* is met through the program entry requirements, as determined by a placement test. This test, based on rate and accuracy of oral reading, ensures that only students at the appropriate skill levels are included in any given program; thus, groups are homogeneous in the skill to be developed, have the pre skills necessary to advance, but have not yet mastered the skills to be taught.

These APA intervention requirements are met by few educational interventions, the risk of underspecifying being the potential for wide variations in any program actually taught to any specific group, and the possibility of including students for whom the intervention is

inappropriate. It is of interest that the APA guidelines translate so readily to the educational setting, though unsurprising, given that much effective psychotherapy has a strong (therapist) teaching and (client) learning component.

NICHD findings: An Example of Empirically Validated Treatment

In 1985, the Health Research Extension Act directed the National Institute of Child Health and Human Development (NICHD) to coordinate research on reading disability and learning disability such that results of research would meet a number of criteria involving scientific rigour. The intention was to define research characteristics that would ultimately lead to methodologically unassailable findings, and benchmarks of consensual knowledge. Issues addressed include definition and identification of reading disabled students, large scale longitudinal studies, careful sampling, and replication of findings. More than 100 researchers in numerous sites across the USA are involved in this cooperative multidisciplinary research with the view of integrating their research efforts. The rigour of their approach entitles their recently reported findings to the status of empirically validated treatment.

A summary of the findings (Reid Lyon, 1996) is presented below (in italics), and provides another suitable context within which to nest the results of this study.

There are as many girls as boys with reading difficulty, yet schools consistently under-identify girls at a rate of three or four to one. In this study the ratio was close to three to one, and was reasonably consistent across schools. At a time when discrimination in society is increasingly being addressed, this finding is deserving of broader scale research and, if confirmed, dissemination.

Longitudinal studies indicate that of children reading disabled in Year Three, approximately 74% will still be so in Year Nine. The findings in this study provide some cause for optimism that the outcome described above is not inevitable, given suitable intervention. Nevertheless, there were some students who did not make apparent progress. There was no discernible pattern to those who did not progress, nor any pretest variable able to predict membership of the no-progress group. However, the NICHD research suggests possible sources such as declining student motivation. In addition to student variables, program sources such as inadequate intensity and/or duration of treatment should be considered. The *Corrective Reading* program has built-in safeguards (through the use of choral responding, individual turntaking, daily reading checkouts) against otherwise unrecognised resistance to progress. However, such procedures can only provide this assessment information if individual student responses are carefully monitored by teachers.

The ability to read fluently for meaning depends primarily on rapid, automatic decoding and recognition at the level of the single word. In this study, the major literacy outcome variables (word attack and spelling) were selected on that basis, as was the program, one designed to improve these skills.

The best predictor of future reading in young readers is phoneme segmentation ability; this ability forms the foundation for the skills described in above. For this study, the phonemic awareness test chosen (TOPA) requires phoneme segmentation ability, and the *Corrective Reading* program emphasises the development of such capacity. The NICHD findings do not argue for dedicated phonemic awareness intervention for older readers because there is not sufficient evidence that meets their criteria for a consensus to be established.

The basis of the reading deficit (phonological processing) should provide the focus for intervention. Efforts should be directed at explicitly and systematically teaching the connection between these phonological rules and the written word. A phonics emphasis provides advantages for disabled readers over a Whole Language approach. The content and delivery of the reading program in this study is consonant with the best available research to date. However there is still much research needed to determine optimum program components and structure. This is discussed further in a later section.

On September 24, 1997 in the USA, the staff of the federal House Education and Workforce Committee presented their outline draft of the Reading Excellence Act. The legislation (passed on October, 23, 1997) states that all the programs to receive support must be based on reliable and replicable research on reading. Thus, the notion of empirically validated practice is in the process of being introduced into education. It is mooted that Reid Lyon, the NICHD director, will play a large part in determining which applications for funding meet the criteria for reliable and replicable research, and take account of the NICHD findings above (Goodman, 1997). This event may well cause repercussions in Australian education.

Social Validity

Another literature providing a context within which to examine results of this study is that of social validity. The concept involves the social desirability and usefulness of an intervention. Arising out of consumer satisfaction indices, the concept has expanded along several dimensions. The type of information collected may be subjective, that is based on the

participants' or others' judgements about the initial need for, and subsequently the value of, the intervention.

In this study, decisions about the need for intervention were made initially by class teachers, on the basis of their observations of the cohort of struggling readers in their grade. As noted earlier, this does not guarantee that all students in need (e.g., females) will be detected. This suggests that group screening using normed assessment (a currently contentious issue in Australian education) may be valuable in ensuring such detection. Normative information involves comparison with a "normal" reference group of age peers. In this study it was available at entry through the placement test, and at post-intervention by analysis of students' pretest and posttest scores. Data is displayed in the normed graphs in Chapters 8 and 9 for each of the tests that have norms. Of interest is the degree to which the intervention has shifted performance towards or into the normal range.

Apart from the type of information gathered, social validity includes a consideration of the intervention process itself - how goals are selected, how satisfactory to consumers are the lesson procedures, and how satisfactory are the outcomes. Indications of each of these elements can be obtained at the beginning and end of the intervention to enable comparison. Kennedy (1992) observed that most of the social validity studies have emphasised the subjective assessment of the value of the intervention.

In terms of the value of the current intervention, some schools did use a questionnaire designed by the author to elicit subjective post intervention data from home teachers and parents, but results were not formally assessed. The only information about student acceptance of the program was incidental, obtained in discussion with teachers and students. In general, the students enjoyed the program, perhaps because it was different to their usual routine, but a number also commented on the success they were achieving. Observation of the students' enthusiasm in classes in which the teacher was warmly brisk, suggests that enjoyment and acceptability may be closely related in students of this age. In other classes where teachers were less comfortable, more sombre, the student demeanour was similar. Gaining social validity information prior, during and after the program from participants may have value in aligning the impact of classroom atmosphere and teacher style on student outcome.

Kennedy (1992) perceives a particular value in including goals and procedures in the social validity framework in those studies in which the primary goal is some form of system change. By contrast, studies directed primarily at knowledge building need not be so

concerned with acceptability issues. Given the potential value to the education system of interventions such as performed in this study, future studies may do well to incorporate such social validity measures in their design.

Clinical Significance

Another term from the psychotherapy evaluation research discussed, by Jacobson and Truax (1991), is clinical significance. The authors make the important point that the efficacy of a treatment cannot be determined solely by statistical procedure because judgements about efficacy are predicated on external standards. Whereas, statistical analyses relate to the probability of a clear and reliable effect occurring from treatment, efficacy questions relate to the worthwhileness of the intervention. For example, a treatment may reliably reduce head banging by 30% from 3000 times per hour to 2100 times per hour. Despite a educationally significant finding and a large effect, it is unlikely that the techniques would be adopted because the intervention is not sufficiently worthwhile.

The standards chosen to ascertain clinical significance may vary, of course. The authors provide several potential indicators. *What percentage of clients showed improvement?* In this study, the highest percentages of students demonstrating improvement of 1 SD (or more) occurred for the processes displaying the larger effect sizes. For example, for the TOPA test 92.5% of those treated improved compared to 4.2% of the controls. For Word Attack, the figures were 91.8% improved against 65.3%, and for spelling 42.5% improved against 8.3%.

Another criterion involves *the recognition by significant others of discernible change*. A questionnaire for parents and teachers (Appendix B) was developed to attempt to address this question. The responses were strongly suggestive of recognised improvement in several reading-related dimensions, but insufficient questionnaires were circulated to warrant deeper analysis.

The complete elimination of the problem appears a worthy objective; although, in this educational intervention it is probably unrealistic. On the other hand, those interventions with a focus on earlier intervention may aspire to such a laudable objective. A more reasonable criterion in a remedial framework could involve reaching or approaching performance levels appropriate for the student's age/grade. An indication of movement towards this may be seen through the use of norms provided in standardised tests such as the TOPA, Word Attack, Spelling, Digit Span.

The effect size statistics are able to provide an indication of the effect of treatment on the mean and standard deviation of the scores of the experimental and control groups as a consequence of the intervention. Another interpretation of effect size is as a coefficient of acceleration. Given that average students continue to progress at an average velocity, students who have fallen behind must accelerate their learning if they are to make up ground on their normally achieving peers. This is a considerable challenge for instruction - to increase the rate of slow learning students to one above the normal. It is unsurprising then, that in educational research, relatively few interventions have large effect sizes. Slavin (1990) considers that effect sizes around 0.25 are educationally significant - the mean effect size of 60 studies he reviewed was 0.27.

The normed graphs presented in Chapters 8 and 9 provide additional information in that they display movement in relation to published norms. In most cases, the reading problem was not eliminated in the relatively short term of the intervention, yet the extent of gains in relation to the normative group gives cause for some optimism, and surely, a determination to continue to intervene.

That *treatment should leave participants less vulnerable to various problems subsequently* is also a worthy criterion. It may be examined in longitudinal studies that measure, for example, high school graduation rates, various follow-up measures of reading, thinking and reasoning, and grade-point averages (Gersten et al., 1988). It may also be argued from a theoretical perspective that significant reading improvement reduces the risk of general education failure through helping to avoid the insidious Matthew Effects discussed in Chapter 4. Reading is usually considered pivotal in all academic subjects; thus, improvement may have inoculative effects across the curriculum. From a somewhat different perspective, Share (1995) argued that students must achieve a certain level of facility with decoding before a self-teaching mechanism allows them to make continuous independent progress from that stage. There is no quantitative measure to pinpoint when that state is reached; nevertheless, the marked improvement in decoding effected through this program suggests that risk factors for future reading and other educational problems are reduced through participation in the program.

Another Wider Issue: The Under-identification of Reading Difficulty in Females

The ratio of almost three to one boys to girls identified by their schools as reading disabled should be a major issue for all concerned with education. State and national testing programs may have been strongly criticised by many involved in the education community -

unions, teachers' colleges, and various education consultants. A valuable outcome of such testing, however, may be the identification of discrimination against a large proportion of the population - females. If this testing results in a similar finding, then teachers may become more sensitive to its occurrence. It is likely that, for this sensitivity to be enhanced, more systematic screening will be required than has occurred in schools in recent years.

Teachers have been exhorted in pre-service and in-service training to rely on informal reading observation to gather information about the reading progress of their students. There has been a parallel argument intended to dissuade teachers from using formal standardised reading assessment. These arguments are based on the Whole Language position (described in detail in Chapter 3) that reading is a natural process, unique for each individual, and not amenable to standardised testing. Further, the argument continues, the essence of reading involves the joint author/reader construction of meaning, a collaboration opaque to the scrutiny of word-level reading tests (or subtests). The identification of this discriminated-against group of students will require an adjustment (perhaps, revocation) of that view. Such an outcome would be doubly valuable, as it is now acknowledged, at least within the research community, that word level assessment is very appropriate, in fact a vital element in screening for reading problems.

Methodological Considerations

Research Design

The purpose of this section is to examine the design of the thesis to consider whether it is likely that uncontrolled factors can more reasonably account for the results than the experimental intervention

Internal Validity

Internal validity refers to the degree to which error variance is controlled within the experiment. Random assignment of students to experimental and control groups is preferred because the groups can be considered equivalent on all but the independent variable. This was not feasible in the present thesis as the groups were independently established across a number of schools and their selection for the thesis was based on opportunity, as is often the case in school-based studies.

The experimental sub-groups were small, each of five to ten children identified by their school as in need of reading intervention. It was not feasible in this thesis to devise a control group (using random selection) that would have an alternative experimental program of similar duration, intensity, teaching style and quality. Such organisation is ideal as it

precludes the threat to internal validity posed by a rival explanation for any differential improvement in the experimental group. Rival explanations could include the novelty effect, the effect of teacher enthusiasm, or small group bonding leading to improved attentiveness of participants. On the other hand, it would not have been desirable to place students in a program ostensibly similar to the intervention for which no real benefit to the student was anticipated. In such a case ethics intrude and reduce experimental design options.

The control group comprised students identified as having the same requirements for reading intervention as the experimental group but who were currently on a waiting list. The interval between pretest and posttest was similar (approximately 7 months) for the two conditions.

The second line of defence of internal validity involves restricting the impact of extraneous variables. In this thesis, the various groups were selected on the basis of their score on a reading test (The *Corrective Reading* Decoding Placement Test, 1988). As a consequence the experimental and control groups are considered equivalent, at least in rate and accuracy, with respect to measured reading ability. The groups were all drawn from the northern and western suburbs of Melbourne (see SES data in Chapter 7), and extraneous variables such as socio-economic status, or the numbers of students with English as a second language should be evenly distributed across experimental and control groups.

When random allocation is not feasible, ensuring the groups are matched on potentially contaminating variables is sometimes attempted; however, it requires equal group sizes to match each member of the experimental group with one from the control group. Instead, a post hoc examination of likely differentially acting extraneous variables was performed within the inferential analyses.

Regarding possible effects of novelty (Hawthorne effect), Hempenstall (1988) followed the progress of a group of students in the earlier (Engelmann et al., 1978) edition of the program over a two year period. The strong effects continued for the two years of the program, suggesting that novelty is not a reasonable explanation. In the first year, two groups received the program (Level B), and each demonstrated similar gains in that year of the program. In the second year only one group was maintained on the program (Level C), and it continued to display a similar rate of progress, while the second group received normal classroom instruction. The second group did not progress beyond their previous year's attainment during that second (no intervention) year. A study by Branwhite (1983) also followed experimental and control groups with similar results - in that only those students

participating in this program over a two year period continued to make accelerated reading progress each year. In Branwhite's study, the novelty explanation is even less likely as the control group received a different intervention program in the first year, but did not progress until receiving the *Corrective Reading* program in the second year.

Despite the results from previous research, it is true that the design of the present thesis cannot entirely rule out the possibility that Hawthorne effects provide at least some influence on the results. There is, however, an indication that such effects are unlikely to be large, based on the assumption that Hawthorne effects tend to be most influential in the short term, diminishing over time. There were five students in this thesis who appeared in two experimental groups because they were at a school that followed the Level A program immediately with Level B. The results were of interest despite the small size of the group, because the data obtained sheds some light on the dose-response issue, and on related questions such as: Are the effects ephemeral? Is a novelty effect responsible for initial improvements? The effect sizes presented in Chapters 8 and 9 indicate that students continue to progress in their phonological processing ability when participating in the next level of the program, and these effects mirrored those of the original program. The results suggest that novelty effects are an unlikely prime cause of the change. Further, they suggest that the instructional sequence within the programs, and from one to the next, are appropriate for the entry skill levels of the students.

