

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**CAUSAL METACOGNITIVE-MOTIVATIONAL MODELS OF
READING COMPREHENSION IN READING DISABLED
AND NORMAL ACHIEVING READERS**

27
6389

**A thesis presented in partial
fulfilment of the requirements for
the degree of**

DOCTOR OF PHILOSOPHY IN PSYCHOLOGY

JOYCE ANNE PEREIRA-LAIRD

Massey University

1996

371.264

Per

02 20

ABSTRACT

Contemporary models of reading indicate that reading achievement and impairment are the products of the complex interaction of motivational, cognitive and metacognitive processes. Most previous research has relied on correlational studies to examine the links amongst these variables. Given the complex relationships of these variables, research designs which examine these constructs simultaneously and which establish causal relationships are needed. The dearth of interactive research with different populations is surprising considering that reliable and theoretically meaningful models that are generally invariant across subpopulations would contribute much towards theoretical parsimony and progress of educational research.

In light of the above considerations, the present study was designed with the primary goal of replicating and extending a previous test of a structural model of reading achievement. The main goal was to explain and predict both reading achievement and impairment from the complex and multicomponential perspective of a model of metacognition. Specifically, this involved an examination of the causal influences of young adolescent students' attributional style, and self-efficacy on metacognitive knowledge and their use of cognitive and metacognitive strategies. In addition, these same variables were examined to see how they related to their reading comprehension performance and how the relationships differed in normal achieving (NA) and reading disabled children (RD). A secondary goal of this study was the investigation of variables that would distinguish between RD and NA readers.

There were three phases involved in the present study. Phase 1 concerned sample selection and involved administration of a short-form of the Wechsler Intelligence Scale for Children - Revised. Selection of RD children was based on a six-stage multidefinitonal approach. A sample of NA readers with reading achievement consistent with their age was also identified. A total of 203 NA readers and 204 RD readers were selected to participate in this study.

The data were collected in Phases 2 and 3. Phase 2 involved administration of two self-report questionnaires which examined children's attributional style, use of strategies, metacognitive knowledge, and self-efficacy for reading. Phase 3 involved individually administered reading interviews. All questionnaires and reading interviews were administered within a two week period.

The relationships among general intellectual ability, attributions, self-efficacy, metacognitive knowledge, metacognitive/cognitive strategy use, and reading comprehension in NA and RD children were evaluated using four models. The results were analyzed using structural equation modelling procedures. The proposed models provided a statistically adequate fit for the obtained data, accounting for about 60% of the variance in student performance. Several structural relationships were similar across groups suggesting that the metacognitive-motivational systems of NA and RD children were rather similar. For instance, the relationships between attributional style (as a single latent construct), efficacy, metacognitive knowledge, cognitive strategy use and metacognitive strategy use were similar across groups.

Nearly all of the structural correlations and the direct and indirect coefficients were in the theoretically expected direction. In both groups, students' adaptive attributional beliefs significantly predicted self-efficacy and metacognitive knowledge. However, when the separate effects of attributional style were examined for each outcome, the results revealed that adaptive attributional style for failure was the only significant predictor of metacognitive knowledge. Furthermore, the attributional components varied in their impact on self-efficacy and these differential effects also varied across groups.

An important contribution of this study was the incorporation of "strategy use" in the model. When combined strategy use (both metacognitive and cognitive) was included in the model, metacognitive knowledge no longer had a direct impact on reading performance (comprehension), instead combined strategy use played a significant role in mediating this relationship. Self-efficacy as well as metacognitive knowledge predicted combined strategy use which in turn predicted reading comprehension. Closer examination of the components of combined strategy use revealed that only "metacognitive strategy use" directly predicted reading comprehension across groups. The mediating role played by cognitive strategy use in the relationship between metacognitive knowledge and comprehension performance differed across groups. Self-efficacy directly and positively predicted metacognitive knowledge and metacognitive/cognitive strategy use.

The results suggested that attributional style plays a pivotal role in the metacognitive development of both NA and RD children. A number of causal paths distinguished good from poor readers. They were the paths between ability and success/failure attributions, ability and performance, success/failure attributions and efficacy, cognitive strategy use and performance, and efficacy and performance. On the whole, motivational variables were more important in determining comprehension for RD children whilst metacognitive and cognitive strategy use variables were more important for achieving readers. The failure to develop an enriched metacognitive system in RD children was ascribed partially to the effects of their self-defeating attributions.

Attributional beliefs, self-efficacy, metacognitive knowledge, and cognitive strategy use uniquely discriminated between NA and RD children. These findings suggest that metacognitive and motivational variables combine effectively to distinguish between RD and NA readers. The results also provide support for the utility of adopting a multidefinitonal approach in defining RD children.

The findings from this study advance the argument that reading achievement and impairment should be studied using a multicomponential framework. The implications of this study's research findings for classroom practice and research methodology are reviewed. Limitations of the present study were also discussed.

DEDICATION

This dissertation is dedicated to the memory of

my Mother, *Thelma A. Pereira*

my Father, *Anthony P. Pereira*

and my Big Papa, *Elias P. Pereira*

ACKNOWLEDGEMENTS

The completion of this dissertation was made possible through the encouragement and willing involvement of many people. First and foremost, I would like to thank my cherished husband, Dr. Richard Laird, for his love, patience, personal support, encouragement, and understanding. Not only did Richie proof read my final draft, but willingly accepted a number of other responsibilities which were instrumental in my being able to complete this dissertation.

I am indebted to my chief supervisor, Dr. Frank Deane, for his sympathetic understanding, advice, and ongoing encouragement throughout my research. His unique sense of humour helped alleviate some of the stress involved in completing this dissertation. I am also grateful to my second supervisor, Dr. Julie Bunnell, for her guidance, advice, support, encouragement, and her willingness to lend me a sympathetic ear when I needed it most. Overall, I wish to thank my supervisors for their positive attitude, constructive criticism, and excellent feedback on the drafts of my dissertation.

I am particularly indebted to Mr. John Horwood of the Christchurch Health and Development Research Unit for his expertise and invaluable advice throughout the data analysis stages of my research. His patience and expertise in teaching me LISREL was gratefully appreciated. His positive attitude and readiness to help, coupled with his constructive criticism, was a major factor contributing to the completion of this research. Thanks also to Monica Skinner for her patience and time in helping me get started on LISREL.

Thanks also goes to Dr. Stephen von Tetzchner, Dr. John Spicer, Associate Professor James Chapman, and Mrs Shannon Roache who assisted me at various stages of my dissertation. Special thanks to Associate Professor James Chapman and Dr. Gus Habermann for providing helpful comments on my final draft.

In addition, thanks to the Professor Nigel Long (HOD), Michael Donnelly, Tony True, Harvey Jones, Carol Vincent, the Secretaries, and Technicians of the Psychology Department for their patience and endurance in the seemingly endless tasks of printing, photocopying, giving advice on aspects of computing, and obtaining equipment and funding for my research. I am grateful to others, including some staff in the Psychology, Statistics, and Computing Services

Departments, as well as friends, for their advice or interest in the progress of my research. I also wish to acknowledge the help of some postgraduate psychology students who helped me conduct the interviews for my research. I am also deeply indebted to the school principals who allowed me access to the students who participated in the study.

I would like to thank Dr. Alan Winton and his family for their kindness, thoughtfulness, support, and encouragement throughout my dissertation. I would also like to specially thank my immediate and extended family, particularly my two sweet and precious brothers (Terence and Melwin) and their families, for their unswerving loyalty, their generous love, kind support, and encouragement in everything I have undertaken. Thanks also to my parents and my Big papa, who moulded me into the person I am now, and who have given me the inspiration and courage to come this far in education.

Finally, and most significantly, thanks to GOD, who I owe everything in my life. As I journeyed through this particular path, I discovered that God was always there by my side, and that he was at the heart of every person on every road I took towards completing my dissertation.

TABLE OF CONTENTS

	Page
Abstract	iii
Dedication	vi
Acknowledgements	vii
Table of Contents	ix
List of Tables	xv
List of Figures	xvi
CHAPTER 1 INTRODUCTION AND OVERVIEW	1
1.1 Identifying some concerns in the achievement literature	1
1.2 Overview of the research to be undertaken	4
1.3 Outline of chapters	5
CHAPTER 2 DEFINITIONS OF LEARNING AND READING DISABILITY	7
2.1 A definition that receives current consensus	9
2.2 Prevalence of learning disabilities	10
2.3 The concept of reading disability	11
2.3.1 Differences between reading disabled (RD) and poor readers	11
2.3.2 Variability of terms to describe reading disability	12
2.4 Common elements or assumptions across definitions	12
2.4.1 Intrinsic problem	13
2.4.2 Ability-achievement discrepancy	13
2.4.3 Exclusionary criteria	14
2.4.4 Uneven growth pattern	15
2.4.5 Specificity of academic problem	15
2.4.6 Low academic achievement	16
2.5 Discrepancy approach to defining reading disability	16
2.5.1 Limitations of the discrepancy-based approach	17
2.5.1.1 Limitations of using IQ in the assessment of discrepancy	17
2.5.1.2 Construct validity of the discrepancy approach	19
(i) Replicability of nonnormal distribution	19
(ii) Validity of differentiating learning disabilities based on IQ discrepancies	20
2.6 Multifaceted approaches to defining reading disability	21
2.7 Conclusion	23
CHAPTER 3 METACOGNITION AND COGNITION AND THEIR RELATIONSHIP TO READING	24
3.1 Definition of metacognition	25
3.2 Theoretical issues concerning metacognition	26
3.2.1 The concept of strategy	27
3.2.2 Distinction between components of metacognition	27
3.2.3 Distinction between cognition and metacognition	28
3.2.3.1 Activities involved in cognitive and metacognitive strategies	29
(i) Reading for meaning and reading for remembering	30
3.2.4 Conceptual difficulties in linking metacognition to reading disabilities	31

3.3	Methodological issues concerning cognition and metacognition	32
3.3.1	Verbal report	32
3.3.2	Self-report questionnaires	33
3.3.3	Error detection	34
3.3.4	Conclusion	35
3.4	Empirical research regarding the relationships between cognition, metacognition, and performance	35
3.4.1	Relationship between metacognition and comprehension	36
3.4.2	Relationship between metacognitive knowledge and performance	36
3.4.3	Relationship between cognitive/metacognitive strategy use and performance	37
3.4.4	Relationship between metacognitive knowledge and cognitive/metacognitive strategy use	40
3.4.4.1	Inactive versus inefficient view of cognitive strategy deficits	43
3.5	Summary and conclusion	44
CHAPTER 4 MOTIVATIONAL FACTORS IN ACADEMIC PERFORMANCE		46
4.1	Self-efficacy	46
4.2	Causal Attributions	48
4.3	Relationship between self-efficacy and attributions	52
4.4	Methodological problems in measuring attributions and self-efficacy	52
4.4.1	Attributions	52
4.4.2	Self-efficacy	55
4.5	Summary	55
CHAPTER 5 THE RELATIONSHIPS BETWEEN METACOGNITIVE, COGNITIVE, AND MOTIVATIONAL VARIABLES		57
5.1	Triple Alliance Model	57
5.2	Borkowski's model of metacognition	58
5.2.1	Components in the metacognition model	59
5.2.2	Development of metacognitive components in a child	61
5.2.3	Relationships between components of the model	62
5.2.4	Implications of the metacognition model for RD children	64
5.2.5	Empirical research supporting aspects of Borkowski's model	66
5.3	The need for causal models	70
5.3.1	Causal models of reading achievement	71
5.3.1.1	Limitations of Carr et al.'s (1991) model	73
	(i) Methodological limitations	73
	(ii) Hypothesized relationships	75
	(iii) Statistical analyses	75
5.3.1.2	Summary	76
CHAPTER 6 RATIONALE FOR PRESENT STUDY		77
6.1	Primary goals of study	77
6.1.1	Testable interactive or multicomponential causal models	77
6.1.2	Issues in need of investigation	78
6.1.2.1	Replication of Carr et al.'s (1991) model	78
6.1.2.2	The mediating role of metacognitive knowledge and strategy use on the relationship between motivational beliefs and performance	79
6.1.2.3	The relative impact of self-efficacy and attributional style on metacognitive knowledge	81