The *Corrective Reading* program involves concentration and commitment, and the secondary students in the 1988 study were not always enthusiastic, or supportive of their peers. Explanations based on escaping from the normal curriculum to an exciting environment did not appear plausible in that case and in other cases observed in secondary schools. In the programs observed in this thesis, however, it was true that the students appeared to enjoy the experience.

Maturation-based explanations assert that normal developmental maturation rather than the intervention can account for any changes. The students in this study were in Years Two to Six, and in the earlier (1988) study in Years Seven and Eight. If maturation were to have a major effect the incidence of reading problems should decline over time rather than continuing (Juel, 1988) or even worsening (Stanovich, 1986). Given the variation in ages described above, it is also unlikely that coincidental maturational "bursts" occurred across all these ages simultaneously. In any case, if maturation is to play a role (even over such a brief period), it should be equally evident in the control groups.

As the experimental and control groups were in a variety of schools (State and Catholic) it seems unlikely that any extraneous events over the period of the program (historical threats to internal validity) could coincidentally affect only the experimental group.

Any effects on students of the test or testing procedure should have been equally distributed across both groups. These include student effects such as being sensitised by the pretest, practice effects, and negative reactions to posttesting.

As testing was performed by a number of people (eight), it is conceivable that there could be variations in the accuracy of the test administration. However, all seven additional testers were Masters students or qualified teachers, trained by the author in administering and scoring the tests. Training involved the provision of a written test manual containing administration and scoring instructions, modelling by the author using a child in the study as a subject, and practice by the tester in which the author provided corrective feedback.

In most cases the tester in the pretest did not administer the posttest. This related to availability rather than design; however, there was no pattern across either the experimental or control groups.

Statistical regression is another threat to internal validity; however, the groups were similar in assessed reading level - it was the basis of their selection - and in both cohorts means did not differ significantly on most pretest instruments. In those schools in which there were both control and experimental groups the decision about which group received the treatment first was not based on problem severity. In other words, one would not expect regression toward the population mean to occur differentially across the groups.

Other potential hazards include the possibility of dropouts affecting the results. There were students not included in the results because of absences at the time of pretest or posttest, or inter-school transfers. This was true for both the experimental and control groups. In Level A, there were twelve experimental and two control dropouts; in Level B, eight experimental and two control dropouts; in the *100 Lessons* program, one control dropout. These represent a small proportion of the total, and unlikely to have had a major effect on results, particularly given the effect sizes obtained. Additionally, it is difficult to imagine any systematic pattern to these absences.

Issues of selection may jeopardise group comparability. For example, it is conceivable that schools prepared to provide a special reading program differ in important aspects from schools that are either unable to or choose not to do so. These school qualities may be efficacious in enhancing reading development but not obvious until the program's

commencement, and the subsequent student progress falsely attributed to program effect. However, the control group comprised wait-list students, and was drawn from the same schools as those in the experimental group.

Because the experimental group consisted of a number of smaller groups, taught by numerous teachers in different settings, it could be that variation in treatment may be sufficient to mask program effects. One slight possibility is that some or most teachers ignored the program guidelines, and their own various modifications had no deleterious effects on outcomes; that is, regardless of how the program was altered it was similarly effective. Presumably this argument implies that teachers per se are the major agents of change, and the program design is of little importance in this process. This is an interesting assertion but specious, since there would be far fewer reading problems in existence if it were true that teachers are universally and similarly effective.

Additionally, the program designers went to some pains to assess the differential effects of altering even minor program elements, in order that the content and delivery of the program was optimal. For example, in Operation Follow-Through, teachers received a great deal of pre-service and during-program support. The level of this support was based on the theoretical and empirical importance of the principles of program design and program delivery. These principles were precisely explicated in the programs because they were considered contributory to student outcome. During program construction a variety of techniques of program delivery were tried and evaluated, and differences in student outcome were observed with variation in these factors. Additionally, the program designers had noted that teachers' straying from guidelines was associated with reduced outcome for students. The final program construction was based on the optimum mix of content and delivery practices as evidenced by their trials, and hence seen as the gold standard. For the *Corrective Reading* program, at least 10 revisions were completed prior to the publication of the current version (Hanner & Engelmann, 1984). Despite the care involved in determining the sequence, content, and delivery elements of the current program, it is an ongoing process of finer-grained analysis of the program that allows for continual revisions and subsequently improved efficiency and effectiveness.

Another of the threats to internal validity - variation in treatment - only needed consideration in the event that there were no observable treatment effects of significance. As there were observed treatment effects then any such variation was not drastically deleterious; however, it may still have reduced observed effects below their potential.

In this thesis, it was recognised that the control over treatment variation could not be as overarching as recommended by the programmers. This difficulty related to the large amount of time required to provide on going support and monitoring over a number of sites; and secondly, on the relationship between the author and schools. Although in some cases the author was instrumental in the school's adoption of the program, in others he was considered largely as an observer, able to provide for the school program evaluation results in return for access to students for his research. This relationship was not conducive to enforcing strict program adherence. Nevertheless, in most settings the programs were being implemented for the first time, and most teachers were anxious to consult with another more experienced with the program. The author developed and provided a manual containing advice on day-to-day running, and exhortations to follow the treatment manual. He spent several hours with new teachers providing information, and modelling the presentation and correction procedures. In one case, post-graduate students were also assigned to follow the progress of the program, and to monitor and report any difficulties as they arose.

In some schools the program was implemented by teachers reasonably experienced with the program, and it was possible to link new teachers with them to provide initial and ongoing support.

An interesting question with tightly structured programs requiring faithful administration is the degree to which the recommended training conditions can be met in the various school settings that comprise the real world of education. How much training and monitoring is necessary for the program's true potential to be met? How much training and monitoring will suffice in most cases for a lesser but still educationally significant effect? This issue is explored in more detail later in this chapter.

External Validity

External validity involves the confidence that any findings are not restricted to the group of students in the study, but can reasonably be generalised to other students in different places, and at other times.

Potential threats include the possibility that the pretesting process itself has an impact on students' responsiveness to the program, and if that is the case then the results would not be generalisable to a non-pretested population. However, program pretesting (as distinct from this study's formal pretest) in the form of a Placement Test is a required element in the *Corrective Reading* program, and hence results achieved through the program are not claimed to be available to a non-pretested population. It would be of concern if the particular structure

of the pretesting for this thesis was unique in terms of test content or tester qualities, and it is the combination of any such aspects of the test battery that restricts generalisation of results. It is helpful to consider the results of other studies (reviewed earlier) in which different schedules of pretesting over almost a twenty year period have produced quite consistently good results. Such a program history, of which this thesis forms but one contribution, invites the conclusion that the pretest schedule should not be considered a threat to external validity.

As the participants of the study were not volunteers it is unlikely that there could be an interaction of selection and treatment among individual students. It may be argued that the schools which agreed to adopt the reading program have qualities that other future adopters of the program do not, and it is the presence of such unidentified qualities that accounts for success, thus limiting generalisability. As indicated earlier, the argument is an inductive one, and this study should be considered in the context of many others in which there has been significant variation in many characteristics (student SES, levels of command of language, school size, location, public or private nature). Results have been consistently reported as impressive, despite population variations, whenever the key elements of the program have been adopted (appropriate pre-selection of students based on reading criteria, and fidelity to the program structure and content).

The same argument applies to interactions of treatment with settings (in this study there were numerous settings), and with individual student or school history. The more frequently a carefully detailed procedure is applied in different settings and time frames the more confidence one has that the interactions described above do not play a role sufficient to limit generalisability. In other words, the same random errors become increasingly unlikely across a variety of studies.

Construct Validity

Construct validity of the variables used in the study refers to the degree to which they are well defined and measured. A weakness in this thesis involves the use of only one measure of each variable. The decision to use only one measure made the study more manageable at the cost of potentially under-representing the constructs involved. Parker (1990) refers to the use of multiple measures of a given construct as a means of triangulating the construct. Wagner, Torgesen, and Rashotte (1994) view multiple measures as enabling measurement error reduction through removing task variance from the measures. They describe the procedure as allowing measurement of the true substance of the construct - "latent variables rather than observed variables" (p. 76).

This problem of the distinction of a construct from its measurement is a long-standing one in research, and some (e.g., Morris, 1994) have even argued for a gold standard reference list of tests to help with the operational definition of reading constructs. In this thesis an attempt was made to use tests (or test formats) that have been used successfully in other studies, are generally considered to be a measure of the relevant construct, and that have acceptable reliability and have clear guidelines for administration. These issues were covered in detail in Chapter 7.

Statistical Conclusion Validity

Statistical conclusion validity (Parker, 1990) involves using statistical procedures appropriate to the conclusions reached. An important issue is the power of the procedure to find an experimental effect when one exists, that is, avoiding a Type II error. Power depends on several features: the size of the sample, the effect size, and the probability of a Type I error. Cohen (1988) recommends that power be set at .8, and he expresses concern that few studies reach that level. The higher the power, the more likely one is to find an experimental effect when one occurs.

In this thesis, acceptable power was set at .8 as Cohen recommended. Alpha was set at $\alpha = .05$ (unless otherwise indicated) as a reasonably low probability of inventing a significant difference. The next element in the equation is effect size. By using Cohen's (1988) tables it was possible to consider effect size to assist determining appropriate sample size. In a meta-analytic review of the literature of Direct Instruction programs, including those used in this study, White (1988) reported a large effect size (0.88). However, in this thesis with numerous dependent variables it was not evident that large effect sizes would occur for all variables. Accordingly, Cohen's (1988) table was used to select a sample size (200) that allowed for a low effect size of $d = 0.25$, whilst maintaining power at .8. This decision produced a strong likelihood of finding significant relationships should they exist. As it eventuated, effect sizes ranged from moderate to large, and hence smaller samples would have sufficed to provide adequate power in most cases.

Reliability of tests used has an effect on the power of the statistics as the amount of error variance rises. The mean reliability of the tests used in this thesis $r = .85$ is considered a moderate level. Analysis of simple gain scores between pretests and posttests also provides reliability hazards because of the high ratio of error variance to true variance, and hence requires caution. This issue is further discussed later in the chapter.

If the program is administered in a non-standard manner, statistical conclusions are threatened. The program design and implementation instructions are intended to minimise such variations, as were the teacher monitoring and training procedures used during the thesis and discussed in detail elsewhere. This threat to statistical conclusion validity was paid particular attention because of the number of sites from which the experimental and control groups were drawn.

Choice of Analyses

The selection of statistical procedures for a non-equivalent group design study presents some challenges. Cook and Campbell (1979) argue that no one procedure is ideal for all non-equivalent group designs, that each can introduce substantial and different biases, and that an analysis of the “structure and interrelationships of the data” (p. 186) is necessary to find the optimum instrument. The optimum instrument is one that is able to partial out the effect of selection differences from the treatment effect.

Given the pretest/posttest nature of the design the analysis of simple gain is an attractive option. However, Anastasi (1988) argues that the reliability of gain scores inevitably suffers even when the original scores possess reasonable reliability. Dugard and Todman (1995) are disparaging about the continued prevalence of such analyses, claiming increasing agreement among commentators that the use of analysis of change scores and repeated measure analyses of variance (ANOVA) are generally inappropriate in such designs.

They contend that the use of repeated measures ANOVA in mixed factorial designs, or the one-way ANOVA of change scores, assumes that there is a randomisation within the experiment’s participants of factor levels (pretest and posttest). This is clearly an impossibility as the order of pretests and posttests is fixed.

Further, they argue, the change score is correlated with the pretest scores, and hence does not have the desired effect of reducing residual variation - an objective of collecting pretest information. In their view, an analysis of covariance (ANCOVA) is a more sensitive measure of change because it allows the pretest information to increase the power of the test. The ANCOVA is more sensitive to small real effects (compared to an ANOVA) because a smaller error term is produced. Hence, there are larger main effects and interaction sums of squares, an advantage achieved because the part of the within-cell variance attributable to the covariate is able to be partialled out.

The ANCOVA and ANOVA assumptions must be considered if the procedure is to be appropriately used. A shared assumption is that the relationship between pretest and posttest

scores is linear. In the event of such lack of linearity data transformation should be performed. A second assumption involves the posttest scores and the homogeneity of their variance. Dugard and Todman argue for the ANCOVA's robustness in this regard with the proviso that group sizes are not greatly dissimilar, and there is a more or less normal distribution of the covariate scores.

Myers and Wells (1991, reported in Dugard & Todman, 1995) make reference to several assumptions for ANCOVA. The first involving independence of treatment and covariate requires that the covariate is assessed prior to the intervention (as it is in this case); and that participants are randomly allocated to treatments (as was not possible in this thesis). The second of these assumptions also can be considered achieved by the use of random assignment; it involves measurement of a fixed effect covariate errorlessly. A further requirement (tested prior to the ANCOVA procedure) is for parallel regression lines, posttest on pretest, an assumption not relevant to ANOVA.

Mok and Wheldall (1995) are more cautious about the use of ANCOVA, being particularly concerned about the assumptions of homogeneity of regression, and errorless (or at least reliable) measurement of control variables. Whereas, they applaud the concern to use the most sensitive tools, these authors warn against the potential increase in Type I errors when ANCOVA is used rather than gain scores. In non-randomised designs the compensating for initial differences implied by the use of pretest scores as covariates can only be truly achieved when there is a perfect correlation "between the predictor and those attributes for which it is seeking to compensate" (p. 200). Thus, the ANCOVA is reasonably precise if group pretests are similar (as was the case in most analyses for this thesis).

Mok and Wheldall consider gain scores can have good reliability and cite several sources in support. They highlight the advantage for gain scores in retaining the measurement unit of the test, and consider their use especially appropriate for non randomised studies employing unequal sized groups. Since the ANOVA for the comparison of gain scores is identical to the analysis of time-by-treatment interaction in a two-factor ANOVA, their argument is as true for the 2x2 mixed ANOVA as for gain score analysis.

It is thus arguable whether there is a single most appropriate statistical procedure for this thesis. A compromise suggested by Mok and Wheldall is to use multiple measures in addition to the ANCOVA, such as effect sizes and repeated measure ANOVA's. If the results are consistent then one may argue that assumptions violated will be different for the various statistics employed, and therefore less likely to lead to spurious conclusions.