6.1.2.4	The effect of self-efficacy on strategy use and performance	81
6.1.2.5	The influence and relative impact of cognitive and metacognitive strategy use on performance	82
6.1.2.6	The relative impact of success and failure attributions on efficacy and metacognitive knowledge	83
6.2	Hypothesized Models	84
6.2.1	Proposed General Model 1	85
6.2.1.1	Hypotheses	88
6.2.2	Proposed General Model 2	88
6.2.2.1	Hypotheses	90
6.2.3	Proposed General Model 3	90
6.2.3.1	Hypotheses	92
6.2.4	Proposed General Model 4	93
6.2.4.1	Hypotheses:	95
6.3	Secondary goals of study	95
6.3.1	Differences in motivational, cognitive, and metacognitive variables across groups	95
6.3.1.1	Hypotheses:	95
6.4	Summary	96
CHAPTER 7 STRUCTURAL EQUATION MODELLING		97
7.1	An introduction to structural equation modelling and the LISREL model	98
7.1.1	Model specification	100
7.1.1.1	Measurement model	101
7.1.1.2	Structural model	102
7.1.2	Other parameter matrices of LISREL	103
7.1.3	Types of model parameters	104
7.1.4	Model identification	107
7.1.5	Parameter estimation	108
7.1.6	Model evaluation	109
7.1.6.1	Goodness-of-fit indices	110
	(i) Measures of overall fit	110
	(ii) Detailed measures of fit	113
7.1.6.2	Comparison of nested models	115
7.2	Multisample model	117
7.3	Justification for LISREL-type models	118
7.4	LISREL formulation of reading performance models	120
7.4.1	Specification of Model 1 for the single sample situation	120
7.4.1.1	Measurement model	123
7.4.1.2	Structural model	125
7.4.1.3	Disturbance assumptions	125
7.4.1.4	Model assumptions	126
7.4.1.5	Identification of model	128
7.4.2	Method effects in self-report data	128
7.4.2.1	Respecification of Model 1 to include correlated errors	130
7.4.3	The multisample situation	132
7.4.3.1	Respecification of Model 1 to include the multisample situation	133
	(i) Measurement model	134
	(ii) Structural model	135
	(iii) Disturbance assumptions	136
	(iv) Identification of model	137

7.5	Advantages and limitations of using SEM techniques	138
7.6	Conclusion	140
CHAPTER 8	METHOD	141
8.1	Overview	141
8.2	Sample	142
8.2.1	Selection of reading disabled (RD) sample	142
8.2.2	Selection of normal achieving (NA) sample	145
8.2.3	Demographic information	146
8.3	Measures	148
8.3.1	Pilot tests	150
8.3.2	Ability	150
8.3.3	Achievement	151
8.3.3.1	Academic achievement	151
8.3.3.2	Reading comprehension performance	151
8.3.4	Causal attribution	152
8.3.5	Self-efficacy	153
8.3.6	Metacognitive knowledge for reading	154
8.3.7	Strategy use	157
8.3.7.1	Self-report measures of strategy use	157
8.3.7.2	Taped concurrent interview: Behavioral measure of strategy use	158
	(i) Passage	159
	(ii) Monitoring device (bleep)	161
	(iii) Interview coding form	161
8.4	Procedure	164
8.4.1	First phase	164
8.4.2	Second phase	165
8.4.3	Third phase	166
8.4.3.1	Training session for using the bleep	167
8.4.3.2	Interview	168
8.4.3.3	Reliability	169
8.5	Data analyses	171
CHAPTER 9	RESULTS	172
9.1	Descriptive statistics and correlations	172
9.2	Validation of selection procedures in defining RD children	176
9.3	Model testing	178
9.3.1	Preliminary LISREL analyses	180
9.3.1.1	Model modifications	181
9.3.1.2	Sensitivity analyses	183
9.3.1.3	Alternative factor structures for multifaceted constructs	184
	(i) Attributional style	185
	(ii) Strategy use	185
9.3.2	Test of hypothesized models	186
9.3.2.1	Test of proposed theoretical Model 1	190
	(i) Measurement model	190
	(ii) Structural model	196
9.3.2.2	Test of proposed theoretical Model 2	201
	(i) Measurement model	201
	(ii) Structural model	207

9.3.2.3	Test of proposed theoretical Model 3	211
	(i) Measurement model	211
	(ii) Structural model	216
9.3.2.4	Test of proposed theoretical Model 4	220
	(i) Measurement model	220
	(ii) Structural model	225
9.3.2.5	Summary of findings	230
9.3.3	Supplementary analyses	231
9.3.3.1	Test of impact of ethnicity and gender on hypothesized models	231
9.3.3.2	Replication of results using random sampling technique	231
CHAPTER 10 DISCUSSION		232
10.1	Brief overview of purpose and findings of study	232
10.2	Relationships which were consistent across NA and RD groups	234
10.2.1	Relationships between ability, the self-system, and metacognitive knowledge	235
10.2.2	The effects of self-system constructs and metacognitive knowledge on strategy use and performance	238
10.3	Relationships that distinguish NA and RD adolescents	241
10.3.1	Relationships between ability, self-system constructs, and performance	242
10.3.2	Relationship between strategy use and performance	243
10.3.3	Relationship between self-efficacy and performance	246
10.3.4	Summary	247
10.4	Distinction between metacognitive components and their relationship to cognitive processes	249
10.5	Implications of findings for Borkowski's model of metacognition	250
10.5.1	Implications for causal ordering of variables	251
10.5.2	Implications for relationships that distinguish between groups	254
10.6	Findings of secondary goals of present study	256
10.7	Educational implications	257
10.8	Limitations and strengths of present study	262
10.8.1	Limitations	262
10.8.2	Strengths	266
10.9	Future research	268
10.10	Summary and conclusions	270
REFERENCES		273
APPENDICES		
Appendix A	Information sheet/consent forms	308
Appendix B	Reading Questionnaires I and II, TARSUM measures, scoring procedures for TARSUM metacognitive strategies, and definitions of strategies used in TARSUM	311
Appendix C	Pereira-Laird, J. A. & Deane, F. P. (1995). <i>Validation of a reading self-efficacy scale for adolescents</i> . Manuscript submitted for publication	339

Appendix D	Pereira-Laird, J. A. (1995). <i>Development and validation of a self-report measure of the Reading Strategy Use (RSU) scale.</i> Manuscript submitted for publication	363
Appendix E	Pereira-Laird, J. A. & Deane, F. P. (1996). <i>Development of a behavioral measure of strategy use for adolescents: Think-Aloud Reading Strategy Use Measure (TARSUM).</i> Manuscript submitted for publication	389
Appendix F	Correlation matrices for the four hypothesized models	421
Appendix G	LISREL specification of hypothesized Models 1 through 4	424
Appendix H	Description of findings for nested sequence of models for hypothesized Models 1 through 4	439
Appendix I	Error variance-covariance matrices for structural Models 1 through 4	455
Appendix J	Supplementary analyses: The impact of ethnicity and gender on the four hypothesized models	459

LIST OF TABLES

Table 1	Summary of variables, matrices, and terms used in the analysis of covariance structures	105
Table 2	Demographic characteristics of reading disabled and normal achieving readers . . .	146
Table 3	Ability and achievement scores of reading disabled and normal achieving readers	147
Table 4	Constructs of interest and their corresponding indicators	149
Table 5	Means and standard deviations of variables in the NA and RD reader groups	174
Table 6	Results of discriminant analysis	177
Table 7	Classification results from discriminant function analysis: Group membership percentages predicted by the function	178
Table 8	Factor loadings and reliabilities of indicators for Measurement Model 1	193
Table 9	Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 1	194
Table 10	Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 1	195
Table 11	Correlations of latent constructs for NA and RD readers for Measurement Model 1	196
Table 12	Factor loadings and reliabilities of indicators for Measurement Model 2	204
Table 13	Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 2	205
Table 14	Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 2	206
Table 15	Correlations of latent constructs for NA and RD readers for Measurement Model 2	207
Table 16	Factor loadings and reliabilities of indicators for Measurement Model 3	213
Table 17	Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 3	214
Table 18	Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 3	215
Table 19	Correlations of latent constructs for NA and RD readers for Measurement Model 3	216

Table 20	Factor loadings and reliabilities of indicators for Measurement Model 4	222
Table 21	Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 4	223
Table 22	Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 4	224
Table 23	Correlations of latent constructs for NA and RD readers for Measurement Model 4	225

LIST OF FIGURES

Figure 1	Hypothesized relationships between cognitive, metacognitive, and motivational variables	5
Figure 2	Borkowski's model of metacognition (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990)	60
Figure 3	Carr et al.'s (1991) model of reading performance in achieving and underachieving readers	72
Figure 4	Hypothesized theoretical Model 1 for NA and RD readers	86
Figure 5	Hypothesized theoretical Model 2 for NA and RD readers	89
Figure 6	Hypothesized theoretical Model 3 for NA and RD readers	91
Figure 7	Hypothesized theoretical Model 4 for NA and RD readers	94
Figure 8	Schematic representation of the general LISREL model	106
Figure 9	Hypothesized relationships between cognitive, metacognitive, and motivational, variables in a single sample	121
Figure 10	Hypothesized Model 1 respecified to include correlated errors	131
Figure 11	Hypothesized Model 1 respecified for the multisample situation	133
Figure 12	Findings of revised Model 1 for NA and RD readers	198
Figure 13	Findings of revised Model 2 for NA and RD readers	209
Figure 14	Findings of revised Model 3 for NA and RD readers	218
Figure 15	Findings of revised Model 4 for NA and RD readers	227

CHAPTER 1

INTRODUCTION AND OVERVIEW

1.1 Identifying some concerns in the achievement literature

Few areas of programme development have generated as much research and controversy as the topic of learning disabilities or underachievement. In New Zealand, concern regarding services for students with learning disabilities has been relatively minimal (Chapman, 1985; Chapman & Wilkinson, 1988). However, approximately 10% or more of the New Zealand school population may have a learning disability in one or more subjects (Chapman, St. George, & van Kraayenoord, 1984; Chapman & van Kraayenoord, 1987; Walsh, 1979), and little data are available on the psychological characteristics of these children. Hence, it seems necessary to conduct research with these children so as to provide a basis for assisting teachers and psychologists in their dealings with these children.

There is much controversy in the area of learning disabilities about the identification of the population, primarily because the term "learning disability" is not a unitary concept and does not represent a homogeneous population. An underlying assumption in the field of learning disabilities is that learning disabled individuals have profiles of learning and cognitive skills that are distinctly different from those of individuals who fail to learn due to general intellectual, environmental, or educational deficiencies. Specifically, the term learning disabilities is used to describe the difficulties of children who are failing in one or more basic achievement skill areas despite normal intelligence. Often it is further stipulated that the discrepancy between ability and achievement is not primarily the result of a visual, hearing, or motor handicap, emotional disturbance, mental retardation or due to environmental, cultural or economic disadvantage, poor instructional methods, and sensory or acquired neurological deficits (American National Joint Committee on Learning Disabilities Memorandum, NJCLD, 1988, cited in Hammill, 1990). If a student's depressed performance is evident in all academic areas, the term "slow learner" could be appropriate, although other causes such as mental retardation, or motivational issues would have to be considered (Swanson & Keogh, 1990).

Difficulties in reading have long been recognized as one of the most frequent and distressing indications of learning disabilities in children. Children who have difficulties in reading are

severely impaired in their ability to acquire information from written text. Nowhere is their handicap more evident than in academic settings, where reading is the major medium for acquiring knowledge and skills (Daneman, 1991). Individuals with low reading achievement have lower academic achievement, too. It is not uncommon to find that reading comprehension ability is highly correlated with school performance ($r \geq .60$) in subjects as diverse as literature and science (Bloom, 1976; Perfetti, 1976). Moreover, reading comprehension seems to account in large part for the high correlations in performance on different school subject areas. For example, the .40 correlation between achievement in science and literature approaches zero when the effects of reading comprehension are partialled out (Bloom, 1976). In other words, reading comprehension seems to be the major common denominator in most of school learning.

In New Zealand, reading difficulties present the most common form of learning disability (van Kraayenoord & Elkins, 1994). Much attention has been directed to identifying children with reading difficulties and providing them with remedial assistance. An example of this is the Reading Recovery Programme, which is widely implemented at the junior school level (Clay, 1985). In the present study it was considered important to study young adolescent (Form 2) children, given empirical evidence that reading disability can present itself as a chronic, persistent, and pervasive problem that affects all academic areas over time (Kavale, 1988). For these children, reading instruction is still a designated part of the curriculum and represents the final setting where the teaching of reading occurs. The teaching of reading skills is particularly crucial for this age group since they are about to make their transition to secondary school where reading skills are seen as necessary for independent reading in content areas. In most countries, a significant amount of time is spent studying from textbooks by students in these upper primary grades (J. Y. Cole & Sticht, 1981). Considering the limited resources available to assist reading disabled children in the New Zealand secondary system, difficulties arising from comprehension problems experienced on entry into secondary school are likely to compound. This would in turn affect a student's future ability to master the content of different subject areas in the secondary school system.

Much of the reading disability research is dominated by single word reading and recognition tasks, which represent lower level skills. At least in the psychological literature, relatively little attention has been paid to perhaps more ecologically valid tasks such as comprehension of connected text. In the upper primary grades, as the demands of school move beyond the lower-level basic skills of decoding, a focus on the decoding aspect presents a limited view

of reading. In fact, one general conclusion that emerges from the literature is that word recognition ability alone cannot account for why some readers are better than others; the processes involved in comprehending and "absorbing" the text meaning are important determinants of reading success, too (Daneman, 1991). For these reasons, the present study investigates reading comprehension.

Two important questions pertaining to reading achievement that have yet to be investigated systematically are: (a) What causal factors determine reading comprehension achievement or impairment? and (b) How do the patterns of causal factors vary to influence reading comprehension achievement and impairment? These complex questions have stimulated decades of extensive research, aimed at identifying and explaining some salient variable or sets of variables as they relate to reading achievement.

One intuitively important variable that surfaced early from this line of research is general intellectual ability. This variable has been shown to correlate quite highly with reading achievement (e.g., Stanovich, Cunningham, & Feeman, 1984). While intellectual ability "explains" a great deal of variance in reading achievement, it by no means accounts for all of it. As a result of this lingering deficiency, researchers have begun to examine new sets of variables to determine their theoretical and practical significance for reading achievement. Among them are metacognitive, cognitive, and motivational factors. Although research in the 1980s made great strides towards understanding the contributions of these factors singly to individual performance differences, new theoretical and research paradigms are required to test more complex models.

During the past decade, a number of theories have been proposed to describe how students regulate their own learning and reading (e.g., Borkowski, Carr, Rellinger, & Pressley, 1990; Corno, 1986; McCombs, 1986, 1989; Paris & Byrnes, 1989; Schunk, 1986, 1989a; Wang & Peeverly, 1986). According to Zimmerman (1989a), most of these multicomponential approaches share a common definition of *self-regulated learning*: the extent to which an individual is an active agent in his or her own learning or reading process, metacognitively, motivationally, and behaviourally. In other words, the focus is on how students direct their own learning or reading, selecting cognitive, metacognitive, and sometimes behavioural strategies that maximize their task-related efforts (Corno, 1986). Hence, these multicomponential approaches highlight the complexity of the reading process as a variable, interactive, and dynamic process (Paris & Oka, 1986; Wixson & Peters, 1984).

Most multicomponential research involving reading achievement and impairment have used correlational methods. To understand and predict reading comprehension performance, causal models need to be developed that focus on the complex interrelationships of these cognitive, metacognitive, and motivational components. This presented the focus of the present study. One theoretical model, developed by Borkowski and his colleagues, seems pertinent to the present study in that it explicates how the relationships between cognitive, metacognitive, and motivational factors differ in reading disabled and normal achieving children (Borkowski et al., 1992; Borkowski, Estrada, Milstead, & Hale, 1989; Borkowski & Turner, 1990; Groteluschen, Borkowski, & Hale, 1990). This model suggests that reading comprehension achievement and its counterpart, underachievement, are the products of complex interactions involving cognitive, metacognitive, and motivational processes.

1.2 Overview of the research to be undertaken

The present study addresses reading impairment and achievement from a multicomponential perspective involving motivational, metacognitive, and cognitive processes. More specifically, the main goal was to determine possible differences in the causal relationships among cognitive, metacognitive, and motivational constructs in normal achieving and reading disabled children. This study also explores the implications and causal ordering of certain selected components of Borkowski's metacognitive model for understanding reading achievement and underachievement (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). The variables include intellectual ability, attributional beliefs, self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use.

Four causal models were tested across two groups of readers to address six main questions:

1. How do reading disabled (RD) and normal achieving (NA) children differ with respect to the causal relationships of the aforementioned variables? x
2. Does metacognitive knowledge mediate the relationship between motivational beliefs (efficacy, attributions) and strategy use? x
3. Does strategy use mediate the relationship between metacognitive knowledge and reading comprehension?
4. What is the relative impact of cognitive and metacognitive strategy use on reading comprehension?
5. What is the differential impact of self-efficacy and attributions on metacognitive knowledge? x

6. Do success and failure attributions differentially influence metacognitive knowledge and self-efficacy?

In the interests of clarity and to provide a conceptual map for the reader, a preview of the hypothesized causal ordering of the main variables in the study is presented below in Figure 1.

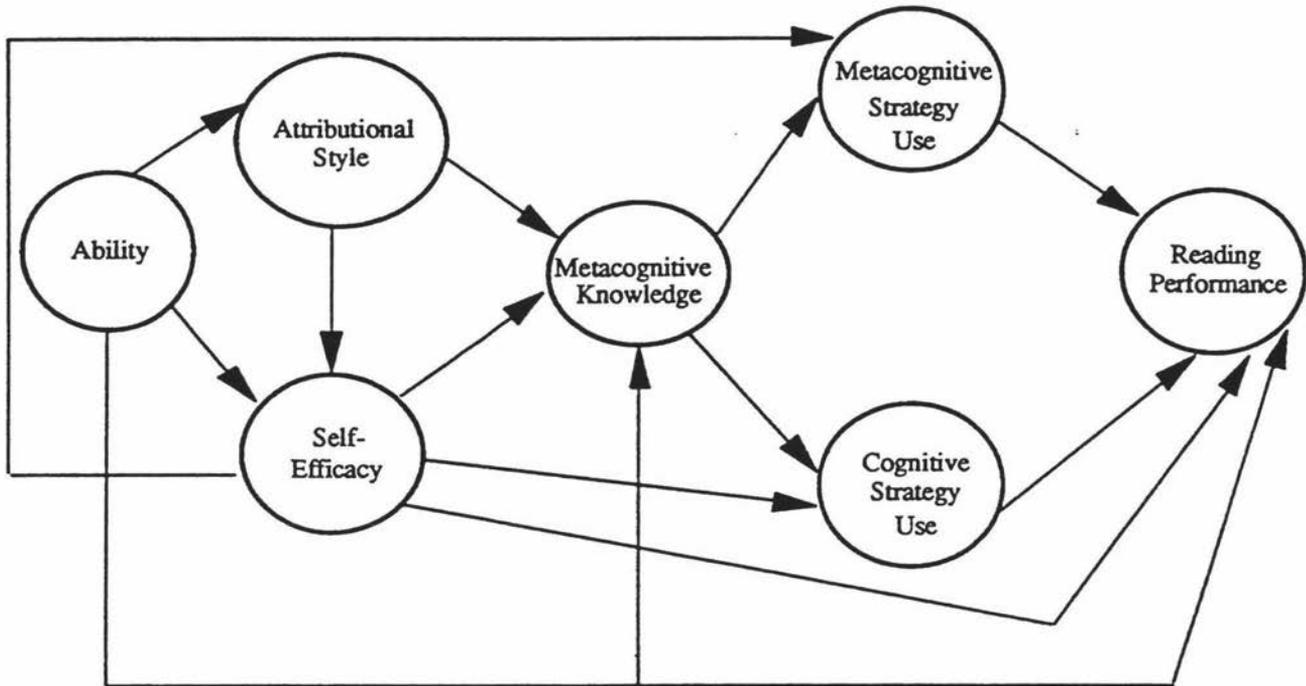


Figure 1. Hypothesized relationships between cognitive, metacognitive, and motivational variables.

1.3 Outline of chapters

The following chapters review more closely some aspects of the literature which are directly relevant to the present study. Chapter 2 provides a discussion of the definition of reading disability and the issues related to operationally defining RD children. The validity of the most frequently used method of defining RD children is examined, followed by a discussion of multifaceted approaches to defining RD children.

Chapters 3 through 5 discuss the potential loci of individual or skill level differences in reading. However, these chapters do not attempt an exhaustive review of the potential loci as two considerations delimit its scope. First, the chapters are concerned with sources of differences that might be encountered within the individual and sources that are manipulable and amenable to instruction. Second, the chapters are concerned only with accounting for the

range of factors of contemporary models of reading. As such, Chapters 3 through 5 present a discussion of three main theoretical perspectives of viewing reading disability, namely: metacognitive/cognitive, motivational, and multicomponential.

More specifically, Chapter 3 discusses cognitive and metacognitive research, beginning with theoretical and methodological issues relating to metacognition and cognition. Empirical evidence which suggests that learning disabled (LD) or reading disabled (RD) children are deficient in their cognitive and metacognitive processes is also discussed, culminating in a number of possible lines of research. This is followed by a discussion of the importance of motivation in the LD or RD individual in Chapter 4. This fourth chapter examines the role of self-efficacy and attributions in influencing achievement. The research suggests that reading disabled children often develop motivational problems as a consequence of their reading difficulties, including low self-efficacy and maladaptive attributions for success and failure. The measurement issues related to these variables are also examined.

In Chapter 5, two theoretical models that take a multicomponential perspective of reading achievement/impairment are described. A detailed description of Borkowski's model is provided as it represents the theoretical framework for the current study (Borkowski et al., 1990, 1992). These theories are useful in providing explanations for the causal relationships among cognitive, metacognitive, motivational, and performance variables and how they differ for RD and NA readers. Although such causal research with RD individuals is practically nonexistent, correlational research indicates a strong relationship amongst metacognitive knowledge, cognitive/metacognitive strategies, motivation, and academic achievement. A causal model of reading based on recent research (Carr, Borkowski, & Maxwell, 1991) that is pertinent to the present study is discussed in some depth. Taken as a whole, Chapters 1 through 5 cover conceptual and definitional, metacognitive, cognitive, and motivational aspects of reading achievement and impairment.

In Chapter 6, the rationale and scope of the present study are specified, along with the research hypotheses. Chapter 7 presents an introduction to theoretical aspects of structural equation modelling and describes the specification of the first hypothesized model into a Linear Structural Relationships (LISREL) model. Chapter 8 reports the methodology whilst Chapter 9 presents the results of the present study. Chapter 10 provides a discussion of the present research findings and the implications of these findings to theory, educational practice, and future research.

CHAPTER 2

DEFINITIONS OF LEARNING AND READING DISABILITY

Since the inception of the "learning disability" label in the 1960s, controversy has surrounded its definition, terminology, etiology, assessment, and remediation (Kavale & Forness, 1992; Torgesen & Wong, 1986). Consequently, this has led to difficulties in formulating a comprehensive theory (Weller, 1987). A large number of definitions have been proposed over the last few decades, each generally reflecting alterations in emphasis or wording rather than substantive changes (French, Ellsworth, & Amoruso, 1995). The learning disability definition has been used in two very different ways. First, it identifies a particular kind of learning problem for further scientific study (Torgesen, 1994). This has resulted in the generation of research that focuses on individual differences in learning and performance of special category children. Second, the definition has stimulated a political/social movement that endeavours to obtain as much funding as possible for this category of children (Murphy, 1992; Torgesen, 1991, 1994).

The term learning disability was originally conceptualized to describe a special education category constituted of students who did not "fit" into other exceptionalities (Mercer, King-Sears, & Mercer, 1990). These students were not achieving commensurate with their intellectual ability. To date, research has shown that there is no single characteristic or syndrome of behaviours that typifies a learning disabled (LD) individual. Instead, LD individuals exhibit an assortment of symptoms depending upon their specific disability, its severity, their unimpaired abilities, and how well the individuals handle the problem (Chalfant, 1987).

The United States federal regulations require the establishment of a severe discrepancy between ability and achievement in one or more areas to define learning disability. However, nothing is specified concerning assessment of ability or the achievement areas, nor is the severity of the discrepancy denoted (Reschly, 1987). The definitional problem can also be ascribed to the fact that so many individuals from different professions (e.g., education, medicine, psychology, and language pathology) have taken an interest in these children. These professionals inevitably have different theoretical orientations and biases that influence how they define this group of children (Swanson, 1991a).

As a consequence, considerable variation exists in how learning disabilities are determined across settings and audiences (Chalfant, 1989; Keogh, 1988a), resulting in the research literature being permeated with conflicting findings. The ambiguity of the definition and the identification criteria for learning disabilities has prompted some to call learning disabilities a "sociologic sponge" that has the flexibility to absorb individuals with a diverse array of problems (Senf, 1987).

In New Zealand, no official recognition has been given to "learning disability" as a category of special education. The needs of children with learning disabilities in New Zealand are the responsibility of educators in regular settings which means that the creation of a learning disability category in New Zealand law is very unlikely. Prospects for these children appear promising, as schools must by law meet the requirements of children with special needs. However, in practice, there are only 255 Resource Teachers to serve a population of 2968 identified underachieving students as of 1 July 1995 (Ministry of Education, 1996). It is therefore unlikely that LD children will find their needs met, given that they are competing with those with physical, mental, and intellectual impairments, who are also mainstreamed (van der Neut, 1992). As one researcher maintains "for students with learning difficulties, whose needs have barely been met, inclusion in the mainstream seems like exclusion from remedial assistance" (Chapman, 1992, p. 369).

Empirical research has seriously challenged most of the basic assumptions about learning disabilities that were advocated in the early literature (Coles, 1987; Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1993; Siegel, 1989a; Stanovich, 1986a, 1993b; Torgesen, 1986). Torgesen (1991) outlined a number of issues that continue to threaten the field of learning disabilities and still engender considerable debate. These include: (a) disputes about basic definitions of learning disability, (b) problems in validating that LD children differ from children with other kinds of learning problems either in their prognosis or their requirement for fundamentally different educational treatments, (c) challenges to the assumption that learning disabilities are physiological in origin, and (d) problems involving the conceptualization and classification of the heterogeneous LD population.

Although it is beyond the scope of this thesis to provide an exhaustive review of all these controversies, some of the criticisms that have been levelled at the discrepancy approach will be examined briefly in section 2.5, as it represents one of the core disputes in the field of

learning disability. For more complete reviews see Coles (1987), Kavale (1988), Keogh (1988a), Lyon (1994), and Torgesen (1991, 1994).

The following sections will discuss a definition that currently receives the broadest consensus in the United States, prevalence of learning disabilities, the concept of reading disability, the common elements that cut across the wide array of definitions, and finally, the discrepancy and multifaceted approaches to defining reading disability.