Given the variation of opinion over the correct statistical procedure, a decision was made to follow the conservative position of Mok and Wheldall by using several tools. As it eventuated, in most cases statistical significance was so strongly established that assumption violation became of lesser concern. Bearing in mind the relatively large sample sizes, in most cases it was obvious by visual inspection that educationally significant differences were present. In particular, the pattern of results for the experimental group was consistently different to that of the control group.

Null Hypothesis Testing

A number of writers have expressed concern about the practice of null hypothesis testing (Cohen, 1994; Hammond, 1996; Thompson, 1996) as the epitome of psychological data analysis. They argue that statistical significance tests are often inappropriately used and misinterpreted, leading to the unfortunate outcome that genuine experimental effects may not be detected, and conversely that chance effects may be assigned undue importance in a given field of knowledge. There is a corresponding interest in the use of statistical estimation as a substitute, particularly various measures of effect size.

A significant advantage of this approach is the ease of comparison between different studies if the same metric is chosen. Another advantage involves the independence of the size of sample on effect size. In a hypothesis testing approach, increasing sample size may only increase the capacity of a study to detect tiny, possibly inconsequential effects. With effect size estimation, precision simply increases with increasing sample size; thus, the larger the sample the more confidence accrues that the resultant effect size measure is a true representation of the relationship between the relevant variables.

It is also worth noting that the effect size provides additional complementary information to the p level. Findings demonstrating high statistical significance but small effects may have little or no practical value for participants. Highly educationally significant findings may arise from large effect size, large sample size or both. Thus, the examination of effect sizes helps differentiate studies in which highly educationally significant results depend on a large sample size (i.e., a small real world effect) from studies in which (a) the existence of a large sample merely adds weight to the reliability of an estimated effect size, or (b) a large effect size occurs regardless of sample size.

Replicability is also an important component in any discussion of results. It is considered by Stanovich (1996) as one of the major hallmarks of any genuine claim to knowledge. True replicability is only provided when independent researchers, without a stake

in a positive outcome, produce similar findings in a different setting. Replication studies relieve doubts about generalisability of results, about idiosyncratic samples, and about experimenter bias.

Another approach to replication is internal replicability. Thompson (1996) nominates such analyses as cross validation, jackknife, and bootstrapping. These involve judgements about result stability across sample variations. In this thesis, it is possible to consider several such sub-groupings additional to the major sample of 206 students. The cohort may be considered as two separate groupings: Level A (experimental and control) as one grouping; and, Level B (experimental and control), as a further separate grouping. If similar outcomes are obtained in each independent analysis, then one may be a little more confident about generalisable results. Similarly, a third group comprises the experimental and control groups associated with the beginning reading program *Teach Your Child to Read in 100 Easy Lessons*.

It is also possible to consider groupings based on other membership criteria besides program. For example, effect sizes may be compared across schools, sex, and age groupings. The degree to which such secondary analyses inform the discussion is in dispute. Thompson (1996) concedes that their contribution is less substantial than those of independent replications. However, he argues that what they do add is non-trivial (even if inclined to be inflated), as opposed to that purely illusory contribution to replicability at times attributed to statistical significance testing. Robinson and Levin (1997) are less positive about the value of internal replicability analyses, pointing to the unavoidable limitations imposed by single sample characteristics, and by potential experimenter bias.

In this thesis, different students (of varied ages), in a number school settings, with various teachers, and on three similarly designed, but non-identical programs demonstrated comparable levels of improvement. This heterogeneity makes extraneous variables such as site effects or teacher effects easier to dismiss as alternative explanations of measured experimental effects (Cohen, 1990). It is relevant to note that membership of any one particular experimental cohort (other than program membership) was not predictive of outcome. Thus, it can be argued that the sample was not homogeneous in age, SES, sex, location, teacher, and school characteristics. As regards experimenter bias, the risks are relatively small in that the experimenter was not the teacher, and was only one of a number of test administrators and scorers; however, the experimenter was aware of whether the students assessed at any given time were experimental or control students.

Effect Size Calculation

The calculation of effect size d was based upon the ratio of the difference between the group means at pretest and posttest (separately for experimental and control groups) and the pooled standard deviation of that group at pretest and posttest. Pooled standard deviation was chosen because it more closely represents the population parameter (Hunter & Schmidt, 1990). The method of separate calculations was chosen because on occasions there were obvious differences between experimental and control groups at pretest (e.g., in TOPA and Spelling) in favour of the latter. In this case, the choice of the traditional measure (the ratio of the group mean difference at posttest and the standard deviation) would seriously underestimate the magnitude of the experimental effect, given the experimental group's initially lower scores. Additionally, the use of effect size separately for the control group provides acknowledgment that the control group was attending school, an environment in which reading related skills are expected to develop, and hence subject to an effect over time. The use of a second effect size score provides additional information to the traditional score, that is, the extent to which schooling alone adds to the development of the skills under analysis.

All effect sizes were calculated using the Hunter-Schmidt error correction procedure (Hunter & Schmidt, 1990) because it makes allowance for measurement error in the dependent variable. The authors argue that effect size attenuation occurs due to the use of less than perfect tests, a characteristic reflected in the standard deviation. Dividing the calculated effect size by the square root of the test reliability thus provides a truer picture of the size of effect. In practice, the alteration is not large, decreasing as test reliability increases.

Further Research

A consideration of the statistical sensitivity of the experiment suggests areas for improving control of error variance. Treatment variability is a major issue in experiments occurring across numerous sites and with numerous teachers. The reading program design (through the use of scripted teacher manuals) inhibits, but of itself cannot eliminate such variability. Both pre-program teacher training and within-program teacher monitoring are variables that could be manipulated in studies to investigate optimum cost-benefit ratios. Environmental factors may differ across sites - scheduled lesson time and duration, lesson frequency, class mix (e.g., drawn from one or several grades), and group size.

There are also potential error sources in the pretest and posttest procedures. A number of different trained testers were used, but no inter-rater reliability controls were attempted. Testing facilities varied markedly across schools, from one tester per room to several per room in some cases. Such uncontrolled factors make a contribution to the error term, and better organisation (within the limitations imposed by schools) may enable error reduction.

The preference of experimental reading measures over standardised tests is common in the research literature. The test consumers' need for quick, readily administered assessments can lead to instruments with small item pools and quite steep item gradients (Lovett, Barron, Forbes, Cuksts, & Steinbach, 1994). Such test characteristics reduce sensitivity to the detection of other than large gains, and it is only increasing the number of participants that small though possibly important gains may be demonstrated educationally. Fortunately, the gains were large enough to be detectable in the Word Attack test, despite its item gradient. The TOPA test too provided a problem, notably ceiling effects. The TOPA was chosen because of the support for oddity tasks as a focus for phonemic awareness assessment, and because of the availability of norms. However, the ceiling effect suggests that either the test include increasingly difficult items of the same type, or include another stage such as phoneme deletion, a more complex test of phonemic awareness.

The issue of multiple measures has previously been examined, but is acknowledged as a limitation in this study. For example, the variance on any one memory measure contains that associated with working memory, that associated with the task structure (e.g., instructions, materials) used to estimate the construct, and error. To minimise the second source of variance and maximise the first multiple measures of a construct are advisable. For example, Swanson and Alexander (1997) chose five different measures of working memory. Salthouse (1990) suggested that no single working memory measure can provide a true picture of working memory because of the influence of task specific factors.

Low subject variability is enabled in one sense in that students were selected in each of the programs on the basis of their performance on a reading test. However, there was no attempt to control for intelligence, or command of English. Many studies routinely discard students with measured intelligence level below IQ 80, with social-emotional difficulties, and those for whom English is a second language. Certainly subject variability could be further reduced by their exclusion. The program designers however argue that their programs' effectiveness is primarily related to design characteristics, and hence should be consistent across a range of learner differences. An analysis of the research on Direct Instruction

programs (Adams, 1996; Lockery & Maggs, 1982; White, 1988) provides support for this argument.

A source of individual differences not accounted for in this study was students for whom English was a second language (ESL). This information would have been helpful in determining the usefulness of this program for such a sub-group. In fact, the geographical area from which the students for this study originated is well known for the high proportion of such students, and the range of first languages represented. A post hoc attempt was made to investigate the proportion of ESL students making little or no gain from the program. Contact with teachers revealed that there was not a high proportion of ESL students among the no-progress group. In fact, teachers' subjective impressions were that such students tended to make substantial gains with the Direct Instruction programs. If this is so, perhaps the explanation lies in the possibility that their problems are not directly related to phonemic awareness, but rather to lack of experience with English words, a situation intensively addressed in the reading programs.

A future research focus entails a fine grained analysis of the components of the reading program. This includes the proportions of purely phonemic awareness (orally dominated activities) relative to activities with phoneme-grapheme involvement. In phonemic awareness training, activities have included word identity, rhyming, sound categorisation, tapping, blending and segmenting to name a few. Before being able to determine an optimum range and sequence of such activities a better understanding of the nature of phonemic awareness (and its relationship to other phonological processes) is required.

There is increasing acceptance that phonemic awareness is a general ability with several levels of complexity across a range of dimensions (Yopp, 1988), but there is not unanimity. There is still much to be discovered about the relationship between the tasks, for example, the degree to which differences in phonemic awareness tasks are due to extraneous task demands, such as memory processes. There are usually significant correlations between the various measures of phonemic awareness, the lack of a perfect correlation ascribed to the superimposition of additional task demands beyond that of a pure measure of phonemic awareness. Some (e.g., Wagner, Torgesen & colleagues) have attempted to partial-out these extraneous task demands by using multiple measures, and extracting the latent variable - one free of the contaminants including various sources of unreliability. Another interesting and related issue involves the relationship (if any) of the various phonological processes. Do the three constructs they propose - phonemic awareness, phonological memory, and lexical

access represent different abilities, or are they related in some way? Answers to each of these questions will play a role in enhancing understanding of the reading process, and in improving instruction, in particular, to at-risk students.

In this study, the improvements in decoding skills resulting from participation in the *Corrective Reading* program have been impressive whether assessed using visual inspection, program mastery tests, multivariate analysis, effect size, teacher and parent interview, or in comparison to norms. These effects were not constrained by age, sex or school attended. However, the study did not assess progress on real reading tasks. As discussed in Chapters 1 and 7, other studies have noted positive changes in various reading tasks, including comprehension, following the *Corrective Reading* program and other phonics emphasis programs. It is usually argued that the ability to decode previously refractory words leads to rapid, accurate, and effortless orthographic reading, when practice is adequate. When words are read effortlessly, attention to comprehension processes is maximised, thus enabling the student's entire oral language to be accessed, and consequent gains in assessed comprehension (Hoover & Gough, 1990). In this study, the story ends at the point of decoding skill gains. Further studies may examine whether students make use of these skills in everyday reading or prefer to return to partial letter cue, and context based guessing (McGuinness et al., 1995). Studies might also consider means of increasing volume of reading, ideally using controlled text mirroring the development of skills in the *Corrective Reading* program. For example, in the *100 Lessons* program it is possible to make use of Distar Library (Engelmann & Bruner, 1977), a set of readers using the same orthography as the reading program, as an adjunct. A chart for determining which book is appropriate for a given lesson is presented in Appendix C. Increasing appropriate practice opportunities is intended to enhance generalisation of reading skills to everyday reading, and to enable the development of orthographic images sufficient to allow sight word recognition.

Concluding Comments

This study with students who have experienced some years of reading failure adds to the scientific literature supporting the value of intensive systematic code-emphasis instruction. However, its implications extend beyond the validation of specific instructional procedures. The research occurred within a number of school settings, and the interventions were shown to be portable (i.e., effective despite inter-school differences), and viable (i.e., able to be incorporated into existing school structures and timetables). They also proved to be

inexpensive (group intervention; no specific funding), outcomes were visible beyond formal assessment and non-trivial (effects noticed by parents and home group teachers).

The content of the intervention was supported by past and recent theoretical evidence about the development of reading, and its underpinning skills, and by a dramatically rapid accretion of empirical evidence around phonological processing as a primary focus for intervention efforts. For the participating schools, the adoption of this reading program represents an ideal realised all too rarely in educational systems - the adoption and incorporation into school literacy policy of an intervention on the basis of its demonstrated effectiveness with the population it is designed to serve.

REFERENCES

- Ackerman, P. T., & Dykman, R. A. (1993). Phonological processes, confrontational naming, and immediate memory in dyslexia. *Journal of Learning Disabilities, 26*, 597-609.
- Adams, G., & Englemann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.
- Adams, M. (1996). Are decoding and literature-based programs incompatible? In L. R. Putnam (Ed.), *How to become a better reading teacher: Strategies for assessment and intervention*, (pp. 183-190). Hillsdale, NJ: Prentice Hall.
- Adams, M. J. (1990). *Beginning to read: Thinking & learning about print..* Cambridge, MA: MIT Press.
- Adams, M. J. (1991). Beginning to read: A critique by literacy professionals and a response by Marilyn Jager Adams. *The Reading Teacher, 44*, 370-395.
- Adams, M. J., & Bruck, M. (1993). Word recognition: The interface of educational policies and scientific research. *Reading and Writing: An Inter-disciplinary Journal, 5*, 113-139.
- Alexander, A. W., Anderson, H. G., Heilman, P. C., Voeller, K. K. S., & Torgesen, J. K. (1991). Phonological awareness training and remediation of analytic decoding deficits in a group of severe dyslexics. *Annals of Dyslexia, 41*, 193-206.
- Allington, R. L. (1984). Content coverage and contextual reading in reading groups. *Journal of Reading Behaviour, 16*, 85-96.
- Allington, R. L., Stuetzel, H., Shake, M., & Lamarche, S. (1986). What is remedial reading? A descriptive study. *Reading Research and Instruction, 26* (1), 15-30.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychological Association. (1993, October). *Final report of the task force on promotion and dissemination of psychological procedures*. New York: American Psychological Association, Division of Clinical Psychologists (Division 12).
- Anastasi, A. (1988). *Psychological testing (6th ed.)*. New York: Harper & Row.
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Champaign, IL: The

- Centre for the Study of Reading.
- Arter, J. A., & Jenkins, J. R. V. (1979). Differential diagnosis-prescriptive teaching: A critical appraisal. *Review of Educational Research, 49*, 517-555.
- Ashman, A., & Elkins, J. (1990). *Educating children with special needs*. New York: Prentice Hall.
- Australian Government House of Representatives Enquiry. (1993). *The Literacy challenge*. Canberra: Australian Printing Office.
- Badger, L. (1984). Providing experiences for reading development. In *Early literacy inservice course* (Unit 5, pp. 19-25). South Australia: Education Department of South Australia.
- Badian, N. A. (1993). Phonemic awareness, naming, visual symbol processing, and reading. *Reading & Writing: An Interdisciplinary Journal, 5*, 87-100.
- Badian, N. A. (1994). Preschool prediction: Orthographic and phonological skills and reading. *Annals of Dyslexia, 44*, 3-25.
- Baker, S. K., Kameenui, E. J., Simmons, D. C., & Stahl, S. A. (1994). Beginning reading: Educational tools for diverse learners. *School Psychology Review, 23*, 372-391.
- Ball, E. W. (1993). Phonological awareness: What's important and to whom? *Reading & Writing: An Interdisciplinary Journal, 5*, 141-159.
- Ball, E. W., & Blachman, B. A. (1988). Phoneme segmentation training: Effect on reading readiness. *Annals of Dyslexia, 38*, 208-225.
- Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition & developmental spelling. *Reading Research Quarterly, 25*, 49-66.
- Barker, T. A., & Torgesen, J. K. (1995). An evaluation of computer-assisted instruction in phonological awareness with below-average readers. *Journal of Computing Research, 13*, 89-103.
- Barmby, S., Bonham, D., Lawry, F., & Nissner, H. (1986). *Real literacy: Shifting the focus..* Victoria, Australia: Maroondah Region Publication.
- Barnes, D. (1985). Why not Direct Instruction? *The Australian Educational & Developmental Psychologist, 1*, 59-62.
- Bateman, B. (1991). Teaching word recognition to slow-learning children. *Reading, Writing and Learning Disabilities, 7*, 1-16.