2.1 A definition that receives current consensus

The conceptual definition of learning disability that is currently most widely accepted is that of the American National Joint Committee on Learning Disabilities (NJCLD) (Hammill, 1990). This definition states:

Learning disabilities is a general term that refers to a heterogeneous group of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning, or mathematical abilities. These disorders are intrinsic to the individual, presumed to be due to central nervous system dysfunction, and may occur across the life span. Problems in self-regulatory behaviours, social perception, and social interaction may exist with learning disabilities but do not by themselves constitute a learning disability. Although learning disabilities may occur concomitantly with other handicapping conditions (for example, sensory impairment, mental retardation, serious emotional disturbance) or with extrinsic influences (such as cultural differences, insufficient or inappropriate instruction), they are not the result of those conditions or influences. (NJCLD Memorandum, 1988, cited in Hammill, 1990, p. 77)

The NJCLD definition has countered criticisms arising from earlier United States federal definitions and raises a number of implicit but important points. First, it highlights the misinterpretation that LD individuals form a homogeneous group (implied in earlier definitions) which in turn led to the belief that there was standardized assessment and remediation for all LD individuals. Second, it avoids the controversial phrase *basic psychological processes* (found in earlier definitions), and the ill-defined conditions that had

caused much confusion (e.g., perceptual handicaps, dyslexia, and minimal brain dysfunction). Third, it suggests that learning disabilities are not limited to children only. Finally, it suggests that there may be many different and equally valid theories of learning disability and thus efforts to identify a single, general theory of learning disability are inconsistent with the definition (see first element of NJCLD definition).

Despite the current consensus it receives, the NJCLD definition still does not make explicit how reading or learning disabilities should be operationally defined. The definition given in the Diagnostic and Statistical Manual for Mental Disorders (DSM-IV; American Psychiatric Association, 1994) attempts to do this to some extent. For example, the key diagnostic criterion for developmental reading disorder is as follows: "Reading achievement, as measured by individually administered standardized tests of reading accuracy or comprehension, is substantially below that expected given the person's chronological age, measured intelligence, and age-appropriate education" (p. 50).

2.2 Prevalence of learning disabilities

The learning disability condition has been called a hoax, yet at other times it has been claimed that up to 50% of the population is affected (Ashman & Elkins, 1994). Usually, it has been reported that prevalence estimates have ranged from 1% to 30% of the school population (Lerner, 1993; Silver, 1988). In New Zealand, prevalence figures are conservative, indicating that approximately 10% of school children have a learning difficulty (Chapman & van Kraayenoord, 1987; Walsh, 1979).

As many as 85% to 90% of the learning disabled population have some type of difficulty in the area of reading (Kaluger & Kolson, 1978). Consequently, reading disability is the most frequently reported academic problem for students with learning disabilities (Deshler, Schumaker, & Lenz, 1984; French et al., 1995; C. Norman & Zigmond, 1980) and represents the focus of the current study. Needless to say, reading is an essential skill and an inability to read has important implications at school and in later life. It is little wonder that so much attention has focused on this area.

In New Zealand, prevalence studies based on a sample of young adolescents estimated that 7% have a learning difficulty in reading vocabulary, 5% in reading comprehension, and nearly

7% in mathematics (Walsh, 1979). In another New Zealand study, it was shown that 8% of a sample of Form 1 children have difficulties in both reading and mathematics (Chapman, St. George, & van Kraayenoord, 1983). A New Zealand survey based on the perceptions of educators of special needs children in regular classrooms indicated that about 7% of students had learning difficulties or mild intellectual disabilities (H. Norman, Sritheran, & Ridding, 1984).

2.3 The concept of reading disability

2.3.1 Differences between reading disabled (RD) and poor readers

Implicit in the majority of definitions of reading disability is the assumption that an individual should be reading at a level predicted either by his or her measured intelligence, chronological age, or grade placement. The concept of reading disability is based upon observations of children who have "unexpected" reading failures despite adequate intelligence, conventional instruction, and socioeducational opportunity. It is often assumed that the causes of poor reading in such children are quite distinct from the causes of poor reading associated with lower IQ. Stanovich (1991a, 1991b, 1993a) suggested that theoretical interest in the reading disability concept has been generated by the assumption that the reading difficulties of the RD child arise from problems different from those depicting the "garden-variety" poor reader (a term coined by Gough & Tunmer, 1986). In other words, a child who is deficient in reading achievement, but whose performance is consistent with his/her aptitude (intelligence), differs in cognitive characteristics and remedial potential from a child who shows a significant discrepancy between reading achievement and aptitude. Alternatively, the reading problems of the RD child may be so much more severe than those of the garden-variety poor reader that it forms a qualitative difference.

An individual-differences model of reading disabilities, called the phonological-core variable-difference model, has been proposed by Stanovich (1988, 1991a) to explain differences between RD and garden-variety poor readers. According to the phonological-core variable-difference model, almost all poor readers have a phonological deficit. Moreover, the model proposes that further processing deficits surface as one moves in the multidimensional space from pure dyslexics to garden variety poor readers. The model rests on the assumption of specificity (see section 2.4.5) inherent in current definitions of reading disability. In addition, it assumes multidimensional continuity for reading ability and its related cognitive subskills, and this assumption has received substantial empirical support (S. E. Shaywitz, Escobar,

Shaywitz, Fletcher, & Makuch, 1992; Silva, McGee, & Williams, 1985). This model also suggests that the definition of reading disability depends on where the severity cutoff score is placed on a reading dimension (Fletcher et al., 1993).

2.3.2 Variability of terms to describe reading disability

Many different terms have been used to describe individuals who demonstrate reading problems. The term used is often dependent on who is using the label and their particular theoretical orientation. Some of the more frequently used terms are *specific reading disability*, *reading disability*, *dyslexia*, and *developmental dyslexia*. More general terms such as *poor reader*, and *learning disabled*, have also been used to describe children with reading disabilities (Kamhi & Catts, 1989; Prior, Sanson, Smart, & Oberklaid, 1995). Many of these terms overlap when they are operationally defined, with the same terms used for different children or different terms encompassing the same children (Downing & Leong, 1982).

The term *dyslexia* is frequently used within a medical setting to describe an individual whose reading problems stem from some type of neurological dysfunction (R. L. Taylor & Sternberg, 1989). In reality, reading problems can stem from a variety of causes and the incidence of "pure" dyslexia is not high (Lerner, 1993; R. L. Taylor & Sternberg, 1989). Consequently, in the educational literature, dyslexia implies a more severe reading disability for which remediation is difficult (Hallahan, Kauffman, & Lloyd, 1985).

While the distinction between learning disability and reading disability can be made in a conceptual sense, in practice, the distinction may be more difficult and perhaps *unnecessary*. Haring and Bateman (1977) consider, for example, that with minor exceptions the terms learning disability and reading disability apply to those children who are not learning to read to expectation.

2.4 Common elements or assumptions across definitions

Although Hammill (1990) contended that "consensus is near" (p. 82), research by others revealed that considerable non-conformity persists in the definition and operationalization of reading/learning disabilities (Mercer et al., 1990). The lack of a generally accepted definition of reading disabilities, along with disagreement over procedures for operationalizing reading disability definitions, makes the identification of a RD sample difficult. Chapman (1985)

suggests that this difficulty is further aggravated in the New Zealand context because of the lack of official state provision of remedial teaching services for RD children. Countries where such provisions are made (e.g., the United States) facilitate access by researchers to formally identified "RD" children, according to some set of legislated criteria. However, since New Zealand does not have such a provision, the problems associated with identification and assessment are compounded further by the absence of a RD group receiving remediation.

Considering the number of definitions that exist in the literature, it may be useful to examine the common elements that occur amongst the existing definitions, in order to provide some coherence toward an operational definition of reading disability for the current study. There are six commonly held assumptions which cut across the array of attributes associated with reading and learning disabilities which bring some consensus to the various definitions. These include (a) intrinsic problem (e.g., neurological dysfunction), (b) discrepancy between achievement and potential, (c) exclusionary criteria, (d) uneven growth pattern, (e) specificity principle, and (f) low academic performance. Each of these elements are examined in the following section.

2.4.1 Intrinsic problem

The most popular early inference was that reading disability was intrinsic to the individual as it was a direct result of neurological impairments (see Doris, 1986, 1993, for a review). Such a position is generally accepted in the field (Shinn, Tindal, Spira, & Marston, 1987), and is also adopted in New Zealand (New Zealand Federation of Specific Learning Disabilities, 1985). However, to date there is insufficient evidence to support the causal explanation of neurological bases for learning or reading disabilities. For example, in their review of the literature, Kavale and Forness (1985) found little support for neurological dysfunctions as a basis for differentiating between LD and NA children. As Clay (1972) points out, terms such as "brain damage" do not provide directions for remedial intervention. Recent work derived from an information processing paradigm also implicates a range of within-individual problems but emphasizes processing problems rather than structural ones (Swanson, 1988).

2.4.2 Ability-achievement discrepancy

One element frequently found in definitions of reading disability is the identification of a gap between what the child is potentially capable of learning and what the child has in fact

learned. Thus, the assumption is that RD individuals do not function at levels consistent with their intellectual potential. In practical terms, this has involved a statistical assessment of the difference between a child's objectively measured reading achievement and his or her general intelligence (Frankenberger & Harper, 1987; Kavale, 1987; C. R. Reynolds, 1985b). A variety of formulae and methods have been used to compute the IQ-achievement discrepancy (see C. R. Reynolds, 1985a, 1985b, 1992; Strawser, 1993). However, objections have been raised concerning the mathematical soundness of the formulae and the psychometric adequacy of the scores that are used in the equation (Hammill, 1990).

Considerable discussion has centred on the discrepancy model, and a number of investigators have described inconsistencies in decision making related to method of assessment, formulae, or the statistics applied (Moats & Lyon, 1993; Shepard, 1980). These different approaches to discrepancy measurement have led to different research sample characteristics and prevalence estimates (Moats & Lyon, 1993). Moreover, tests frequently used to measure an individual's potential and reading performance level have been questioned in terms of their validity and reliability (Farr & Carey, 1986; Stanovich, 1986b). There is also disagreement as to what and how much constitutes a meaningful discrepancy (French et al., 1995). Most studies have operationalized reading disability on the basis of a discrepancy of greater than 1.0 or 1.5 standard deviations between IQ and reading achievement scores (e.g., Fletcher et al., 1994; Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1992; B. A. Shaywitz, Fletcher, Holahan, & Shaywitz, 1992). Fletcher et al. (1993) suggested that at the very least, it must be demonstrated that the child has normal intelligence. How the discrepancy is determined has been another major area of concern for educators because standards vary across states in America and professionals disagree on the best method for determining discrepancy.

2.4.3 Exclusionary criteria

The third assumption or element concerns the exclusion of other causes which limit learning disabilities to those students whose learning problems are not explained by other diagnostic categories (e.g., mental retardation). This issue is highlighted in the last sentence of the NJCLD definition (1988, cited in Hammill, 1990; see section 2.1). However, an individual may demonstrate one of the handicaps mentioned in the NJCLD definition and still be identified as LD provided the handicap is not the primary cause of the learning problem. There is increasing acceptance amongst researchers that other conditions may occur

concomitantly with reading/learning disabilities (French et al., 1995). They also acknowledge that it is difficult to ascertain which problem is primary and which is secondary.

2.4.4 Uneven growth pattern

The concept of an uneven growth pattern refers to an uneven or irregular development of the various components of mental ability. It is assumed that mental ability is not a single capacity, but a combination of many underlying mental abilities. Lerner (1993) suggests that component mental abilities do not develop normally for LD children, resulting in intra-individual differences in the different mental processes (Kirk & Chalfant, 1984). This means that in the LD child, while some components develop in an anticipated sequence or rate, others lag in their development, thereby emerging as symptoms of the learning difficulty.

2.4.5 Specificity of academic problem

The assumption of specificity is found in nearly all legal and psychometric definitions of reading disability (Stanovich, 1986a). It is argued that a reading disabled child exhibits unexpected failures in certain, but not all, academic or educational tasks (e.g., reading, arithmetic, or spelling). The failure is unexpected, based on the child's general cognitive level, which is almost always operationally defined as performance on an IQ test. For example, RD children are posited to have a cognitive deficit that is reasonably specific to the cognitive requirements of the reading task. This assumption of specificity is an important concept because it is one of the primary means used to distinguish between RD or LD children and other low achievers (e.g., slow learners or those with borderline retardation). However, the assumption has been undermined recently by the findings of individual difference research. This research suggests that RD or LD children are deficient in a large number of cognitive processes when compared to normal achieving (NA) children. Two important factors that may have confounded the findings are the heterogeneous nature of samples used and the lack of strict IQ criteria (Stanovich, 1986b).

Stanovich (1986a, 1986b) has also proposed the "Matthew effect" hypothesis as another possible explanation for the large number of cognitive deficits found in RD children. He used the Biblical analogy (from the book of Matthew) to describe how differences in reading volume or experience between motivated and unmotivated students led to a "rich-get-richer"

and "poor-get-poorer" phenomenon. Stanovich suggested that the initial and specific difficulties of RD children may develop into more generalized deficits "due to the behavioural/cognitive/motivational spinoffs from failure at such a crucial educational task as reading" (Stanovich, 1986b, p. 389). Empirical support for Matthew effects (e.g., Juel, 1988), raises the alarming prospect that a young RD child might actually evolve into a garden-variety poor reader. In fact, research has demonstrated the great difficulty in obtaining specific impairments in RD children who have been in school for an appreciable amount of time (Torgesen, 1990).

2.4.6 Low academic achievement

Learning disabled individuals encounter a variety of academic problems. The NJCLD (1988, cited in Hammill, 1990) definition lists six specific areas in which learning disabilities can be observed (see section 2.1). Although the operational definitions in the United States vary from state to state, they include difficulty in academic learning as a primary component of their definition of learning/reading disabilities (Frankenberger & Fronzaglio, 1991). A number of researchers who examined the diversity of characteristics common to learning disability have found poor performance to be the most obvious and consistent finding (Keogh, Major-Kingsley, Omori-Gordan, & Reid, 1982; C. Norman & Zigmond, 1980).

2.5 Discrepancy approach to defining reading disability

Despite the continuing lack of agreement on both conceptual and operational levels as to what constitutes an appropriate definition of reading disability, most current operational definitions involve discrepancy (Frankenberger & Fronzaglio, 1991; Frankenberger & Harper, 1987; Prior et al., 1995; Stanovich, 1993a; Szuszkiewicz & Symons, 1993). Although there is dissension on a number of issues regarding the classification of RD children, the core conflict centres on the operationalization and theoretical justification of the discrepancy criterion (Stanovich, 1991a, 1991b). In the last few years, Stanovich (1991a, 1991b) has proposed a reexamination of the concept of discrepancy for defining reading disability. Researchers and theoreticians have been reluctant to take on the additional problems of other criteria due to the conceptual confusion surrounding the discrepancy criterion (Stanovich, 1991a). Because it is ubiquitous and the source of much theoretical perplexity, some researchers have argued that reading disability should be defined solely on the basis of tests of reading achievement (Fletcher et

al., 1989; Siegel, 1988, 1989a). As the discrepancy-based approach is still widely and currently used by researchers for defining RD children (Szuskiewicz & Symons, 1993), it warrants further discussion here, particularly with reference to its limitations.

2.5.1 Limitations of the discrepancy-based approach

The arguments surrounding the discrepancy criterion can be classified into two types. First, there are statistical and measurement complications involved in the operationalization of severe discrepancy. Specifically, this concerns the use of IQ tests to measure potential. The second relates to the construct validity of the discrepancy definition. Neither the definitions nor the criteria for their operationalization have been subjected to extensive analyses of validity and reliability. As Rourke (1983) suggested, these definitions have rarely emerged from empirical investigation.

2.5.1.1 Limitations of using IQ in the assessment of discrepancy

The use of IQ in the diagnosis of reading/learning disability has been called into question by several researchers (Aaron, 1991; Berninger, Hart, Abbott, & Karovsky, 1992; Fletcher et al., 1993; Juel, 1988; Lyon, 1989; Morris, 1993; Rispens, van Yeren, & van Duijn, 1991; Share, McGee, & Silva, 1989; Share & Silva, 1987; Siegel, 1988, 1989a, 1989b, 1990; Stanovich, 1986a, 1986b, 1989, 1991a, 1991b, 1993a, 1993b; Torgesen, 1993). There are several major problems in using the intelligence test and the issues raised may be summarized as following:

1. IQ tests do not constitute an adequate representation of a unitary construct such as potential (Stanovich, 1991b).
2. The assumption underlying the use of IQ as a measure of reading potential is that IQ exerts a unidirectional influence on reading performance but is not itself affected by the reading experience. However, research evidence has verified that reading itself is a moderately powerful determinant of vocabulary growth, verbal intelligence, and general comprehension ability (Juel, 1988; Share et al., 1989; Share & Silva, 1987; Stanovich, 1991a; Swanson, 1991a, 1993). This means that the reading problems of RD children increasingly interfere with their knowledge acquisition and, thus, their performance on IQ tests (Siegel, 1989a). Consequently, IQ may not be a good index of expected achievement because scores may drop as RD children grow older.

3. Researchers cannot agree on the type of IQ score (Full, Verbal, Performance). As a result, variations in IQ tests result in somewhat different subtypes or subgroups of RD children being identified (Stanovich, 1991b). Furthermore, the widely used Verbal Scale of the WISC-R does not include assessment of phonological awareness, a specific area considered fundamental to some prominent theories of reading disability (Berninger et al., 1992; Morris, 1993; Stanovich, 1993a; Torgesen, 1993).
4. IQ test performance may be negatively influenced by attentional disorder or reduced motivation, resulting in low scores that mask actual cognitive ability (Swanson, 1993).
5. Computation of discrepancies between achievement and IQ scores involve complex statistical procedures and diagnostic interpretation rather than merely straightforward comparison (Fletcher et al., 1993; Swanson, 1993).
6. The role of IQ cutoffs in studies of learning/reading disabilities represents another concern to some researchers (e.g., Siegel, 1988). In excluding all children with IQs below 85 or 90 from the reading disability category, one is suggesting that only children who have normal intelligence can possibly be reading disabled. Empirical findings question the validity of differentiating RD or LD children according to IQ cutoffs of 80 and above (Siegel, 1988, 1989a, 1989b, 1992).

More extreme positions query the relevance of IQ, given that small differences in learning disability identification rates have been found between methods that include IQ and those that exclude the use of IQ (Rispen et al., 1991). Basal or threshold levels of IQ have been recommended to be important elements of identification by researchers who were concerned that children with low IQs and learning problems will not have the prerequisite discrepancy because of that IQ (Leong, 1989; Siegel, 1989a, 1989b).

To overcome problems with the use of IQ, researchers have suggested a variety of cognitive processing skills as alternatives to the use of IQ scores as indices of aptitude in defining reading disability. For instance, researchers have suggested measuring the discrepancy between reading ability and listening comprehension (Spring & French, 1990), or between reading ability and phonological skills (Stanovich, 1988). However, these approaches still present problems, particularly with respect to Matthew effects (see Fletcher et al., 1993). Overall, it can be concluded that discrepancy assessment procedures are threatened by findings

that the acquisition of literacy promotes the very cognitive skills assessed by aptitude measures (Stanovich, 1993a, 1993b).

2.5.1.2 Construct validity of the discrepancy approach

According to Fletcher et al. (1993), there are two levels at which definitions based on discrepancies of IQ and achievement can be addressed. The first is replicability of the nonnormal distribution, and the second concerns the validity of distinguishing achievement impairments based on IQ discrepancies.

(i) *Replicability of nonnormal distribution*

The epidemiological work of Rutter and Yule (1975) provided the most important empirical evidence which supports the foundational assumptions of the concept of reading disability. Rutter and Yule (1975) argued that there are two subtypes of children with reading difficulties, "backward readers" (garden-variety poor readers) who read at IQ appropriate levels and "specific reading disabled" who read below expected levels according to their IQ scores. By using regression procedures, Rutter and Yule separated general reading backwardness from specific reading retardation, showing that the distribution of reading impairment was bimodal. This distinction has been widely accepted in the literature. The rationale for this distinction is that there is a somewhat discontinuous "hump" in the lower end of the distribution of reading scores (in addition to the usual central mode found near the distribution's mean). Evidence of such a hump suggests a discrete pathology model.

In the last decade, several epidemiological studies have addressed the replicability of the nonnormal distribution (Olson, Kliegl, Davidson, & Foltz, 1985; Share, McGee, McKenzie, Williams, & Silva, 1987; S. E. Shaywitz et al., 1992; Silva et al., 1985). In general, these studies did not find evidence for a hump in the distribution of reading skills. There is much converging evidence that indicates that the hump in the Rutter and Yule (1975) study was a statistical artefact, possibly involving ceiling effects on the tests (Share et al., 1987; Silva et al., 1985; van der Wissel & Zegers, 1985).

The bivariate distribution of IQ and reading was found to be normally distributed as were the univariate distributions of both IQ and reading scores. Consequently, the data from these

studies did not support the commonly held notion that reading problems as defined by discrepancy-based definition represented a unique syndrome. Rather the studies suggested that:

dyslexia goes along a continuum that blends imperceptibly with normal reading ability. These results indicate that no distinct cut-off point exists to distinguish children with dyslexia clearly from children of normal reading ability; rather, the dyslexic children simply represent a lower portion of a continuum of reading capabilities. (S. E. Shaywitz et al., 1992, p. 148)

(ii) *Validity of differentiating learning disabilities based on IQ discrepancies*

A large number of studies have addressed the degree to which reading or learning disabled individuals can be separated based on whether their achievement scores are discrepant with their IQs (Fletcher et al., 1989, 1992, 1993, 1994; Share et al., 1987; S. E. Shaywitz et al., 1992; Silva et al., 1985). In general, the results suggest that very few differences exist between IQ-based discrepancy defined children and children who are discrepant in achievement relative to age but not IQ scores (low achievement definition). More specifically, research has demonstrated that there were few differences in cognitive ability, phonological processing, and other reading related skills among RD children whose scores were discrepant or not discrepant with IQ (e.g., Hurford et al., 1993; Hurford, Schauf, Bunce, Blaich, & Moore, 1994).

However, there were some major limitations to these validity studies (Fletcher et al., 1994). For instance, both the discrepancy definitions, and the measures of achievement and aptitude varied across studies. Also, no *a priori* hypotheses were formulated to specify how discrepant and nondiscrepant poor readers might be differentiated on the cognitive, linguistic, or neuropsychological variables. Despite these limitations, in general, the classification hypothesis underlying the distinction between specific reading disabled and the garden-variety poor reader is at best weak and of questionable validity. Although these studies do not discredit the belief that reading disabilities exist, they refute the view that reading disabilities based on discrepancies with IQ form a discrete entity (Fletcher, 1992).

2.6 Multifaceted approaches to defining reading disability

A number of researchers have suggested taking a multidimensional approach to defining reading disabilities (e.g., Johnson, 1988; McKinney, 1988; Senf, 1986; Tindal & Marston, 1986). In a similar vein, Keogh (1987) suggested that rather than trying to constrain the various characteristics of LD or RD children into a single, all-encompassing definition, it would be preferable to acknowledge that these different characteristics depict different types of learning disabilities. Keogh's (1987) approach reflects the view of other writers (e.g., McKinney, 1988) who see merit in identifying reading disability subtypes. For instance, a number of investigators have made efforts to categorize these children into more homogeneous subgroups based on neuropsychological processing measures (Petrauskas & Rourke, 1979), reading and spelling abilities (Boder, 1973), and visual-spatial and linguistic measures (Spren & Haaf, 1986). Unfortunately, the subtyping literature is often confusing and contradictory, resulting in disagreement amongst theorists on the taxonomy of subtypes (Kamhi & Catts, 1989). For a more thorough discussion of reading subtypes, the reader is referred to a number of excellent reviews and books (Hooper & Willis, 1989; Kavale & Forness, 1987; Keogh, 1988a; McKinney, 1988; Senf, 1986; Torgesen, 1982a).

Identifying LD or RD children in terms of their "intraindividual" differences (Chalfant, 1989; Kirk & Chalfant, 1984) represents another form of the multidimensional approach used by some researchers (e.g., Evans, 1992). In this approach, LD children are characterized by discrepancies between intelligence and achievement, by differences between psychological processes or developmental abilities, and by differences between performance on different tasks or between academic areas. Thus, the mentally retarded child or slow learner is likely to show fairly consistent, low levels of achievement. In contrast, the LD child would show an inconsistent profile as he or she usually achieves well or adequately in some areas, but poorly in others. This concept of an inconsistent ability profile has been emphasized by Keogh (1988b), as a marker for learning disabilities as opposed to the ability-achievement discrepancy that characterizes both mental retardation and learning disabilities.

There is also growing advocacy for viewing decision making concerning definition as a multistage rather than a single stage process. Increasingly, multistage or multidefinitional assessments are advocated as one way of improving the diagnosis of reading disabilities, thus reducing the number of false-positive findings (Berninger et al., 1992; Kavale, Forness, & Lorschach, 1991; Leong, 1985, 1987). For instance, Kavale et al. (1991) suggested that a

suitable operational definition for a multivariate phenomenon like learning disability can be obtained by extending the view that a definition can include a number of elements or concepts (that may be defined operationally) rather than a single element. The elements chosen, known as "operational interpretations", could be based on validated components representing the areas which have sufficient theoretical importance to the learning disability concept. Kavale et al. illustrated how operational interpretations may be combined to form a comprehensive operational definition for classifying learning disabilities. They proposed a five level hierarchical sequenced operational definition of learning disabilities based on a number of operational interpretations extracted from what is known about the learning disability phenomenon. In their model, discrepancy was represented in the first level followed by four additional levels. The presence of academic problems and poor learning efficiency were represented in levels 2 and 3, followed by the presence of specific skill deficits on the last two levels.

In a similar vein, Leong (1985, 1987) proposed a two-stage framework of diagnosis incorporating standardized achievement tests and teachers' estimates of children's performance at the assessment stage, and more refined, individual tasks at the diagnostic stage. Empirical support is provided by Berninger et al. (1992) for a two-stage approach analogous to the framework of Leong (1985, 1987). In stage 1 classroom teachers were asked to identify "low functioning" children in a variety of beginning reading skills (e.g., orthographic and phonological coding, and perceptual motor integration skills). Stage 2 concentrated on a more comprehensive psychoeducational assessment for children whose problems persisted.