- Bay, M., Staver, J. R., Bryan, T., Hale, J. B. (1992). Science instruction for the mildly handicapped: Direct instruction versus discovery teaching. *Journal of Research in Science Teaching*, 29, 555-570.
- Becker, W. C. (1977). Teaching reading and language to the disadvantaged. What we have learned from field research. *Harvard Educational Review*, 47, 518-543.
- Becker, W. C. (1986). *Applied psychology for teachers: A behavioural-cognitive approach*. Chicago, USA: Science Research Associates.
- Becker, W. C. (Ed.).(1988). Direct Instruction: A general case for teaching a general case.[Special issue]. *Education and Treatment of Children*, 11, 301-396.
- Becker, W. C., & Gersten, R. (1982). A follow-up of Follow Through: The later effects of the Direct Instruction model on children in fifth & sixth grades. *American Educational Research Journal*, 19, 75-92.
- Bentin, S., & Leshem, H. (1993). On the interaction between phonological awareness and reading acquisition: It's a two-way street. *Annals of Dyslexia*, 43, 125-148.
- Bentin, S., Hammer, R., & Cahan, S. (1991). The effect of aging and first grade schooling on the development of phonological awareness. *Psychological Science*, 2, 271-274.
- Bereiter, C., & Kurland, M. (1981). A constructive look at Follow Through results. *Interchange*, 12, 1-22.
- Berninger, V. W., Thalberg, S. P., DeBruyn, I., & Smith, R. (1987). Preventing reading disabilities by assessing and remediating phonemic skills. *School Psychology Review*, 16, 554-565.
- Berninger, V. W., & Abbott, R. D. (1994). Redefining learning disabilities. Moving beyond aptitude-achievement discrepancies to failure to respond to validated treatment protocols. In G. Reid Lyon (Ed), *Frames of reference for the assessment of learning disabilities. New views on measurement issues* (pp. 163-184). MD: Brooks Publishing.
- Binder, C., & Watkins, C. L. (1990). Precision Teaching and Direct Instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly*, 3 (4), 74-96.
- Bishop, D. & Butterworth, G. (1980). Verbal-performance discrepancies: Relationship to both risk and specific reading retardation. *Cortex*, 16, 375-389.
- Blachman, B. A. (1984). Language analysis skills and early reading acquisition. In G. P. Wallach, & K. G. Butler (Eds.), *Language learning disabilities in school age children* (pp. 271-287). Baltimore: Williams & Wilkins.

- Blachman, B. A. (1991). Early intervention for children's reading problems: Clinical applications of the research in phonological awareness. *Topics in Language Disorders, 12* (1), 51-65.
- Blachman, B. A. (1994). What we have learned from longitudinal studies of phonological processing and reading, and some unanswered questions: A response to Torgesen, Wagner, & Rashotte. *Journal of Learning Disabilities, 27*, 287-291.
- Blachman, B. A., Ball, E. W., Black, R. S., & Tangel, D. M. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: Does it make a difference? *Reading & Writing: An Interdisciplinary Journal, 6* (1), 1-18.
- Bond, G. L., & Dykstra, R. (1967). The co-operative research program in first grade reading instruction. *Reading Research Quarterly, 2*, 5-142.
- Boomer, G. (1988). Standards & literacy. Two hundred years on the road to literacy: Where to from here? *Directions: Literacy*. Supplement to *Education Victoria*., June. Melbourne, Australia: State Board of Education.
- Bowers, P. G. (1995). Tracing symbol naming speeds unique contributions to reading disabilities over time. *Reading and Writing: An Interdisciplinary Journal, 7*, 189-216.
- Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability: Multiple measures of a singular process. *Journal of Experimental Child Psychology, 51*, 195-219.
- Bowey, J. A. (1996). Phonological sensitivity as a proximal contributor to phonological recoding skills in children's reading. *Australian Journal of Psychology, 48*(3), 113-118.
- Bowey, J. A., & Francis, J. (1991). Phonological analysis as a function of age and exposure to reading instruction. *Applied Psycholinguistics, 12*, 91-121.
- Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design study of phonological skills underlying Fourth-Grade children's word reading difficulties. *Child Development, 63*, 999-1011.
- Bradley, L. (1990). Rhyming connections in learning to read and spell. In P. D. Pumphrey, & C. D. Elliot (Eds.), *Children's difficulties in reading, spelling and writing: Challenges and responses* (pp. 83-100). London: The Falmer Press.
- Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read - A causal connection. *Nature, 301*, 419-421.

- Brady, S. (1991). The role of working memory in reading disability. In S. Brady, & D. Shankweiler, (Eds.), *Phonological processes in literacy* (pp. 129-152). Hillsdale, NJ: Erlbaum.
- Brady, S. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia*, 44, 26-59.
- Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city children. *Annals of Dyslexia*, 44, 26-59.
- Branwhite, A. B. (1983). Boosting reading skills by direct instruction. *British Journal of Educational Psychology*, 53, 291-298.
- Brigance A. H. (1992). *Comprehensive Inventory of Basic Skills*. Melbourne, Australia: Hawker Brownlow.
- Brown, I. S., & Felton, R. H. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. *Reading & Writing: An Interdisciplinary Journal*, 2, 223-241.
- Bruck, M. (1988). The word recognition and spelling of dyslexic children. *Reading Research Quarterly*, 23, 51-69.
- Bruck, M. (1990). Word-recognition skill of adults with childhood diagnoses of dyslexia. *Developmental Psychology*, 26 (3), 439-454.
- Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology*, 28, 874-886.
- Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Child Psychology*, 50, 156-178.
- Bruck, M., & Treiman, R. (1992). Learning to pronounce words: The limitations of analogies. *Reading Research Quarterly*, 27, 374-389.
- Bryant, P. (1990). Phonological development and reading. In P. D. Pumphrey, & C. D. Elliot (Eds.), *Children's difficulties in reading, spelling and writing: Challenges and responses* (pp. 63-82). London: The Falmer Press.
- Bryant, P. E., Bradley, L., Maclean, M., & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language*, 16, 407-428.
- Bryant, P., & Bradley, L. (1985). *Children's reading problems: Psychology & education*. UK: Basil Blackwell.
- Bryen, D. N., & Gerber, A. (1987). Metalinguistic abilities and reading: A focus on

- phonological awareness. *Reading, Writing & Learning Disabilities*, 3, 357-367.
- Burt, J. S., & Butterworth, P. (1996). Spelling in adults: Orthographic transparency, learning new letter string and reading accuracy. *European Journal of Cognitive Psychology*, 8(1), 3-43.
- Byrne, B. (1991). The role of phonological awareness in reading acquisition. *Australian Journal of Reading*, 2, 133-139.
- Byrne, B. (1996). The learnability of the alphabetic principle: Children's initial hypotheses about how print represents spoken speech. *Applied Psycholinguistics*, 17, 401-426.
- Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81, 313-321.
- Byrne, B., & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology*, 82, 805-812.
- Byrne, B., & Fielding-Barnsley, R. (1991a). *Sound Foundations*. Artamon, Australia: Leyden Educational Publishers.
- Byrne, B., & Fielding-Barnsley, R. (1991b). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, 83, 451-455.
- Byrne, B., & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A one-year follow-up. *Journal of Educational Psychology*, 85, 104-111.
- Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3- year follow-up and a new preschool trial. *Journal of Educational Psychology*, 87, 488-503.
- Byrne, B., Fielding-Barnsley, R., & Ashley, L. (1996). What does the child bring to the task of learning to read? A summary of the New England Reading Acquisition project. *Australian Journal of Psychology*, 48 (3), 119-123.
- Cairney, T. H., Lowe, K., McKenzie, P., & Petrakis, D. (1993). *Investigating literacy in correctional institutions*. NSW, Australia: Department of Employment, Education and Training.
- Cambourne, B. (1979). How important is theory to the reading teacher? *Australian Journal of Reading*, 2, 78-90.

- Campbell, M. L. (1983). *Corrective Reading* program evaluated with secondary students in San Diego. *Direct Instruction News*, pp. 1, 23.
- Cantwell, A., & Rubin, H. (1992). Object naming ability of adults with written language difficulties. *Annals of Dyslexia*, 42, 179-195.
- Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology*, 55, 1-30.
- Carbo, M. (1988). Debunking the great phonics myth. *Phi Delta Kappan*, 70, 226-40.
- Carbo, M. (1992). Giving unequal learners an equal chance: A reply to a biased critique of learning styles. *Remedial and Special Education*, 13(1), 19-29.
- Carnine, D. (1982). *Approaches to reading: Psycholinguistics and Direct Instruction methods*. Paper presented at the Conference of the Australian Association for Direct Instruction, Kuringai CAE, NSW.
- Carnine, D. (1991). Curricular interventions for teaching higher order thinking to all students: Introduction to the special series. *Journal of Learning Disabilities*, 24, 261-269.
- Carnine, D. (1995). Trustworthiness, useability, and accessibility of educational research. *Journal of Behavioral Education*, 5, 251-258.
- Carnine, D., & Silbert, J. (1979). *Direct instruction reading*. Columbus, OH: Charles Merrill.
- Casazza, M. E. (1993). Using a model of direct instruction to teach summary writing in a college reading class. *Journal of Reading*, 37, 202-208.
- Castle, J. M., Riach, J., & Nicholson, T. (1994). Getting off to a better start in reading and spelling: The effects of phonemic awareness instruction within a whole language program. *Journal of Educational Psychology*, 86, 350-359.
- Castles, I. (1994). *SEIFA: Socio-economic indexes for areas*. Canberra: Australian Government Printer.
- Catts, H. W. (1991). Early identification of reading disabilities. *Topics in Language Disorders*, 12 (1), 1-16.
- Catts, H. W. (1996). Defining dyslexia as a developmental language disorder: An expanded view. *Topics in Language Disorders*, 16(2), 14-29.
- Catts, H. W., & Vartiainen, T. (1993). *Sounds Abound*. East Moline, Illinois: LinguSystems.
- Ceci, S. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, 27, 703-722.

- Chall, J. S. (1967). *The great debate*. New York: McGraw Hill.
- Chall, J. S. (1979). The great debate: Ten years later, with a modest proposal for reading stages. In L. B. Resnick, & P. A. Weaver (Eds.), *Theory & Practice in Early Reading*. Vol. 1, pp. 29-54. Hillsdale, NJ: Lawrence Erlbaum.
- Chall, J. S. (1987). Reading & early childhood education: The critical issues. *Principal*, 66, 6-9.
- Chall, J. S. (1989). Learning to read: The great debate 20 years later - a response to "Debunking the great phonics myth. *Phi Delta Kappan*, 70, 521-538.
- Chamberlain, L. A. (1987). Using Direct Instruction in a Victoria, BC resource room. *ADI News*, 7 (1), 7-9.
- Chaney, C. (1990). Evaluating the Whole Language approach to language arts: The pros and cons. *Language, Speech, and Hearing in Schools*, 21, 244-249.
- Clunies-Ross, G. (1990). Some effects of direct instruction in comprehension skills with sixth grade students. *Behaviour Change*, 7, 84-89.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.), Hillsdale, NJ: Lawrence Earlbaum.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Cohen, J. (1994). *The earth is round* ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cole, P., & Chan, L. (1990). *Methods and strategies for special education*. Sydney: Prentice Hall.
- Coltheart, M. (1983). Phonological awareness: A preschool precursor of success in reading. *Nature*, 301, 421.
- Conway, R. (1994, May 6). Time to test students. *Herald Sun*, p.15.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi experimentation. Design and analysis issues for field settings*. Boston: Houghton-Mifflin.
- Cornelissen, P. L., Hansen, P. C., Bradley, L., Stein, J. F. (1996). Analysis of perceptual confusions between nine sets of consonant-vowel sound in normal and dyslexic adults. *Cognition*, 59, 275-306.
- Cornwall, A. (1992). The relationship of phonological awareness, rapid naming and verbal memory to severe reading and spelling disability. *Journal of Learning Disabilities*, 25, 532-538.

- Crowder, R., & Wagner, R. (1992). *The psychology of reading: An introduction*. New York: Oxford.
- Cunningham, A. (1990). Explicit vs implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, 50, 429-444.
- Darch, C. (1993). Direct Instruction: A research-based approach for designing instructional programs. In R. C. Greaves & P. J. McLaughlin (Eds), *Recent advances in special education and rehabilitation*. (pp 88-106). Boston: Andover Medical.
- Davidson, M., & Jenkins, J. R. (1994). Effects of phonemic processes on word reading and spelling. *Journal of Educational Research*, 87, 148-157.
- Davis, A. (1988). A historical perspective. In J. Estill Alexander (Ed.), *Teaching reading* (3rd ed., p. 532-553). USA: Scott, Foresman & Co.
- Davis, J. M., & Spring, C. (1990). The Digit Naming Speed Test: Its power and incremental validity in identifying children with specific reading disabilities. *Psychology in the Schools*, 27, 15-22.
- Davis, W. (1973). *Teaching reading in early England*. London: Pitman.
- Delpit, L. D. (1988). The silenced dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review*, 58, 280-298.
- Denckla, M. B., & Rudel, R. (1976). Rapid automatised naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471-479.
- Department of Employment, Education, and Training. (1991). *Literacy and industry update, 1*, NSW, Australia: DEET.
- Dugard, P., & Todman, J. (1995). Analysis of pretest and posttest control group designs in educational research. *Educational Psychology*, 15, 181-198.
- Dunn, R., Beaudry, J. S., & Klavas, A. (1989). Survey of research on learning styles. *Educational Leadership*, March, 50-58.
- Edelsky, C. (1990). Whose agenda is this anyway? A response to McKenna, Robinson & Miller. *Educational Researcher*, 19, 7-11.
- Eden, G. F., Stein, J. F., Wood, H. M., & Wood, F. B. (1995). Verbal and visual problems in reading disability. *Journal of Learning Disabilities*, 28, 272-290.
- Education Department of South Australia. (1984). *Early literacy inservice course*. South Australia.
- Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior*, 19, 5-31.

- Ehri, L. C. (1993). How English orthography influences phonological knowledge as children learn to read and spell. In R. J. Scholes (Ed.), *Literacy and Language Analysis*, (pp. 21-43). Hillsdale, NJ: Erlbaum.
- Ehri, L. C. (1994). Development of the ability to read words: Update. In R. Ruddell, M. Ruddell, & H. Singer (Eds.). *Theoretical models and processes of reading (4th ed., pp. 323-358)*. Newark, DE: International Reading Association.
- Ehri, L. C. (1995). Phases of development in learning to read word by sight. *Journal of Research in Reading*, 18(2), 116-125.
- Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly*, 27, 13-26.
- Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and in the phonological representation of lexical items. *Annals of Dyslexia*, 44, 205-226.
- Eldredge, J. L. (1991). An experiment with a modified whole language approach in first grade classrooms. *Reading Research & Instruction*, 30 (3), 21-38.
- Eldredge, J. L., Quinn, B., & Butterfield, D. D. (1990). Causal relationships between phonics, reading comprehension, and vocabulary achievement in the second grade. *Journal of Educational Research*, 83, 201-214.
- Engelmann, S. (1980, Feb). Toward the design of faultless instruction: The theoretical basis of concept analysis. *Educational Technology*, 28-36.
- Engelmann, S. (1988). The logic and facts of effective supervision. *Education. & Treatment of Children*, 11 (4), 328-340.
- Engelmann, S., & Bruner, E. C. (1977). *Distar Library*. Chicago: Science Research Associates
- Engelmann, S., & Bruner, E. C. (1988). *Reading Mastery*. Chicago: Science Research Associates.
- Engelmann, S., & Carnine, D. (1982). *Theory of Instruction*. New York: Irvington.
- Engelmann, S., Becker, W. C., Carnine, D., & Gersten, R. (1988). The Direct Instruction Follow Through model: Design and outcomes. *Education and Treatment of Children*, 11(4), 303-317.
- Engelmann, S., Haddox, P., & Bruner, E. (1983). *Teach Your Child To Read In 100 Easy Lessons*. New York: Simon & Schuster.

- Engelmann, S., Hanner, S., & Johnson, G. (1988). *Corrective Reading-Series Guide*. Chicago: Science Research Associates.
- Engelmann, S., Hanner, S., & Johnson, G. (1988). *Corrective Reading: Decoding A*. Chicago: Science Research Associates.
- Engelmann, S., Johnson, G., Becker, W., Meyer, L., Carnine, L., & Becker, J. (1978). *Corrective Reading: Decoding B*. Chicago: Science Research Associates.
- Fabre, T. (1984). *The application of direct instruction in special education: An annotated bibliography*. Unpublished manuscript, University of Oregon.
- Felton, R. H. (1992). Early identification of children at risk for reading disabilities. *Topics in Early Childhood Special Education, 12*, 212-229.
- Felton, R. H. (1993). Effects of instruction on the decoding skills of children with phonological-processing problems. *Journal of Learning Disabilities, 26* (4), 583-589.
- Felton, R. H., & Brown, I. S. (1990). Phonological processes as predictors of specific reading skills in children at risk for reading failure. *Reading & Writing: An Interdisciplinary Journal, 2*, 39-59.
- Felton, R. H., & Pepper, P. P. (1995). Early identification and intervention of phonological deficit in kindergarten and early elementary children at risk for reading disability. *School Psychology Review, 24*, 405-414.
- Felton, R. H., & Wood, F. B. (1989). Cognitive deficits in reading disability and attention deficit disorder. *Journal of Learning Disabilities, 22*, 3-13.
- Fields, B. A. (1986). Direct Instruction and the practicality ethic: A case study and position in curriculum innovation. *Australian Journal of Remedial Education, 18*(1), 17-22.
- Fields, B. A., & Kempe, A. (1992). Corrective feedback in whole language teaching: Implications for children with learning problems. *Australian Journal of Special Education, 16*(2), 22-31.
- Fister, S., & Kemp, K. (1993). Translating research: Classroom application of validated instructional strategies. In R. C. Greaves, & P. J. McLaughlin (Eds), *Recent advances in special education and rehabilitation*. (pp 107-126). Boston: Andover Medical.
- Flesch, R. (1955). *Why Johnny can't read*. New York: Harper & Row.
- Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1993). Classification of learning disabilities. Relation to other childhood disorders. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding of*

- learning disabilities: New views from research and their implications for education and public policies.*(p.153-170). Baltimore: Brooks.
- Flowers, D. L. (1993). Brain basis of dyslexia: A summary of work in progress. *Journal of Learning Disabilities, 26*, 575-582.
- Foorman, B. R. (1995). Research on "the Great Debate": Code-oriented versus Whole Language approaches to reading instruction. *School Psychology Review, 24*, 376-392.
- Foorman, B. R., & Francis, D. J. (1994). Exploring connections among reading, spelling, and phonemic segmentation during first grade. *Reading and Writing: An Interdisciplinary Journal, 6*, 65-91.
- Foorman, B. R., Francis, D. J., Beeler, T., Winikates, D., & Fletcher, J. (1997). Early interventions for children with reading problems: Study designs and preliminary findings. *Learning Disabilities: A Multidisciplinary Journal, 8*, 63-71.
- Foorman, B. R., Francis, D., Novy, D., & Liberman, D. (1991). How letter-sound instruction mediates progress in first grade reading and spelling. *Journal of Educational Psychology, 83*, 456-469.
- Francis, W. N. (1965). *The English Language: An introduction*. New York: Norton.
- Freebody, P., & Byrne, B. (1988). Word-reading strategy in elementary school children: Relation to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly, 23*, 441-453.

- Gable, R. A., & Warren, S. F. (1993). The enduring value of instructional research. In Robert Gable & Steven Warren (Eds.), *Advances in mental retardation and developmental disabilities: Strategies for teaching students with mild to severe mental retardation*. (pp. 1-7). Philadelphia: Jessica Kingsley.
- Gaskins, I. W., Ehri, L. C., Cress, C., O'Hara, C., Donnelly, K. (1996). Procedures for word learning: Making discoveries about words. *The Reading Teacher*, 50 (4), 312-327.
- Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology*, 82, 387-406.
- Gersten, R. (1985). Direct Instruction with special education students: A review of evaluation research. *Journal of Special Education*, 19(1), 41-58.
- Gersten, R., & Dimino, J. (1993). Visions & revisions: A special education perspective on the whole language controversy. *Remedial & Special Education*, 14 (4), 5-13.
- Gersten, R., & Guskey, T. (1985, Fall). Transforming teacher reluctance into a commitment to innovation. *Direct Instruction News*, 11-12.
- Gersten, R., Keating, T., & Becker, W. (1988). The continued impact of the Direct Instruction model: Longitudinal studies of Follow Through students. *Education and Treatment of Children*, 11 (4), 318-327.
- Gersten, R., Woodward, J., & Darch, C. (1986). Direct Instruction: A research based approach to curriculum design and teaching. *Exceptional Children*, 53 (1), 17-31.
- Gettinger, M. (1993). Effects of invented spelling and direct instruction on spelling performance of second-grade boys. *Journal of Applied Behaviour Analysis*, 26, 281-291.
- Gibbs, S. (1996). Categorical speech perception and phonological awareness in the early stages of learning to read. *Language & Communication*, 16 (1), 37-60.
- Gibson, E., & Levin, H. (1978). *The psychology of reading*. MA: MIT.
- Giffen, D. (1980). A letter-a-reply - an invitation to respond. *Reading Around*, 8, 84-88.
- Gillam, R. B., & Van Kleeck, A. (1996). Phonological awareness training and short-term working memory: Clinical implication. *Topics in Language Disorders*, 17 (1), 72-81.
- Glang, A., & Gersten, R. (1987, Winter). Coaching teachers. *Direct Instruction News*, pp. 1, 4, 5, 7.

- Glang, A., Singer, G., Cooley, E., & Tish, N. (1992). Tailoring direct instruction techniques for use with elementary students with brain injury. *Journal of Head Trauma Rehabilitation*, 7, 93-108.
- Glynn, T., Bethune, N., Crooks, T., Ballard, K., & Smith, J. (1992). Reading Recovery in context: Implementation and outcome. *Educational Psychology*, 12, 249-261.
- Gollash, F. (1980). A review of Distar Reading I from a psycholinguistic viewpoint. *Reading Education*, 5, 78-86.
- Good, T. L., & Brophy, J. E. (1987). *Looking in classrooms* (4th ed). New York: Harper & Row.
- Goodman, K. S. (1979). Reading: A psycholinguistic guessing game. In H. Singer, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. (pp. 259-271). Newark, DE: International Reading Association.
- Goodman, K. S. (1986). *What's whole in whole language*. Richmond Hill, Ontario: Scholastic.
- Goodman, K. S. (1989, March). Whole Language is whole: A response to Heymsfeld. *Educational Leadership*, 69-72.
- Goodman, K. S. (1997, Sept. 24). Capturing "America Reads" for a larger agenda, *Education Week*, p. 14.
- Goodman, Y. M. (1991). Beginning to read: A critique by literacy professionals and a response by Marilyn Jager Adams. *The Reading Teacher*, 44, 375-378
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. UK: Erlbaum.
- Goulandris, N., & Snowling, M. (1995). Assessing reading skills. In E. Funnell, & M. Stuart (Eds.), *Learning to read: Psychology in the classroom*. Cambridge, MA: Blackwell.
- Greaves, R. C., & McLaughlin, P. J. (Eds.). (1993). *Recent advances in special education and rehabilitation*. Boston: Andover Medical.
- Gregory, R. P. (1983). Direct Instruction, disadvantaged and handicapped children: A review of the literature and some practical implications. Part 1. *Remedial Education*, 18, 108-114.
- Gregory, R. P., Hackney, C., & Gregory, N. M. (1982a). *Corrective Reading* program: An evaluation. *British Journal of Educational Psychology*, 52, 33-50.
- Gregory, R. P., Hackney, C., & Gregory, N. M. (1982b). *Corrective Reading* program: The use of educational technology in a secondary school. *Reading Education*, 2 (2), 21-25.

- Griffith, P., & Olson, M. W. (1992). Phonemic awareness helps beginning readers break the code. *The Reading Teacher, 45*, 516-523.
- Groff, P. (1990). An analysis of the debate: Teaching reading without conveying phonics information. *Interchange, 21* (4), 1-14.
- Groff, P. (1991). Word recognition and critical reading. *Reading, Writing, and Learning Disabilities, 7*, 17-31.
- Grossen, B. (1995). *The research base for Reading Mastery, SRA*. [On-line]. Available: <http://darkwing.Uoregon.Edu/~adrep/rdgtxt.htm>.
- Grossen, B., & Carnine, D. (1990). Diagramming a logic strategy: Effects on difficult problem types and transfer. *Learning Disability Quarterly, 13*, 168-182.
- Grossen, B., & Carnine, D. (1990). Translating research on initial reading instruction into classroom practice. *Interchange, 21* (4), 15-23.
- Grossen, B., & Kelly, B. F. (1992). The effectiveness of Direct Instruction in a third-world context. *International Review of Education, 38* (1), 81-85.
- Hammond, G. (1996). The objections to null hypothesis testing as a means of analysing psychological data. *Australian Journal of Psychology, 48*, 104-106.
- Hanner, S., & Engelmann, S. (1984, May). Learner verification for the *Corrective Reading* program. *AADI*
- Harper, G. F., Mallette, B., Maheady, L., & Brennan, G. (1993). Classwide student tutoring teams and Direct Instruction as a combined instructional program to teach generalizable strategies for mathematics word problems. *Education & Treatment of Children, 16*, 115-134.
- Hatcher, P., Hulme, C., & Ellis, A. (1994). Ameliorating reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41-57.
- Hempenstall, K. (1988). *Evaluation of Direct Instruction program, Healesville High School*. Unpublished manuscript. Royal Melbourne Institute of Technology.
- Hempenstall, K. (1990a). Direct Instruction: *Experiences of an educational psychologist*. Paper presented at the 13th National Conference of the Australian Behaviour Modification Association. Melbourne, Australia.
- Hempenstall, K. (1990b, July,). School failure: A debilitating condition. *AADI Newsletter*, p.4.
- Hempenstall, K. (1994, Oct. 11). Words should be heard and seen. *The Age*, p.18.