The abovementioned multifaceted approaches appear to be a promising avenue for defining RD children. However, to date little progress has been made in employing such approaches, particularly with regard to the multidefinitional approach. The significance of these approaches lies in their attempt to deemphasize intelligence, especially discrepancy notions, in conceptualizing learning or reading disabilities. To date, intelligence has received far too much consideration in definitions of learning or reading disability, and this has diverted attention from more important variables associated with the concept of learning disabilities (Kavale, 1993). Kavale (1993) suggested that intelligence need not be a "cornerstone" variable for learning or reading disabilities. Rather it could be regarded as one among many variables that need to be considered to capture the true essence of learning/reading disabilities.

2.7 Conclusion

At the moment there is no commonly accepted criterion for identifying RD children, with identification procedures varying across researchers. This review highlights a number of issues relating to the discrepancy approach in defining RD children. The almost total reliance on discrepancy as an identification variable for reading disabilities is neither conceptually nor psychometrically justifiable for a number of reasons. First, lower IQ may be a result, rather than a cause, of reading disability (Stanovich, 1986b). For instance, the intelligence scores of many older children with reading disabilities are affected by their lack of success in basic school subjects because of their inability to read texts and other supplemental material (Stanovich, 1986b; van den Bos, 1989). This phenomenon renders discrepancy definitions of learning and reading disabilities problematic as children grow older. Second, studies have failed to find evidence of cognitive differences between RD and garden-variety poor readers. Third, identification procedures using discrepancy as the *sole criterion* in identifying RD children fail to accurately identify RD students, because discrepancy is a necessary but *insufficient* criterion for determining learning or reading disabilities (Payette, Clarizio, Phillips, & Bennett, 1995; C. R. Reynolds, 1985a, 1985b). For these reasons, a focus on the discrepancy criterion for diagnosing reading disabilities may not be appropriate. Such a unidimensional approach is unlikely to capture the complexity of a multivariate phenomenon like reading disability.

Instead, a multifaceted approach incorporating multidefinitonal components may be a more sensible way of defining reading disability. Because reading is a complex process, taking a multidefinitonal approach to defining reading disabilities is appropriate as it is more likely to capture the complexity of the construct of reading disability. A number of elements or defining features, most of which are considered fundamental indicators of the reading/learning disability concept, can be utilized in this multidefinitonal approach. These elements can be determined from the common elements present across definitions (see section 2.4). Such an approach was used in the present study.

Despite the uncertainty regarding sample definition, there appears to be some consensus regarding the primary importance of cognitive and metacognitive skills for successful academic achievement. This will be discussed in the following chapter.

CHAPTER 3

METACOGNITION AND COGNITION AND THEIR RELATIONSHIP TO READING

"Metacognition" is an important theoretical construct in cognitive psychology. The term was coined by Flavell (1976) in his research on children's memory processes. It was extended to the area of reading by Brown (1980) who suggested that efficient reading, whether to understand, or to remember, is characterised by the use of metacognitively-guided strategies. The area within metacognition that relates to reading comprehension has been termed metacomprehension (Weinstein, 1988).

Over the past two decades, research within a metacognitive view of reading ability-disability has led to a somewhat different view of the etiology of reading disability. Cognitive views posit a lack of ability in one or more component areas of cognitive processing, whereas a metacognitive view posits a lack of strategy use, the so-called *production deficiency* (Brown, 1974; Flavell, 1970). Central to many cognitive views of learning is the concept of *control* or *executive processes* and this is critical to the understanding of metacognition (Brown, 1975; Campione, 1987). Kuhn (1992) states that the executive process involves selecting, controlling, and monitoring the cognitive skills necessary for achieving a problem solving goal. This executive process is an important part of metacognition, and is considered a kind of "higher-level" cognition.

A number of researchers have provided empirical support for the important role that metacognition plays in cognitive development and academic learning (Ghatala, 1986; O'Sullivan & Pressley, 1984; Paris, Wasik, & van der Westhuizen, 1988). Impressive though this work is, there are still some difficulties with the concept of metacognition that could benefit from additional research (Kurtz, 1991; R. E. Reynolds & Wade, 1986). The difficulties relevant to this discussion centre on definitional, conceptual, and methodological issues relating to metacognition and establishing a clear causal relationship between the components of metacognition and cognitive performance. Kurtz (1991) suggested that a critical area in the concept of metacognition is its distinction from cognitive processes which are not metacognitive.

In this chapter, the theoretical construct of metacognition is defined. In addition a number of theoretical and methodological issues related to this construct are discussed. Then empirical research supporting its crucial role in the effective learning and reading of normal achieving and poor or learning disabled readers is discussed.

3.1 Definition of metacognition

Despite the growing body of research supporting the position of metacognition in cognitive theories, there continues to be disagreement surrounding its definition and operationalization (Kurtz, 1991; R. E. Reynolds & Wade, 1986). Several early definitions of metacognition have been proposed (Brown, 1978, 1980; Flavell, 1976, 1979). Two salient features emerge from these definitions and these have received wide acceptance. The first salient feature is that metacognition involves metacognitive knowledge or knowledge about thinking. The second salient feature is that metacognition involves self-regulation or control of thinking. These features are clearly captured in the definition provided by Forrest and Waller (1981):

metacognition is a construct which refers to what a person knows about his/her cognition (in the sense of being consciously aware of the process, and further, of being able to tell you about them in some way), *and* his/her ability to control these cognitions (in the sense of planning cognitive activities, choosing among alternative activities). (p. 2)

As Brown, Bransford, Ferrara, and Campione (1983) have shown, these two components of metacognition confound the problem of definition. For instance, whereas the knowledge component is basically storable, stable, late-developing, and fallible, the regulatory component is not storable, quite unstable, relatively age independent, and task and situation dependent. Nevertheless, this two-faceted definition still dominates the field (Brown et al., 1983; Schneider & Pressley, 1989; Yussen, 1985).

In more recent definitions of metacognition, Borkowski and his colleagues (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990) have elaborated these two subdivisions of metacognition to include complex interactive components (Larson & Gerber, 1992). This model differs from earlier conceptions of metacognition in that it has been expanded to include affective and motivational constructs. As this model of metacognition explicates relations between *metacognitive*, *cognitive*, and *motivational*

constructs, and represents the theoretical basis of the current study, it will be discussed in detail later (see Chapter 5, section 5.2). For the present, components in the dominant two pronged definition will be discussed. It should be noted that this includes some of the components in Borkowski's model of metacognition.

The first aspect of metacognition has been variously described as awareness (Carr et al., 1991), knowledge about cognition (Brown, 1975), cognition about cognition (Flavell, Miller, & Miller, 1993), metacognitive knowledge (van Kraayenoord, 1986), stable and stable information about one's own cognitive processes (Baker & Brown, 1984a; Slife, Weiss, & Bell, 1985), cognitive self-appraisal (Paris et al., 1988; Paris & Winograd, 1990a), and specific strategy knowledge (Borkowski et al., 1990). In the present study, metacognitive knowledge will be used to denote the first aspect of metacognition: awareness.

Paris, Lipson, and Wixson (1983) categorised metacognitive knowledge in terms of declarative, procedural, and conditional knowledge, because cognitive self-appraisal involves *what* you know, *how* you think, and *when* and *why* to apply strategies. Declarative knowledge about reading includes stable knowledge about the structure and goals of the task. Procedural knowledge refers to information about how to perform strategies related to each of the reading subprocesses. Conditional knowledge involves application and transformation of strategies to accommodate various reading conditions and circumstances. However, Flavell (1985) suggested that metacognitive knowledge is not always cleanly categorizable, but highly interactive.

The second aspect of metacognition has been variously referred to as executive processes (Groteluschen et al., 1990), metacognitive acquisition procedures (Borkowski et al., 1990), metacognitive strategy use (Pokay & Blumenfeld, 1990), self-regulatory strategies (Pintrich & De Groot, 1990), self-management (Paris & Winograd, 1990b), cognitive self-regulation (Brown, 1978), and metacontrol skills (Larson & Gerber, 1987). In the present study, metacognitive strategy use will be used to denote the second aspect of metacognition: executive processes.

3.2 Theoretical issues concerning metacognition

In this section, some of the conceptual difficulties concerning metacognition will be discussed. These include defining what a strategy is, distinguishing metacognitive components, and

differentiating cognition from metacognition. The difficulty in linking metacognition to reading disability is also discussed.

3.2.1 The concept of strategy

The centrality of "strategies" for understanding text is widely acknowledged (e.g., Alexander & Judy, 1988; Brown et al., 1983; Garner, 1990; Paris, 1988; Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989). However, a definition of strategy is required since there is no consensus regarding this (Paris, Wasik, & Turner, 1991). Related discussion in the literature concerning differentiation of cognitive and metacognitive activity does little to clarify the central notion of a strategy (see section 3.2.3). Gerber (1983) warned that the concept of strategy would become progressively obscured and lose its apparent explanatory power unless its meaning was clarified. The reader is referred to Pereira-Laird and Deane (1996) for more detailed information concerning this issue. For the purposes of the present study, Rowe and Rayford's (1984, cited in Pritchard, 1990) definition of a strategy is adopted, where a strategy is defined as a *deliberate action* that readers take voluntarily to comprehend what they read.

A variety of labels have been used by researchers across studies to describe the strategies that are used with textual material. Among other labels, these include "study strategies" (Brown et al., 1983; Wade, Trathen, & Schraw, 1990), "learning strategies" (Alvermann & Moore, 1991; Pintrich & De Groot, 1990; Weinstein & Mayer, 1986), "reading strategies" (Chan, 1995; Paris et al., 1991), "reading processing strategies" (Pritchard, 1990) and "text-processing strategies" (Paris et al., 1991). Irrespective of the label, the reader uses a combination of metacognitive strategies to monitor understanding (Baker & Brown, 1984a; Brown et al., 1983; Garner, 1987; Paris et al., 1991) and cognitive strategies to learn, remember, and understand text (Pintrich & De Groot, 1990).

3.2.2 Distinction between components of metacognition

There has been considerable debate in the last decade or two about the distinction between knowledge about cognition and regulation of cognition. For instance, Brown and Palincsar (1982) commented that these two forms of metacognition are "closely related" and that "each supports the other recursively" (p. 1). They suggest that any attempt to separate the two constructs results in "oversimplification".

On the other hand, Lawson (1984) argued that the term metacognition should only apply to metacognitive knowledge rather than to control processes. He sees the two dimensions as logically distinct, with metacognitive knowledge resulting from the operation of executive control processes, but acknowledges that metacognitive knowledge is one source of influence on executive processes. Nevertheless, it is argued here that metacognitive knowledge involves executive control and as such both are interwoven components of metacognition. To argue theoretically that the components of metacognition are interwoven does not imply that research into metacognition will not distinguish between knowledge of cognition and control of cognition (van Kraayenoord, 1986). In fact, there are a number of researchers who have empirically confirmed the occurrence of both facets (Brown, 1981; Cornoldi, 1990; Cross & Paris, 1988; Pereira-Laird, 1996; Schraw & Dennison, 1994). It is argued here that the separate and joint contributions that each facet makes to reading comprehension are worthy of investigation. To date, the differential effects of these metacognitive components to reading comprehension performance are unclear and remain uninvestigated. The present study attempts to explore and clarify these effects.

3.2.3 Distinction between cognition and metacognition

The confusion over the definition of metacognition parallels a difficulty in distinguishing between cognition and metacognition. More specifically, there are problems differentiating between metacognitive strategies and cognitive strategies (A. St. George, van Kraayenoord, & Chapman, 1985). It appears that at times the distinction is unclear with researchers often being quick to label any strategic action as metacognitive (Brown et al., 1983). Perhaps a resolution of the conflict is that proposed by Flavell (1979) who states that "Cognitive strategies are invoked to *make* cognitive progress, metacognitive strategies to *monitor* it" (p. 909). In other words, one way of viewing the relationship between them is that cognition is involved in doing, whereas metacognition is involved in choosing and planning what to do and monitoring what is being done. However, Brown (1987) notes that in some cases a strategy may be invoked for either purpose or may achieve both a cognitive and a metacognitive goal. She describes an example as follows: "Asking yourself questions about the chapter might function either to improve your knowledge (a cognitive function) or to monitor it (a metacognitive function)" (p. 67). Thus, a specific activity can represent the strategy itself (looking for main points), its monitoring function (a metacognitive activity), and a reflection of metacognitive knowledge that the strategy is appropriate for the given situation.

It is argued here that metacognitive and cognitive strategies are distinct components as they have been distinguished empirically in a large number of research studies (Anthony, 1994; Pereira-Laird, 1996; Pokay & Blumenfeld, 1990; Pressley et al., 1991; Zimmerman & Martinez-Pons, 1990). One such study is that by Pintrich and De Groot (1990) who successfully used factor analysis to differentiate these strategies. Even though there was a substantial correlation (.83) between these strategies, the metacognitive strategy use measure predicted the students' academic performance much better. The authors concluded the two types of strategies were conceptually distinct based on their finding that cognitive strategy use without the concomitant use of self-regulation was not conducive for academic performance.

A recent study by Pereira-Laird (1996), using more powerful techniques such as confirmatory factor analyses, has also provided evidence for the distinction between cognitive and metacognitive strategies, particularly in the area of reading. Analyses of the strategy use items revealed the presence of two factors, providing discriminant validity for cognitive strategy use and metacognitive strategy use. Confirmation of several predicted conceptual relationships between the two strategy use scales and other reading related factors (i.e., achievement, anxiety, perceived value, and efficacy) provided evidence of construct validity.