- Hempenstall, K. (1995a). *The Picture Naming Test*. Unpublished manuscript. Royal Melbourne Institute of Technology.
- Hempenstall, K. (1995b, Mar. 28). Tackling the guarantee of early failure at reading. *The Age*, p. 8.
- Hempenstall, K. (1995c). *Brigance Spelling: Reliability*. Unpublished manuscript. Royal Melbourne Institute of Technology.
- Hempenstall, K. (1996). The gulf between educational research and policy: The example of Direct Instruction and Whole Language. *Behaviour Change*, 13, 33-46.
- Hendrickson, J. M., & Frank, A. R. (1993). Engagement & performance feedback: Enhancing the classroom achievement of students with mild mental disabilities. In R. A. Gable, & S. F. Warren (Eds.), *Advances in mental retardation and developmental disabilities: Strategies for teaching students with mild to severe mental retardation*. Philadelphia: Jessica Kingsley.
- Henry, M. K. (1993). The role of decoding in reading research and instruction. *Reading and Writing: An Interdisciplinary Journal*, 5, 105-112.
- Herr, C. (1989). Using *Corrective Reading* with adults. *ADI News*, 8(2), 18-21.
- Heshusius, L. (1991). Curriculum based assessment and direct instruction: Critical reflections on fundamental assumptions. *Exceptional Children*, 57, 315-328.
- Heymsfeld, C. R. (1989, March). Filling the hole in whole language. *Educational Leadership*, 65-68.
- Heymsfeld, C. R. (1992). The remedial child in the whole-language cooperative classroom. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 8, 257-273.
- Hill, P. (1995). *School effectiveness and improvement: Present realities and future possibilities*. Dean's Lecture: Paper presented at Melbourne University, May 24. Available at http://www.edfac.unimelb.edu.au/Seminars/dean_lec/list.html
- Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading and Writing: An Inter-Disciplinary Journal*, 7, 171-188.
- Holdsworth, P. (1984). *Corrective Reading* tested in UK. *ADI News*, 4 (2), 1-4.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Hornsby, D., Sukarna, D., & Parry, J. (1986). *Read on: A conference approach to reading*. Sydney: Martin Educational.

- House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. (1978). No simple answer: Critique of the Follow Through evaluation. *Harvard Educational Review, 48*, 128-160.
- Huang, H. S., & Hanley, R. (1994). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*, 73-98
- Hudson, A. (1997). Empirically validated interventions in education: A comment on King. *Behaviour Change, 14*, 25-28.
- Hulme, C., & Roodenrys, S. (1995). Practitioner review: Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry, 36*, 373-398.
- Hulme, C., & Snowling, M. (1992). Phonological deficits in dyslexia: A "Sound" reappraisal of the verbal deficit hypothesis? In N. Singh, & I. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 270-301). New York: Springer Verlag.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Hurford, D. P., Darrow, L., Edwards, T., Howerton, C., Mote, C., Schauf, J., & Caffey, P. (1993). An examination of phonemic processing abilities in children during their first grade year. *Journal of Learning Disabilities, 26* (3), 166-177.
- Hurford, D. P., Schauf, J. D., Bunce, L., Blauch, T., & Moore, K. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities, 27*, 371-382.
- Indrisano, R., & Chall, J. S. (1995). Literacy development. *Journal of Education, 177*, 63-83.
- Iversen, S., & Tunmer, W. E. (1993). Phonological processing and the Reading Recovery program. *Journal of Educational Psychology, 85*, 112-126.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Johnson, B., & Stone, E. (1991). Is whole language restructuring our classroom? *Contemporary Education, 62*, 102-104.
- Jorm, A. F., & Share, D. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics, 4*, 103-147.
- Jorm, A. F., Share, D., McLean, R., & Matthews, R. (1984). Phonological recoding and learning to read: A longitudinal study. *Applied Psycholinguistics, 5*, 201-207.

- Joyce, B., Weil, M., & Showers, B. (1992). *Models of teaching* (4th ed.). Massachusetts: Allyn & Bacon.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.
- Juel, C. (1993). The spelling-sound code in reading. In S. Yussen, & M. Smith (Eds.), *Reading across the life span* (pp. 95-109). New York: Springer-Verlag
- Kameenui, E. J., & Shannon, P. (1988). Point/counterpoint: Direct instruction reconsidered. In J. E. Readence & R. S. Baldwin (Eds.), *Thirty-seventh Yearbook of the National Reading Conference* (pp. 36-41). Chicago: National Reading Conference.
- Kameenui, E. J., & Simmons, D. C. (1990). *Designing instructional strategies: The prevention of academic learning problems*. Columbus, OH: Merrill.
- Kassendorf, S. J., & McQuaid, P. (1987). *Corrective Reading* evaluation study. *Direct Instruction News*, Fall, p. 9.
- Kavale, K. (1990). Variances & verities in learning disability interventions. In T. Scruggs, & B. Wong (Eds.), *Intervention research in learning disabilities*. (pp 3-33). New York: Springer Verlag.
- Kennedy, C. H. (1992). Trends in the measurement of social validity. *The Behavior Analyst, 15*, 147-156.
- Kinder, D., & Carnine, D. (1991). Direct Instruction: What it is and what it is becoming. *Journal of Behavioural Education, 1*, 193-213.
- Kirk, S. A., & Bateman, B. (1962). Diagnosis and remediation of learning disabilities. *Exceptional Children, 29*, 73-78.
- Kirsch, I., & Guthrie, J. (1984). Adult reading practices for work & leisure. *Adult Education Quarterly, 34*, 213-232.
- Kirtley, C., Bryant, P. E., Maclean, M. J., & Bradley, L. L. (1989). Rhyme, rime and the onset of reading. *Journal of Experimental Child Psychology, 48*, 224-245.
- Klesius, J. P., Griffith, P. L., & Zielonka, P. (1991). A whole language and traditional instruction comparison: Overall effectiveness and development of the alphabetic principle. *Reading Research & Instruction, 30* (2), 47-61.
- Kuder, S. J. (1990). Effectiveness of the DISTAR reading program for children with learning disabilities. *Journal of Learning Disabilities, 23*(1), 69-71.
- Kuder, S. J. (1991). Language abilities and progress in a Direct Instruction reading program for students with learning disabilities. *Journal of Learning Disabilities, 24*, 124-127.

- Lamb, S. J., & Gregory, A. H. (1993). The relationship between music and reading in beginning readers. *Educational Psychology, 13*, 19-27.
- Landerl, K., Frith, U., & Wimmer, H. (1996). Intrusion of orthographic knowledge on phoneme awareness: Strong in normal readers, weak in dyslexic readers. *Applied Psycholinguistics, 17* (1), 1-14.
- Lehto, J., (1995). Working memory and school achievement in the ninth form. *Educational Psychology, 15*(3), 271-281.
- Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology, 55*, 297-328.
- Leonard, C. M., Lombardino, L. J., Mercado, L. R., Browd, S. R., Breier, J. I., & Agee, O. F. (1996). Cerebral asymmetry and cognitive development in children: A magnetic resonance imaging study. *Psychological Science, 7* (2), 89-95.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*, 431-461.
- Liberman, I. Y., & Liberman, A. M. (1990). Whole language vs code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia, 40*, 51-75.
- Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education, 6*, 8-17.
- Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). Ann Arbor: University of Michigan Press.
- Lindamood, C., & Lindamood, P. C. (1969). *Auditory Discrimination in Depth*. USA: Riverside Publishers.
- Lindamood, C., Bell, N., & Lindamood, P. (1992). Issues in phonological awareness assessment. *Annals of Dyslexia, 42*, 242-259.
- Lindamood, P. C. (1994). Issues in researching the link between phonological awareness, learning disabilities, and spelling. In G. R. Lyon, (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 351-374). Baltimore, PH: Brookes Publishing Co.

- Lindsley, O. (1992). Why aren't effective teaching tools widely adopted. *Journal of Applied Behaviour Analysis, 25*, 21-26.
- Lloyd, J., Epstein, M., & Cullinan, D. (1981). Direct teaching for learning disabilities. In J. Gottlieb, & S. Strichart (Eds), *Perspectives on handicapping conditions: Current research and applications in learning disabilities*. University of California.
- Lockery, L., & Maggs, A. (1982). Direct instruction research in Australia: A ten year analysis. *Educational Psychology, 2*, 263-288.
- Lovett, M. W., Border, S. L., De Luca, T., Lacerenza, L., Benson, N. J., & Brackstone D. (1994). Treating the core deficit of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. *Developmental Psychology, 30*, 805-822.
- Lovett, M. W., & Steinbach, K. A. (1997). The effectiveness of remedial programs for reading disabled children of different ages: Does the benefit decrease for older children? *Learning Disability Quarterly, 20*, 189-209.
- Lukatela, K., Carello, C., Shankweiler, D., & Liberman, I. Y. (1995). Phonological awareness in illiterates: Observations from Serbo-Croatian. *Applied Psycholinguistics, 16*, 463-487
- Lundberg, I., Frost, J., & Petersen, O. P., (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263-284.
- Lyon, G. R. (1996). Learning disabilities. *The Future of Children: Special Education for Children with Disabilities, 6* (1), 54-76.
- Lyon, G. R., & Moats, L. C. (1997, November/December). Critical conceptual and methodological considerations in reading intervention research. *Journal of Learning Disabilities, 30*, 578-588.
- MacGinitie, W.H. (1991). Reading Instruction: Plus ça change. *Educational Leadership, 91*, 55-58.
- Maggs, A., & Murdoch, R. (1979). Teaching low performers in upper-primary and lower-secondary to read by direct instruction methods. *Reading Education, 4* (1), 35-39.
- Maggs, A., & White, R. (1982). The educational psychologist: Facing a new era. *Psychology in the Schools, 19*, 129-134.
- Majsterek, D. J., & Ellenwood, A. E. (1995). Phonological awareness and beginning reading: Evaluation of a school-based screening procedure. *Journal of Learning Disabilities,*

28, 449-456.

Mann, V. A. (1993). Phoneme awareness and future reading ability. *Journal of Learning Disabilities, 26*, 259-269.

Mann, V. A., & Brady, S. (1988). Reading disability: The role of language deficiencies. *Journal of Consulting and Clinical Psychology, 56*, 811-816.

Mann, V. A., Tobin, P., & Wilson R. (1987). Measuring phonological awareness through the invented spelling of kindergarten children. *Merrill-Palmer Quarterly, 33*, 365-391.

- Mather, N. (1992) Whole language reading instruction for students with learning disabilities: Caught in the cross fire. *Learning Disabilities Practice*, 7, 87-95.
- Mazzocco, M., Denckla, M., Singer, H., Scanlon, D., Vellutino, F., & Reiss, A. (1997). Neurogenic and neurodevelopmental pathways to learning disabilities. *Learning Disabilities: A Multidisciplinary Journal*, 8, 31-42.
- McCaslin, M. M. (1989). Whole language: Theory, instruction, and future implementation. *The Elementary School Journal*, 90, 223-229.
- McFaul, S. A. (1983, April). An examination of Direct Instruction. *Educational Leadership*, 67-69.
- McGee, L. M., & Lomax, R. G. (1990). On combining apples & oranges: A response to Stahl & Miller. *Review of Educational Research*, 60, 133-140.
- McGuinness, D., McGuinness, C., & Donohue, J. (1995). Phonological training and the alphabet principle: Evidence for reciprocal causality. *Reading Research Quarterly*, 30, 830-852.
- McLean, C., & Moore, J. (1985). Holy Family Education Centre Literacy Program. *NSW Journal of Special Education*, 3, 35-44.
- Medley, D. M. (1982). Teacher effectiveness. In H. E. Mitzel (Ed.), *Encyclopaedia of Educational Research* (5th ed., Vol. 4). (pp. 1894-1903). New York: The Free Press.
- Meyer, L. A., Gersten, R. M., & Gutkin, J. (1983). Direct Instruction: A Project Follow Through success story in an inner-city school. *The Elementary School Journal*, 84, 241-252.
- Moats, L. C. (1994a). Strategy assessment in perspective. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 607-614). Baltimore, PH: Brookes Publishing Co.
- Moats, L. C. (1994b). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia*, 44, 81-102.
- Moats, L. C. (1994c). Assessment of spelling in learning disabilities research. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 333-350). Baltimore, PH: Brookes Publishing Co.
- Mok, M., & Wheldall, K. (1995). Some reservations about the use of analysis of covariance in educational research. *Educational Psychology*, 15, 199-202.
- Morais, J. (1991). Phonological awareness: A bridge between language & literacy. In D.

- Sawyer, & B. Fox (Eds.), *Phonological awareness in reading: The evolution of current perspectives*. (pp. 31 - 71). New York: Springer-Verlag.
- Morais, J., Alegria, J., & Content, A. (1987). The relationship between segmental analysis and alphabetic literacy: An interactive view. *Cahiers de Psychologie Cognitive*, 7, 1-24.
- Morris, R. (1994). A review of critical concepts and issues in the measurement of learning disabilities. In G. Reid Lyon (Ed.), *Frames of reference for the assessment of learning disabilities*. Baltimore, MD: Brookes Publishing.
- Munro, J. (1995). Explaining developmental dyslexia: Orthographic processing difficulties. *Australian Journal of Remedial Education*, 27(1), 5-15.
- Murphy, G. (1980). The delivery of educational technical assistance services for children with special learning needs: A comment on Maggs. *The Australian Journal of Special Education*, 4 (1), 27-29.
- Nakano, Y., Kageyama, M., & Kinoshita, S. (1993). Using Direct Instruction to improve teacher performance, academic achievement, and classroom behaviour in a Japanese public junior high school. *Education and Treatment of Children*, 16, 326-343.
- Naslund, J. C., (1994). Understanding literacy and cognition: Trends in psycholinguistic and neuropsychological research. *German Journal of Educational Psychology*, 8 (3/4), 139-151.
- Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation predicts early reading and spelling skills. *Reading Research Quarterly*, 32, 154-167.
- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. WA, DC: US Government Printing Office.
- Newby, R. F., Recht, D. R., & Caldwell, J. (1993). Validation of a clinical method for the diagnosis of two subtypes of dyslexia. *Journal of Psychoeducational Assessment*, 11, 72-83.
- Newman, J. M. (1991). Whole language: A changed universe. *Contemporary Education*, 62, 70-75.
- Nicholls, C. (1980). Kentucky Fried reading scheme. *Reading Education*, 6, 18-22.
- Nicholson, T. (1985). Good readers don't guess. *Reading Psychology*, 6, 181-198.
- Nicholson, T. (1986). Research revisited: Reading is not a guessing game - the great debate revisited. *Reading Psychology*, 7, 197-210.
- Nicholson, T. (1991a). Do children read words better in context or in lists? A classic study revisited. *Journal of Educational Psychology*, 82, 444-450.