3.2.3.1 Activities involved in cognitive and metacognitive strategies

Although the essence of metacognition is still being debated, there seems to be more consensus about the mental activities constituted by it (Haller, Child, & Walberg, 1988). These metacognitive strategies involve planning, monitoring, and regulation activities that take place before, during, and after any thinking act such as reading (Brown, Armbruster, & Baker, 1986; Paris & Lindauer, 1982; Pintrich & Schrauben, 1992).

Planning involves setting goals for reading or studying, skimming a text before reading, formulating questions before reading a text, and task analysis (Pintrich & Schrauben, 1992). Monitoring strategies include tracking attention while reading a text, and self-testing (through paraphrasing, summarizing, or self-questioning) to check for understanding. They also include the use of test-taking strategies (monitoring one's speed and time in an examination or test), and evaluating activities such as confirming assumptions and making predictions. Thus, monitoring strategies alert the reader to breakdowns in attention or comprehension so they can be repaired or "fixed" through the use of regulating strategies (Pintrich & Schrauben, 1992). Regulation activities consist of compensatory strategies to redirect and bolster faltering

comprehension. They are closely related to monitoring strategies. When readers realize they do not understand a portion of text, they can use a variety of regulation strategies to restore comprehension. These can include, among others, rereading, self-questioning, slowing the reading pace, and contrasting textual information with prior knowledge (Pintrich & Schrauben, 1992).

In contrast, cognitive strategies, refer to integrating new material with prior knowledge. Cognitive strategies that students use to acquire, learn, remember, retrieve, and understand the material while reading include rehearsal, elaboration, and organizational strategies (Pintrich & Schrauben, 1992; Weinstein & Mayer, 1986; Zimmerman & Martinez-Pons, 1986, 1988).

Rehearsal strategies involve the repetition of to-be-learned information. They can include repetition of items in an ordered list, copying important portions of the material, underlining or highlighting the material, and verbatim note taking (Weinstein & Mayer, 1986). Elaboration strategies involve creating linkages between an individual's existing knowledge and the new information. These include using mental images, summarizing or paraphrasing text, describing how information relates to existing knowledge, creating analogies, and answering questions (Weinstein & Mayer, 1986). Organizational strategies involve rearranging the information in a way that makes learning meaningful. They include sorting, categorizing, outlining a passage, creating a hierarchy, mapping important concepts that have been read, or creating a diagram to show the relationships of important ideas (Weinstein & Mayer, 1986).

(i) *Reading for meaning and reading for remembering*

The specific strategies for effective reading differ according to the student's goal for reading, either reading for meaning or reading for remembering. Reading for meaning is basically an attempt to comprehend and this often involves comprehension monitoring. On the other hand, reading for remembering (i.e., learning) encompasses strategies of reading for meaning and strategies for retaining textual material. In other words, for effective learning to occur, both strategies for understanding and strategies for remembering are necessary. Learning involves a focus on what the student does while reading; this includes identifying the main ideas in the reading, and testing one's understanding and mastery of the material. The present study includes an examination of strategies for both reading for meaning and learning.

3.2.4 Conceptual difficulties in linking metacognition to reading disabilities

The relationship between metacognition and intelligence has received strong theoretical support in current conceptions of intelligence (Allon, Gutkin, & Bruning, 1994). Therefore, if metacognitive problems are used to explain reading disabled children's deficiencies, it would logically lead to the conclusion that they have lower intelligence. Such a conclusion would undermine both the discrepancy notion and the assumption of specificity. Thus, the linking of metacognitive constructs to reading disabilities has been seriously questioned by Stanovich (1986c) who stated that, "the many recent discussions linking deficiencies in metacognitive functioning with reading disability are actually embarrassing to the assumption of specificity because many individual-difference theorists regard metacognitive awareness as a *sine qua non* of *intelligence!*" (p. 233).

Alternatively, Wong (1985, 1987, 1991) has highlighted the relevance of the theoretical construct of metacognition to reading disabilities. Wong (1991) suggests that the problem highlighted by Stanovich (1986c) can be avoided by invoking Stanovich's (1986b) hypothesis of the "Matthew effects". This would render metacognitive problems as a *second-order* (rather than a first-order) difficulty involving a joint by-product of deficient reading experiences and motivation. Thus, one reason that reading disabled children appear to have mild but pervasive cognitive deficits is the development of characteristics that are considered secondary to their reading failure (Torgesen, 1990, 1991). These secondary characteristics are consequences of reading disability, rather than causes of it. Once acquired, they can further exacerbate the children's ability to perform successfully in school by limiting their ability to acquire new information.

It can also be argued that linking domain specific metacognitive functioning to reading disabilities may not undermine the assumption of specificity, as such a deficit is not likely to depress intelligence. This is because domain specific metacognition represents only one aspect of metacognition. For instance, reading disabled children may have a metacognitive deficit in the reading domain but not in other academic domains. Domain specificity of metacognition has been alluded to by Lester (1988). In fact, a number of researchers have studied metacognition with respect to specific domains (e.g., Garofalo & Lester, 1985; Schoenfeld, 1985).

3.3 Methodological issues concerning cognition and metacognition

Any investigation of the characteristics and value of metacognition and cognition research also depends on instrumentation. Little agreement has been reached about the most effective way of measuring metacognition (McLain, Gridley, & McIntosh, 1991). As a result of the difficulty associated with measuring metacognitive aspects of reading (Jacobs & Paris, 1987), a wide variety of techniques have been used to measure metacognitive knowledge, and cognitive/metacognitive strategy use of readers (Garner, 1982, 1987). These include verbal reports (interview, think-aloud measures), questionnaires, error detection, computer and video studies, reaction time and forced choice tasks. In this section, only commonly used measures will be discussed.

3.3.1 Verbal report

One of the most frequently used methods is the verbal report which includes interviews and *think-aloud* procedures. These procedures have been used to measure both metacognitive knowledge and strategy use. The interview involves asking children open-ended questions about how they think, read, learn, and remember; or what they might think and do in a hypothetical situation. Interviews produce *retrospective* verbalizations, for they elicit reports of metacognitive/cognitive activity already completed. In contrast, think-aloud protocols generate *concurrent* verbalizations about a reading activity that is temporarily interrupted at frequent predetermined and cued points for provision of the verbal report. Research seems to support the use of concurrent reporting (think-aloud protocols) in preference to retrospective reporting as being a more reliable measure of students' strategy usage (Ericsson & Simon, 1980; Meichenbaum, Burland, Gruson, & Cameron, 1985; Meyers & Lytle, 1986).

A number of criticisms have been levelled at verbal reports (Garner, 1987). One major concern is that children may lack the language and verbal facility to discuss the mental events. A second criticism is that verbal report data may include subjects' rationalizations, fabrications, and elicited mimicry because of the demand characteristics of the situation. Moreover, verbal responses may be ambiguous and may be edited by the subjects to conform to their expectations about what the investigator wants to hear. There is some evidence that asking questions during reading alters the way in which subjects process the text, and in addition, interacts with other variables to influence recall (Wade & Trathen, 1989). A third

concern about verbal report data is that their reliability and validity are rarely assessed. For instance, the cueing offered by instructions and probes may invalidate the data.

Another difficulty with verbal reports is that they may not reflect what students are really doing. In other words, students may be engaged in metacognitive strategy use but not report it, or they may report it inaccurately (R. E. Reynolds & Wade, 1986; Schneider, 1985). As an explanation for such anomalies, Flavell (1979) suggested that metacognitive knowledge can be either conscious or automatic. If the activity has reached the stage of automaticity, it is unlikely that children will be able to report its occurrence. Brown (1981) suggested that the same claim can also be made about monitoring activities since they are likely to be "over-learned" in proficient readers. Furthermore, children may not always use the strategies that they have in their repertoires. Finally, even if children use the "correct" strategy and are able to verbalize it, they may not use it effectively or efficiently (R. E. Reynolds & Shirley, 1988). These difficulties may help explain why metacognitive activities, as measured by verbal report data, are not consistently correlated with performance of cognitive behaviours (Baker, 1985). Thorough reviews of the problems surrounding verbal report data are available elsewhere (see Afflerbach & Johnston, 1984; Garner, 1987; Nisbett & Wilson, 1977).

Despite the criticisms, Ericsson and Simon (1980) have asserted that verbal report data can provide a great deal of knowledge about students' metacognitive activity. They suggest that the following factors can increase the accuracy of verbal report data to an acceptable level: (a) not requiring subjects to verbalize at the same time as they are performing the task, (b) minimizing the time interval between processing and reporting, (c) emphasizing to subjects that they should focus primarily on the task, rather than on the reporting of their processes, and (d) collecting data from other sources to check for convergence of data.

3.3.2 Self-report questionnaires

Although relatively infrequent, some researchers have used multiple-choice questionnaires to investigate what children know about metacognition. Two multiple-choice questionnaires currently reported in the literature to measure metacognitive knowledge include the Index of Reading Awareness (Jacobs & Paris, 1987) and the Metacomprehension Strategy Index (Schmitt, 1988, cited in Schmitt, 1990). These multiple-choice formats avoid many of the

problems of verbal facility and elicited answers. However, they force choices on questions and contain an element of guessing not found in verbal reports.

Self-report questionnaires have also been recently used by researchers to measure both cognitive and metacognitive strategy use (Chan, 1995; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990; Schraw & Dennison, 1994). Although this method can overcome some of the limitations of verbal report data, it has certain inherent limitations. These limitations were outlined by Pokay and Blumenfeld (1990) as tapping students' perceptions of what they are doing rather than the accuracy of these perceptions, and not accounting for the efficiency or appropriateness of strategy use. Appropriateness of strategy use was not accounted for because students report strategies they use generally, rather than strategies they use to solve specific tasks. Thus, responses of students in questionnaires refer only to student perceptions and not necessarily to accuracy, effectiveness, or appropriateness of strategy use.

3.3.3 Error detection

One useful approach to gathering data on metacognitive self-regulation processes involves altering texts to include impediments to comprehension (e.g., inconsistent text, nonsense words), and then observing whether students notice them (Baker, 1985; Baker & Brown, 1984a; van Kraayenoord, 1986; Zabucky & Ratner, 1986). This method assumes that if students are engaging in monitoring activities, they will notice the irregularity and take some corrective action. Studies of awareness of deliberately inserted ambiguities, inconsistencies and errors have shown that both good and poor readers do monitor their comprehension by detecting inconsistencies (Bos & Filip, 1984; Winograd & Johnston, 1982). However, poor and learning disabled readers do not monitor their reading spontaneously or as well as their normal achieving peers (e.g., Bos & Filip, 1984; Zabucky & Commander, 1993).

The error-detection paradigm used by researchers to investigate comprehension monitoring is not without limitations. For instance, Baker (1985) discovered that some subjects, when interviewed, had noted the inconsistency but had altered their understanding of the story in order to make it fit meaningfully into the scenario they were creating. As a result of her extensive research, Baker concluded that even with careful analysis of on-line data, errors in inferring what readers do and do not think about their comprehension strategies are common.

3.3.4 Conclusion

In conclusion, there seem to be difficulties mainly with indicators of metacognitive functioning. Verbal report and self-report data are easy to obtain but they can be unreliable and ambiguous. Data obtained from inserting inconsistencies into texts do not seem to encourage the types of metacognitive responses expected. One possible solution to these difficulties seems to be the use of multiple indicators of metacognition (Ericsson & Simon, 1980; R. E. Reynolds & Wade, 1986; Schneider, 1985). While this approach may not guarantee accurate estimates of metacognitive functioning, it will decrease the number of false positive outcomes (R. E. Reynolds & Wade, 1986). In line with this, the present study incorporated self-report questionnaires to supplement think-aloud protocols of strategy use so as to obtain more valid estimations of students' strategy usage. The subjects' in-process reports of strategies can be considered as evidence of the covert mental processes that occur during reading. Although the validity of such verbal reports is not universally accepted (Garner, 1982, 1987; Nisbett & Wilson, 1977), think-aloud procedures have been widely defended as a valid means of gaining information about the reading process (Afflerbach & Johnston, 1984; Baker & Brown, 1984a; Ericsson & Simon, 1980; Hynd, Alvermann, & Carter, 1984). By using a set of different methods that do not share the same sources of error, convergent data on actual strategies used by readers can be collected. Furthermore, the data are not as vulnerable to the same sources of invalidity (R. E. Reynolds & Wade, 1986).

3.4 Empirical research regarding the relationships between cognition, metacognition, and performance

Knowledge of how metacognition predicts and affects cognitive performance and strategy use is crucial if metacognition is to survive as a functional and important construct. In fact, R. E. Reynolds and Wade (1986) stated, "If there is little or no relationship, then one could argue that the construct of metacognition has no real utility; on the other hand, evidence of a clear causal link between metacognition and cognitive behaviour would essentially validate the construct" (p. 311).

In the last decade, increasing attention has been devoted to the role of strategy deficits in learning and reading disabilities, partly as a result of the documented importance of metacognitive and cognitive strategies for independent learning in normal achievers (Brown et al., 1983; Pressley et al., 1990; Pressley, Symons, Snyder, & Cariglia-Bull, 1989).