- Nicholson, T. (1991b). *Overcoming the Matthew effect: Solving reading problems across the curriculum*. Wellington: New Zealand Council for Educational Research.
- Nicholson, T., Bailey, J., & McArthur, J. (1991). Context cues in reading: The gap between research & popular opinion. *Reading, Writing & Learning Disabilities, 7*, 33-41.
- Nicholson, T., Lillas, C., Rzoska, M. A. (1988). Have we been misled by miscues? *The Reading Teacher, Oct.*, 6-10.
- Noon, L., & Maggs, A. (1980). Accelerating written language processes in normal and gifted children: Direct instruction strategies and sequences. *Reading Education, 5* (1), 11-26.
- O'Connor, R. E., Jenkins, J. R., Cole, K. N., & Mills, P. E. (1993). Two approaches to reading instruction with children with disabilities: Does program design make a difference? *Exceptional Children, 59*, 312-323.
- O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*, 532-546.
- O'Connor, R. E., Notary-Syverson, A., & Vadasy, P. F. (1996). Ladders to literacy: The effects of teacher-led phonological activities for kindergarten children with and without disabilities. *Exceptional Children, 63*, 117-130.
- O'Connor, R. E., Slocum, T. A., & Jenkins, J. R. (1995). Transfer among phonological tasks in kindergarten: Essential instructional content. *Journal of Educational Psychology, 87*, 202-217.
- Olson, R. K., & Wise, B. W. (1992). Reading on the computer with orthographic and speech feedback. *Reading and Writing: An Interdisciplinary Journal, 4*, 108-144.
- Olson, R., Forsberg, H., Wise, B., Rack, J. (1994): Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon, (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243-278). Baltimore: PH. Brookes Publishing Co.
- Olson, R., Wise, B., Conners, F., Rack, J. P., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities, 22*, 339-348.
- Parker, R. M. (1990). Power, control and validity in research. *Journal of Learning Disabilities, 23*, 613-620.

- Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S. J., & Frith, C. D. (1996). Is developmental dyslexia a disconnecting syndrome? Evidence from PET scanning. *Brain*, *119*, 143-157.
- Pearson, P. D. (1989). Commentary: Reading the whole language movement. *Elementary School Journal*, *90*, 231-241.
- Penney, R. K. (1988, March). Computability [*sic*] of early intervention programmes with normalization. *Bulletin of the Australian Psychological Society*, 28-29.
- Perfetti, C. A. (1991). Representations and awareness in the acquisition of reading competence. In L. Rieben, & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 33-44). NJ: Lawrence Erlbaum.
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition*. (pp. 145-174). NJ: Lawrence Erlbaum.
- Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, *33*, 283-319.
- Peterson, P. L. (1979, Oct.). Direct instruction: Effective for what and for whom? *Educational Leadership*, 46-48.
- Pollack, M., & Pickarz, J. (1963). *Reading programs & problem readers*. New York: David McKay.
- Polloway, E. A., & Epstein, M. H. (1986). The use of *Corrective Reading* (SRA) with mildly handicapped students. *ADI News*, *5*(2), 2-3.
- Polloway, E. A., Epstein, M. H., Polloway, C. H., Patton, J. R., & Ball, D. W. (1986). *Corrective Reading* program: An analysis of effectiveness with learning disabled and mentally retarded students. *RASE*, *7* (4), 41-47.
- Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology*, *80*, 319-323.
- Pressley, M. & Beard El-Dinary, P. (1997). What we know about translating comprehension-strategies instruction research into practice, *Journal of Learning Disabilities*, *30*, 486-488.
- Prior, M. (1993, June 15). Reading: The myths dispelled. *The Age*, p. 11.
- Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (1995). Reading disability in an Australian community sample. *Australian Journal of Psychology*, *47* (1), 32-37.

- Rack, J. P., Hulme, C., & Snowling, M. (1993). Learning to Read: A theoretical synthesis. *Advances in Child Development & Behaviour, 24*, 99-132.
- Rack, J. P., Snowling, M., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27*, 39-53.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. NJ: Prentice Hall.
- Read, C. (1991). Access to syllable structure in language and learning. In S. A. Brady, & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 119-124). NJ: Lawrence Erlbaum.
- Reetz, L. J., & Hoover, J. H. (1992). The acceptability and utility of five reading approaches as judged by middle school LD students. *Learning Disabilities, Research & Practice, 7*, 11-15.
- Richdale, A. L., Reece, J. E., & Lawson A. (1996). Teachers, children with reading difficulties, and remedial reading assistance in primary schools. *Behaviour Change, 13* (1), 47-61.
- Rist, M. C. (1992, November). Learning by heart. *The Executive Educator, 12-19*.
- Robinson, D. H., & Levin, J. R. (1997). Reflection on statistical and substantive , with a slice of replication. *Educational Researcher, 26*, 21-27.
- Rohl, M., & Pratt, C. (1995). Phonological awareness, verbal working memory and the acquisition of literacy. *Reading and Writing: An Interdisciplinary Journal: 7*, 327-360.
- Rosenshine, B. & Stevens, R. (1984). Classroom instruction in reading. In D. Pearson (Ed.), *Handbook of research on teaching* (3rd. ed.), pp. 376-391. New York: MacMillan.
- Rosenshine, B. V. (1979). Content, time and direct instruction. In P. L. Peterson, & H. J. Walbert (Eds.), *Research on teaching: Concepts, findings and implications*. (pp 28-57). Berkeley, CA: McCutchan.
- Rosenshine, B. V. (1980). *Direct instruction for skill mastery*. Paper presented to the School of Education, University of Milwaukee, Wisconsin.
- Rosenshine, B. V. (1986). Synthesis of research on explicit teaching. *Educational Leadership, 43* (7), 60-69.
- Rosenshine, B. V., & Berliner, D. C. (1978). Academic engaged time. *British Journal of Teacher Education, 4*, 3-16.

- Rosenshine, B. V., & Stevens, R. (1984). Classroom instruction in reading. In D. Pearson (Ed.), *Handbook of research on teaching* (3rd. ed.), pp. 376-391. New York: MacMillan.
- Roth, S. F., & Beck, I. L. (1987). Theoretical & instructional implications of the assessment of two microcomputer word recognition programs. *Reading Research Quarterly*, 22, 197-218.
- Rubin, H., Rotella, T., Schwartz, L., & Bernstein, S. (1991). The effect of phonological analysis training on naming performance. *Reading & Writing: An Interdisciplinary Journal*, 3, 1-10.
- Ruddell, R. B., & Sperlberg, M. (1988) Factors influencing the use of literacy research by the classroom teacher: Research review & new directions. In J. E. Readence (Ed.), *Dialogues in literacy research*. Chicago: National Reading Conference Inc.
- Salthouse, T. A. (1990). Working memory as a processing resource in cognitive ageing. *Developmental Review*, 10, 101-124.
- Sattler, J. M. (1992). *Assessment of children*: Revised and updated third edition. San Diego, California: Jerome M. Sattler, Publisher.
- Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, 21, 439-453.
- Schickedanz, J. A. (1990). The jury is still out on the effects of whole language and language experience approaches for beginning reading: A critique of Stahl & Miller's study. *Review of Educational Research*, 60, 127-131.
- Schneider, W., Kuspert, P., Roth, E., & Vise, M. (In press). Short and long term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology*.
- Schweinhart, L. J., Weikart, D. P., & Larner, W. B. (1986). Consequences of three pre-school curriculum models through age 15. *Early Childhood Research Quarterly*, 1 (1), 15-45.
- Scruggs, T., & Wong, B. (Eds.). (1990). *Intervention research in learning disabilities*. New York:Springer Verlag.
- Shankweiler, D., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and consequences. *Reading and Writing: An Interdisciplinary Journal*, 8, 267-294.
- Shannon, P. (1994). *People who live in glass houses*. In C. B. Smith (Ed.), *Whole Language: The debate*. (pp. 97-120). Bloomington, IN: ERIC.

- Shapiro, J., Nix, G. W., & Foster, S. F. (1990). Auditory perceptual processing in reading disabled children. *Journal of Research in Reading, 13*, 123-132.
- Share, D. L. (1990). Self correction rates in oral reading: Indices of efficient reading or artifact of text difficulty ? *Educational Psychology, 10* , 181-186 .
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218.
- Share DL, McGee R, & Silva P (1989). IQ and reading progress: A test of the capacity notion of IQ. *Journal of the American Academy of Child and Adolescent Psychiatry, 28*, 97-100.
- Share, D. L., & Stanovich, K. E. (1995). Cognitive processes in early reading development: accommodating individual differences into a model of acquisition. *Issues in Education, 1*(1), 1-57.
- Shaywitz, B., Shaywitz, S., Francis, D., Fletcher, J., Pugh, K., Gore, J., Todd Constable, J., Fulbright, R., Skudlarski, P., Liberman, A., Shankweiler, D., Katz, L., Bronen, R., Marchione, K., Holahan, J., Francis, D., Klorman, R., Aram, D., Blachman, B., Stuebing, K., & Lacadie, C. (1997). The Yale Center for the study of learning and attention: Longitudinal and neurobiological studies. *Learning Disabilities: A Multidisciplinary Journal, 8*, 21-29.
- Siegel, L. S. (1993). The development of reading. *Advances in Child Development and Behavior, 24*, 63-97.
- Simmons, D. C. (1992). Perspectives on dyslexia: Commentary on educational concerns. *Journal of Learning Disabilities, 25*, 66-70.
- Simmons, D. C., Gunn, B., Smith, S. B., & Kameenui, E. J. (1995). Phonological awareness: Application of instructional design. *LD Forum, 19*(2), 7-10.
- Simner, M. L. (1995). Reply to the Ministries' reactions to the Canadian Psychological Association's position paper on beginning reading instruction. *Canadian Psychology, 36*, 333-342.
- Singer, H., & Balow, I. H. (1981). Overcoming educational disadvantages. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research review*. (pp. 101-114). Newark, DE: International Reading Association.
- Singh, N. N., & Solman, R. T. (1990). A stimulus control analysis of the picture-word problem: The blocking effect. *Journal of Applied Behaviour Analysis, 23*, 525-532.

- Singh, N. N., Deitz, D. E. D., & Singh, J. (1992). Behavioural approaches. In Nirbay N. Singh, & Ivan L. Beale (Eds.) *Learning disabilities: Nature, theory, and treatment*. New York: Springer-Verlag.
- Slavin, R. E. (1990). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.
- Smith, F. (1975). *Comprehension and learning: A conceptual framework for teachers*. New York: Richard C. Owen.
- Smith, F. (1978). *Understanding reading*. New York: Holt, Rinehart, & Winston.
- Smith, F. (1992). Learning to read: The never-ending debate. *Phi Delta Kappan*, 74, 432-441.
- Smith, P.G. (1991). A practical guide to whole language in the intermediate classroom. *Contemporary Education*, 62, 88-95.
- Smith, S. B, Simmons, D. C., & Kameenui, E. J. (1995). *Synthesis of research on phonological awareness: Principles and implications for reading acquisition*. NCITE Technical Report No 21. (On-Line). Available: <http://darkwing.uoregon.edu/~ncite/reading/PhonoSyn.html>.
- Snider, V. E. (1992). Learning styles and learning to read: A critique. *Remedial and Special Education*, 13 (1), 6-18.
- Snowling, M. J. (1996). Annotation: Contemporary approaches to the teaching of reading. *Journal of Child Psychology & Psychiatry*, 37, 139-148.
- Snowling, M. J., & Hulme, C. (1991). Speech processing and learning to spell, In R. Ellis, & R. Bowler (Eds.), *Language and the creation of literacy* (pp. 33-39). Baltimore: Orton Dyslexia Society.
- Snowling, M. J., Goulandris, N., & Defty, N. (1996). A longitudinal study of reading development in dyslexic children. *Journal of Educational Psychology*, 88, 653-669.
- Solity, J. (1991). An overview of behavioural approaches to teaching children with learning difficulties and the national curriculum. *Educational Psychology*, 11, 151-166.
- Solman, R. & Stanovich, K. E. (1992). Information processing models. In N. Singh, & I. Beale (Eds.), *Learning disabilities: Nature, theory & Treatment*. New York: Springer Verlag.
- Solman, R. T. (1986). Teaching sight words. *Set: Research information for teachers*, (2). Melbourne: ACER.
- Solman, R. T., Singh, N. N., & Kehoe, E. J. (1992). Pictures block the learning of sightwords. *Educational Psychology*, 12, 143-154.

- Solomons, B. (1992). *Phonemic awareness training: An early intervention program*. NSW, Australia: Macquarie University.
- Sommers, J. (1995). Seven-year overview of Direct Instruction programs used in basic skills classes at Big Piney Middle School. *Effective School Practices, Fall*, 29-32.
- Spear-Swerling, L., & Sternberg, R. J. (1994). The road not taken: An integrative theoretical model of reading disability. *Journal of Learning Disabilities, 27*, 91-103.
- Spector, J. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology, 84*, 353-363.
- Spector, J. (1995). Phonemic awareness training: Application of principles of direct instruction. *Reading and Writing Quarterly, 11*, 37-51.
- Spedding, S., & Chan, L. K. S. (1993). Metacognition, word identification, and reading competence. *Contemporary Educational Psychology, 18*, 91-100.
- Spiegel, D. L. (1992). Blending whole language and systematic direct instruction. *The Reading Teacher, 46* (1), 38-44.
- Stage, S. A., & Wagner, R. K. (1992). Developmental of young children' phonological and orthographic knowledge as revealed by their spellings. *Developmental Psychology, 28*, 287-296.
- Stahl, S. A. (1990). Riding the pendulum: A rejoinder to Schickedanz, and McGee and Lomax. *Review of Educational Research, 60*, 141-151.
- Stahl, S. A. (1992). Saying the "p" word: Nine guidelines for exemplary phonics instruction. *The Reading Teacher, 45*, 618-625.
- Stahl, S. A., & Kuhn, M. R. (1995). Does Whole Language or instruction matched to learning styles help children learn to read? *School Psychology Review, 24*, 393-404.
- Stahl, S. A., & Miller, P. D. (1989). Whole language & language experience approaches for beginning reading: A quantitative research synthesis. *Review of Educational Research, 59*, 87-116
- Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.
- Stanovich, K. E. (1985). Explaining the variance in reading ability in terms of psychological processes: What have we learned? *Annals of Dyslexia, 35*, 67-96.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.

- Stanovich, K. E. (1988a). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21*, 590-604
- Stanovich, K. E. (1988b). The right and wrong places to look for the cognitive locus of reading disability. *Annals of Dyslexia, 38*, 154-157.
- Stanovich, K. E. (1992). Speculation on the causes and consequences of individual differences in early reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 307-342). Hillsdale, NJ: Laurence Erlbaum.
- Stanovich, K. E. (1993a). Does reading make you smarter? Literacy and the development of verbal intelligence. *Advances in Child Development & Behaviour, 24*, 133-180.
- Stanovich, K. E. (1993b). The language code: Issues in word recognition. In S. R. Yussen, & M. C. Smith (Eds.), *Reading across the life span* (pp. 111-135). New York: Springer-Verlag.
- Stanovich, K. E. (1994). Romance and reality. *The Reading Teacher, 47*, 280-291.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86* (1), 24-53.
- Statistical Package for the Social Sciences [Computer Software]. (1995). Chicago IL: SPSS. (Version 6.1 for the Apple Macintosh).
- Stebbins, L., St. Pierre, R. G., Proper, E. C., Anderson, R. B., and Cerva, T. R. (1977). *Education as experimentation: A planned variation model (Vol. IV.)*. Cambridge, MA: Abt. Associates.
- Stedman, L., & Kaestle, C. (1987). Literacy & reading performance in the United States, from 1880 to the present. *Reading Research Quarterly, 23* (1), 8-46.
- Stevens, R. J., Slavin, R. E., & Farnish, A. M. (1991). The effects of co-operative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*, 8-16.
- Stodhard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children, *Reading and Writing, 4*, 245-256.
- Stone, J. E. (April 23, 1996). Developmentalism: An obscure but pervasive restriction on educational improvement. *Education Policy Analysis Archives* [On-line] 4, 1950 lines. Available: <http://seamonkey.ed.asu.edu/epaa>.
- Stuart, M. (1995). Prediction and qualitative assessment of five and six-year-old children's

- reading: A longitudinal study. *British Journal of Educational Psychology*, 65, 287-296.
- Stuart, M., (1995). Prediction and qualitative assessment of five and six-year-old children's reading: A longitudinal study. *British Journal of Educational Psychology*, 65, 287-296.
- Swanson, H. L., & Alexander, J. E. (1997). Cognitive processes as predictors of word recognition and reading comprehension in learning disabled and skilled readers: Revisiting the specificity hypothesis. *Journal of Educational Psychology*, 89, 128-158.
- Sykes, S. (1991). A whole language perspective on reading and writing. *Australian Journal of Remedial Education*, 23(2), 23-27.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- Tangel, D. M., & Blachman, B. A. (1992). Effect of phoneme awareness instruction on kindergarten children's invented spelling. *Journal of Reading Behaviour*, 24, 233-261.
- Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, 25 (2), 26-30.
- Thorne, M. T. (1978). "Payment for Reading". The use of the "Corrective Reading Scheme" with junior maladjusted boys. *Remedial Education*, 13, 87-89.
- Torgesen, J. K. (1993). Variations on theory in learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), *Better understanding of learning disabilities: New views from research and their implications for education and public policies* (pp. 153-170). Baltimore: Brooks Publishing.
- Torgesen, J. K., & Bryant, B. (1994a). *Test of Phonological Awareness*. Austin, TX: Pro-Ed.
- Torgesen, J. K., & Bryant, B. (1994b). *Phonological Awareness Training for Reading*. Austin, TX: Pro-Ed.
- Torgesen, J. K., Morgan, S. T., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology*, 84, 364-370.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1994). Longitudinal studies of phonological processing & reading. *Journal of Learning Disabilities*, 27, 276-286.

- Torgesen, J. K., Wagner, R. K., Simmons, K., & Laughon, P. (1990). Identifying phonological coding problems in disabled readers: Naming, counting, or span measures? *Learning Disability Quarterly*, *13*, 236-243.
- Torgesen, J., Wagner, R. K., Rashotte, C., Alexander, A., & Conway, T. (1997). Preventative and remedial interventions for children with severe reading disabilities. *Learning Disabilities: A Multidisciplinary Journal*, *8*, 51-61.
- Treiman, R. (1985). Onsets & rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology*, *39*, 161-181.
- Treiman, R. (1991). The role of intrasyllabic units in learning to read and spell. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 65-106). Hillsdale, NJ: Erlbaum.
- Treiman, R. (1993). *Beginning to spell*. New York: Oxford University Press.
- Trelease, J. (1989). *The New Read Aloud Handbook*. New York: Penguin.
- Troia, G.A., Roth, F.P., & Yeni-Komshian, G.H. (1996). Word frequency and age effects in normally developing children's phonological processing. *Journal of Speech and Hearing Research*, *39*, 1099-1108.
- Tuckman, B. W. (1991). *Educational psychology: From theory to application*. USA: Harcourt Brace Jovanovich.
- Tunmer, W. E., & Hoover, W. A. (1993). Phonological recoding skill and beginning reading. *Reading & Writing: An Interdisciplinary Journal*, *5*, 161-179.
- Uhry, J. K., & Shepherd, M. J., (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly*, *28*, 219-233.
- Vandervelden, M. C., & Siegel, L. S. (1995). Phonological recoding and phoneme awareness in early literacy: A developmental approach. *Reading Research Quarterly*, *30*, 854-875.
- Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology*, *83*, 437-443.
- Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness and reading ability: Evidence from a longitudinal & experimental study. *Merrill-Palmer Quarterly*, *33*, 321-363.

- Vellutino, F. R., Scanlon, D. M., & Spearing, D. (1995). Semantic and phonological coding in poor and normal readers. *Journal of Experimental Child Psychology*, 59, 76-123.
- Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). Component of reading ability: issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues*. Baltimore, PH. Brookes Publishing Co., pp. 279-332.

- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*, 601-638.
- Wagner, R. K. (1988). Causal relations between the development of phonological processing abilities and the acquisition of reading skills: A meta-analysis. *Merrill-Palmer Quarterly, 34*, 261-279.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30* (1), 73-87.
- Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85*, 83-103.
- Walsh, D. J., Price, G. G., & Gillingham, M. G. (1988). The critical but transitory importance of letter naming. *Reading Research Quarterly, 23*, 108-121.
- Watson, D. (1989). Defining and describing whole language. *Elementary School Journal, 90*, 129-142.
- Weaver, C. (1988). *Reading process & practice: From socio-psycholinguistics to whole language*. Portsmouth, NH: Heinemann.
- Weaver, C. (1991). Weighing the claims about "Phonics First". *The Education Digest, April*, 19-22.
- Webster, P. E., & Plante, A. S. (1992). Effects of phonological impairment on word, syllable and phoneme segmentation and reading. *Language, Speech and Hearing Services in Schools, 23*, 176-182.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.) San Antonio: Psychological Corporation.
- Weir, R. (1990). Philosophy, cultural beliefs and literacy. *Interchange, 21* (4), 24-33.
- Weisberg, P., & Savard, C. F. (1993). Teaching preschoolers to read: Don't stop between the sounds when segmenting words. *Education & Treatment of Children, 16* (1), 1-18.

- White, W. A. T. (1988). A meta-analysis of the effects of direct instruction in special education. *Education & Treatment of Children, 11*, 364-374.
- Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology, 72*, 1-15.
- Williams, J. P. (1991a). Beginning to read: A critique by literacy professionals and a response by Marilyn Jager Adams, *The Reading Teacher, 44*, 370-395.
- Williams, J. P. (1991b). The meaning of a phonics base for reading instruction. In The Orton Dyslexia Society (Eds.), *All language and the creation of literacy* (pp. 57-62).
- Winnie, P. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. *Review of Educational Research, 49*, 13-50.
- Wise, B., Olson, R., & Treiman, R. (1990). Subsyllabic units as aids in beginning readers word learning: Onset-rime versus post-vowel segmentation. *Journal of Experimental Child Psychology, 49*, 1-19.
- Wolery, M., Ault, M. J., & Doyle, P. M. (1992). *Teaching students with moderate to severe disabilities*. New York: Longman.
- Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly, 26*, 123-141.
- Wood, F. B., & Felton, R. H. (1994). Separate linguistic and attentional factors in the development of reading. *Topics in Language Disorders, 14* (4), 42-57.
- Woodcock, R. W. (1973). *Woodcock Reading Mastery Tests*. MN: American Guidance Service.
- Woodcock, R. W. (1987). *Woodcock Tests of Reading Mastery -Revised*. MN: American Guidance Service.
- Woodward, J. (1993). The technology of technology-based instruction: Comments on the research, development, and dissemination approach to innovation. *Education & Treatment of Children, 16*, 345-360.
- Yates, G. C. R. (1988). Classroom research into effective teaching. *Australian Journal of Remedial Education, 20* (1), 4-9.
- Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23*, 159-177.
- Yopp, H. K. (1992). Developing phonemic awareness in young children. *The Reading Teacher, 45*, 696-703.

APPENDICES

APPENDIX A

The Picture Naming Test.

(Hempenstall, 1995a)

This test is designed to see how many pictures children can name in one minute. It is a measure of a child's degree of ready accessibility to names (lexical access). It is relevant to reading because it is indicative of how readily children can gain access to a sound, sound sequence, or word meaning.

Instructions to examiners.

1. Read the following script to students.

I'm going to show you some pictures, and I want you to tell me what's in them. I want to see how many pictures you can name in one minute, so go as fast as you can. If you don't know an answer, go onto the next picture. There are more pictures on the next page, so turn over when you finish a page. First we'll have a practice. What do these pictures show?

2. Present the practice page. Prompt moving on if the child spends more than 2 seconds on any one picture. Do not provide answers. Ask the child to read left to right, top to bottom; if necessary, demonstrate without naming the pictures.

3. Present the test pages. Say:

Let's see how many you can name in one minute. Go!

4. Accept reasonable responses, for example, on Page 1, Picture 13, responses which describe the person (*man sick in bed*), or the process or concept (*sick*), are acceptable. Do not accept responses which merely define the name, for example, *something you eat with* is not a correct response to the picture of a fork.

5. Mark errors and omissions separately on score sheet.

Acknowledgement: Pictures derived from TOPA, Test Of Phonological Awareness (Torgesen & Bryant, 1994), PRO-ED.

Acceptable Solutions

PRACTICE PAGE:

1. bat; baseball bat
2. trumpet; horn; cornet
3. bed; double bed
4. cup
5. car; speeding
6. cake; birthday cake
7. hook; fish hook
8. box; case
9. frog
10. pig
11. cow
12. gun; sixgun; revolver

TEST PAGE 1

1. leg
2. lamp; light
3. hand
4. fish
5. fire
6. hat
7. star
8. foot
9. pliers
10. drum
11. tie
12. cup
13. sick
14. pail; bucket
15. two
16. sewing; sew; stitching
17. cake
18. key
19. fall; doll; girl
20. bell; ring

PAGE 2

1. girl
2. chicken; bird
3. goat
4. cat
5. smile; nose & mouth
6. cup
7. bed
8. dig; gardening; boy digging
9. duck; bird
10. arm
11. dog
12. tyre; wheel
13. nest
14. leaf
15. nine
16. mouse
17. conch; shell
18. shine; polish; shoeshine; shoepolish
19. hut; house; home
20. face; smile

PAGE 3

1. fork
2. fan
3. foot
4. shirt
5. heart; loveheart
6. horn; bugle
7. gaol; prison; prisoner; criminal
8. house; home
9. dog
10. top; spin
11. table; desk
12. bat; baseball bat
13. night; moon; cloudy
14. nose
15. nest
16. pin; nail
17. cry; fear
18. steal
19. pot; pan; saucepan
20. pat; pat dog

APPENDIX B

Corrective Reading Program Evaluation: Parents

Your child has been participating in a special reading assistance program at the school, and we would like to find out how useful it has been for your child. We are particularly interested to learn whether you have noticed any changes in your child's reading. We would appreciate your help in filling out this form, and returning it to us as soon as is convenient.

Please underline the words which **best describe** your child's current reading.

In terms of the amount of reading done at home, my child is now reading **much more than a little more than the same as less than** before the program's introduction.

If you have noticed an increase, what type(s) of reading materials does your child favour?

In terms of the skill of reading done at home, my child is now reading **much better than better than the same as worse than** before the program's introduction.

If you have noticed a skill improvement, is it in **speed, accuracy, smoothness, preparedness to read out loud understanding of what is read?**
(You may underline any number of these words.)

In terms of the enjoyment of reading done at home, my child now seems to find reading **much more enjoyable than more enjoyable than the same as less enjoyable than** before the program's introduction.

Do you have any other comments which you think might be helpful to future planning? Please write them below.

Corrective Reading Program Evaluation: Teachers

One or more of your students has been participating in a special reading assistance program at the school, and we would like to find out how useful it has been for him/her. We are particularly interested to learn whether you have noticed any changes in your student's reading, and general performance.

Please underline the words which **best describe** your student's current reading.

In terms of the *amount* of reading done at school, my student is now reading **much more than a little more than the same as less than** before the program's introduction.

If you have noticed an increase, what type(s) of reading materials does your student favour?

In terms of the *skill* of reading done at school, my student is now reading **much better than better than the same as worse than** before the program's introduction.

If you have noticed a *skill improvement*, is it in **speed accuracy smoothness preparedness to read out loud understanding of what is read?** (You may underline any number of these words).

In terms of the *enjoyment* of reading done at school, my student now seems to find reading **much more enjoyable than more enjoyable than the same as less enjoyable than** before the program's introduction.

Is there evidence of change in reading skills in *other* curriculum areas ie., have the skills transferred? The student is **much better than better than the same as worse than** before the program's introduction.

Has there been any change in the student's *attitude, or behaviour* generally? The student is **much better than better than the same as worse than** before the program's introduction.

Do you have any other comments which you think might be helpful to future planning? Please write them below.



APPENDIX C

DISTAR Library Series

Stories with corresponding lesson from “*Teach Your Child to Read in 100 Easy Lessons*”. Allows extra reading practice throughout the program.

<i>Book</i>	<i>Page Numbers</i>	<i>Lesson Numbers</i>
Book A	1 + 2	35
	3 + 4	36
	5 + 6	37
	7 + 8	38
	10 + 11	39
	12, 13, 14	41
Book B	1 - 5	42
	6 - 10	43
	11 - 15	44
	16 - 18	45
	19 - 23	46
Book C	1 - 5	47
	6 - 9	48
	10 - 13	49
	14 - 17	50
	18 - 19	51
	21 - 23	52
Book D	1 - 4	53
	5 - 8	54
	9 - 13	55
	14 - 17	56
	18 - 20	57
	21 - 23	58
Book E	1 - 4	59
	5 - 8	60
	9 - 12	61
	13 - 15	62
	16 - 18	63
	19 - 21	64
	22 - 24	65
	25 - 31	66
Book F	1 - 4	67
	5 - 7	68
	8 - 11	69

12 - 13	70
14 - 15	71
16 - 18	72
19 - 21	73
22 - 23	74

The lesson numbers are not exact matches; however, all books should be completed by Lesson 74 as the DISTAR orthography is discontinued at that lesson