In this section, research is reviewed regarding the relationships between metacognitive knowledge, metacognitive/cognitive strategy use, and performance with normal achieving (NA) and reading disabled (RD) or learning disabled (LD) children. The review will also include research on poor readers as considerably more research has been done with this group of readers than with LD and RD children concerning the relationship between metacognition and reading comprehension (Rottman & Cross, 1990).

3.4.1 Relationship between metacognition and comprehension

An indirect relation between metacognition and comprehension has been demonstrated by showing that metacognitive status varies with age (Cross & Paris, 1988; Paris et al., 1991). Other studies have demonstrated a direct relation between comprehension and metacognition by showing that metacognitive status varies with reading ability (e.g., Paris et al., 1991).

The role of metacognition has also been shown in studies designed to improve reading comprehension by increasing children's awareness and use of strategies (Cross & Paris, 1988; Paris & Jacobs, 1984; Paris, Saarnio, & Cross, 1986). In a meta-analysis of 20 previous studies it was concluded that the average effect size (.71) for the effect of metacognitive intervention on reading performance was one of the larger ones found in educational research (Haller et al., 1988).

The relationship between metacognition and comprehension has also been studied in poor readers. Reviews suggest that less competent readers have limited metacognitive knowledge and have difficulties in monitoring and self-regulating comprehension (Baker & Brown, 1984a, 1984b; Garner, 1987; Haller et al., 1988; Mateos, 1989, cited in Mateos & Alonso, 1991; Wagoner, 1983; Wong, 1987). In summary, the research reviewed here suggests that metacognition may play an important role in reading comprehension.

3.4.2 Relationship between metacognitive knowledge and performance

Research examining metacognitive knowledge about reading strategies was often characterized by comparisons between good and poor readers or between younger and older readers. Such a research paradigm has been used to demonstrate differences between good and poor readers

and is clearly based on the assumption that the factors that differentiate these two groups of readers are the causes of poor readers' disability.

Considerable evidence has accumulated to indicate that reading performance is positively related to metacognitive knowledge about reading (Cross & Paris, 1988; Paris & Oka, 1986; Paris & Winograd, 1990a; Schneider & Pressley, 1989). Lipson, Irwin, and Poth (1986) reported correlations ranging from .30 to .60 between metacognitive knowledge and reading comprehension.

The relationship between metacognitive knowledge and performance also applies to RD children, who may not have acquired as much metacognitive knowledge as other students (Borkowski et al., 1989; Johnston & Winograd, 1985; Wong, 1987). Although there have been inconsistent results, most studies have shown that RD students lack metacognitive knowledge in reading (e.g., Pintrich, Anderman, & Klobucar, 1994; Short & Ryan, 1984). These studies have noted that incompetent readers lack knowledge about the purposes of reading and effective reading strategies, and are less aware of the variables that influence reading performance (Baker & Brown, 1984a; Brown, 1980; Garner, 1987; P. J. Moore, 1982; Paris & Jacobs, 1984; Winograd, 1984; Wong, 1987).

In summary, most of the empirical research suggests that metacognitive knowledge is an important predictor of reading performance in both normal achieving and reading disabled children.

3.4.3 Relationship between cognitive/metacognitive strategy use and performance

Successful readers have been shown by research to utilize a variety of cognitive and metacognitive strategies to improve their reading and learning (French et al., 1995; Garner & Alexander, 1989; Pressley, Borkowski, & Schneider, 1987). Research on learning strategies shows that strategies develop with age and with ability level such that older and more able students are more likely to spontaneously use appropriate cognitive strategies in a learning task than younger and less able students (e.g., Baker & Brown, 1984a; Brown et al., 1983; Paris et al., 1991; Weinstein & Mayer, 1986). The learner seems to gradually develop the ability to spontaneously use a cognitive strategy (Brown et al., 1983) and passes through a stage of "production deficiency" (Brown, 1974; Flavell, 1970), where he/she possesses the

strategy but does not use it spontaneously during learning. For example, Weinstein and Mayer (1986) concluded that by age 7, children were frequently able to use rehearsal strategies if specifically instructed to do so, but may not use them spontaneously with a learning task until age 12.

Learning disabled students have been shown to have a number of deficits in cognitive strategies. Studies of memory performance have shown LD students to be deficient in the use of mnemonic strategies and the clustering of semantically related material (Bauer, 1979; Torgesen, Murphy, & Ivey, 1979). Similarly, studies on selective attention have shown that LD students do not use rehearsal techniques to focus on critically important material to the same extent as NA learners (e.g., Hallahan & Reeve, 1980). Other studies indicate that LD students are deficient in strategies for test-taking, organizing information, and summarization (Scruggs & Mastropieri, 1986; Swanson, 1989; Winograd, 1984). Cognitive strategy deficits have also been found in LD readers on comprehension tasks involving implied information (French et al., 1995; Wong, 1980).

The cognitive strategy deficits of RD children could bear direct relevance to the "inactive learner" hypothesis in that the observable performance of RD students often greatly underestimates the actual competence of these students (Torgesen, 1982b; Torgesen et al., 1979). More specifically, this discrepancy between observed and expected performance has been attributed to failure to employ cognitive strategies actively. However, there is also research evidence that suggest the performance deficits are due to children's inefficient use of cognitive strategies ("inefficient learner" hypothesis). This issue will be discussed in some depth in section 3.4.4.1.

Research suggests that metacognitive strategy deployment and reading achievement are positively related among both primary and secondary school students (Bean, Singer, & Sorter, 1986; Cross & Paris, 1988; Palincsar & Ransom, 1988), and tertiary students (Nist, Simpson, & Hogrebe, 1985). Competent readers have also been observed to characteristically monitor the comprehension and retention of material (French et al., 1995). In contrast, incompetent readers have been shown to be deficient in their use of metacognitive strategies (Campione, 1987; Kaufman, Randlett, & Price, 1985; Pressley & Levin, 1987).

Considerable evidence has accumulated to indicate that RD children experience difficulty with self-regulating mechanisms such as checking, monitoring, testing, revising, planning, and

evaluating when attempting to read (e.g., Bauer & Emhert, 1984; Brown & Palincsar, 1982; Palincsar & Brown, 1984; Pressley & Levin, 1987; Short & Ryan, 1984). Relative to their more successful peers, RD children have been found to exhibit metacognitive deficiencies at all levels of task performance. They are apparently less adept at:

1. evaluating their overall performance when finished with a task (French et al., 1995).
2. planning effective organizational schemes for approaching the task (French et al., 1995; Resnick & Klopfer, 1989), or organizing ideas (Englert & Thomas, 1987; C. C. Thomas, Englert, & Gregg, 1987).
3. monitoring their knowledge, progress or understanding of the task (Baker & Brown, 1984a; Bos & Filip, 1984; Wagoner, 1983; Winograd & Johnston, 1982).
4. regulating their comprehension and monitoring text for errors, confusions, or incongruencies (Bos & Filip, 1984; Davey, 1987; Garner, 1987).
5. adjusting their reading behaviour to match different kinds of content or reading contexts (Brown et al., 1986).

The studies reviewed clearly suggest that LD/RD students can be distinguished from NA students by their failure to spontaneously activate and employ comprehension strategies (Englert, Raphael, Fear, & Anderson, 1988). They are often unaware of appropriate cognitive strategies that facilitate task performance and cannot efficiently initiate, regulate, and monitor the use of such strategies. The self-regulation differences between these children still remain, when IQ is controlled (Short et al., 1991). However, it should be noted that the majority of evidence about the role of self-regulation in comprehension is correlational in nature (Mateos & Alonso, 1991).

Empirical research comparing the relative importance of metacognitive and cognitive strategies in predicting performance is sparse. It is argued that metacognitive and cognitive strategies must work in conjunction with each other if one is to be successful in reading. The cognitive strategies are essential tools to learn text, but the metacognitive strategies provide the monitoring and direction for their use. In other words, children can be taught to use many different cognitive strategies, but if they do not have the necessary metacognitive strategies to determine which cognitive strategy to use in a given situation, they will not be successful readers (cf. Pintrich & De Groot, 1990). Another point to note is that the usage of cognitive strategies is not often indicative of deliberate, planned behaviour. Children may learn that underlining or highlighting is a useful strategy for facilitating text comprehension, but if they

use it indiscriminately, highlighting irrelevant details or the entire text, it is unlikely that they are metacognitively aware.

In summary, the empirical research suggests that both metacognitive and cognitive strategy use are important predictors of reading performance for both RD and NA readers. However, the use of cognitive strategies without the concomitant use of metacognitive strategies may not necessarily improve reading performance. O'Neill (1992) suggested a need for additional research on the use of metacognitive strategies among readers of differing reading abilities, as well as clarification of the relationship between the use of metacognitive strategies and reading comprehension among readers of differing reading ability.

3.4.4 Relationship between metacognitive knowledge and cognitive/metacognitive strategy use

Several empirical studies have demonstrated a positive relationship between metacognitive knowledge and metacognitive strategy use in a variety of domains such as problem solving, mathematics, and reading comprehension (Baker & Brown, 1984a; Schraw, 1994; Slife et al., 1985; Slife & Weaver, 1992). Such a relationship has also been found with learning disabled children and poor readers (e.g., Slife et al., 1985). It has been suggested that while competence in self-regulation only occurs when an individual has high metacognitive knowledge, metacognitive knowledge does not guarantee a high degree of self-regulatory competence (Schraw, 1994).

Empirical research has also indicated that both metacognitive knowledge and cognitive strategy use are positively related (Kurtz, 1990; Pressley et al., 1987; Schneider & Pressley, 1989). In fact, metacognitive knowledge is an essential prerequisite if students are to become competent cognitive strategy users (Baker & Brown, 1984a; Borkowski et al., 1989; Harris & Pressley, 1991; Paris et al., 1983; Winograd & Hare, 1988). Knowledge about strategies has been shown to predict cognitive strategy use in NA children (Borkowski, Peck, Reid, & Kurtz, 1983; Kurtz, Reid, Borkowski, & Cavanaugh, 1982).

However, there is also evidence that knowledge of strategies does not ensure their effective and consistent use (Borkowski et al., 1990; Garner, 1990; Paris et al., 1991). This has been demonstrated particularly with unsuccessful readers. For instance, research has indicated that while RD children may be able to use cognitive strategies (e.g., contextual clues to assist in

comprehension) when directed to do so, they do not know when, where, and how to use these strategies spontaneously (e.g., Swanson, 1989). Thus, although metacognitive knowledge leads to strategy application, evidence shows that this knowledge may be insufficient, particularly for RD students who lack the necessary skills for precise and efficient use of the knowledge.

A variety of explanations have been offered as to why RD children are deficient in their use of cognitive strategies. Several researchers have proposed that RD students' failure to use cognitive strategies effectively, if at all, during reading is influenced by their limited metacognitive knowledge about reading (e.g., Paris & Oka, 1989). For example, research has revealed that poor and LD readers' failure to summarize well is due to their uncertainty about what information to include (Winograd, 1984).

In contrast, others have proposed that these children possess the metacognitive knowledge but fail to apply it spontaneously (Cooney & Swanson, 1987; Short & Weissberg-Benchell, 1989; Swanson, 1989; Torgesen, 1979). Such a view has been supported by the observation that external prompting could mobilize these children to apply the cognitive strategies (see Cooney & Swanson, 1987; Swanson, 1989, for a review), and by empirical research (Belmont, 1989; Ellis, Deshler, Lenz, Schumaker, & Clark, 1991; Short et al., 1991). This characteristic supports the conception of these children as inactive learners (see section 3.4.4.1).

On the other hand, Barclay and Hagen (1982) take a liberal view and allow for the possibility that LD and RD children have deficits in both the knowledge and regulation components of metacognition. There is now increasing support for the view that the poor task performance and the passiveness of RD students' learning are due in part to their poor metacognitive functioning. Such a view seems plausible since efficient learning is characterized by appropriate use of cognitive strategies, and sophisticated metacognition is necessary for efficient cognitive strategy use (Pressley et al., 1987). Metacognition is thought to direct efficient cognitive strategy use in two ways. First, knowledge about specific strategies, including how, when, and why to use them, is necessary in order to implement a strategy. Second, the regulatory aspect of metacognition allows children to monitor cognitive strategy effectiveness, and to modify strategies when faced with new or challenging tasks (Ghatala, Levin, Pressley, & Lodico, 1985). This regulation component of metacognition is helpful in augmenting children's existing metacognitive knowledge about strategies.

Another factor that has been implicated for RD students' deficient cognitive strategy use is motivation. Short and Ryan (1984) have suggested that the strategic differences between RD and NA readers may result from differential willingness or ability to utilize metacognitive knowledge and strategies. Such a view has been supported by research showing that RD children avoid strategy use that requires a high level of effort (Swanson, 1984; see also Guttentag, 1984). The importance of motivational factors in RD or LD children will be discussed in detail in the next chapter.

It should be noted that the first two viewpoints arise from the contradictory research findings concerning the metacognitive knowledge of RD children. While a number of studies indicate that RD children have limited metacognitive knowledge, there are others that suggest that RD children possess metacognitive knowledge. The discrepant results in studies of RD children's metacognitive knowledge may be due to the variety of definitions and instruments used across studies.

The above conflicting viewpoints highlight the following questions: What is the nature of the relationship between metacognitive knowledge and cognitive strategy use in RD and NA children? Do RD children possess less metacognitive knowledge than NA children? It is plausible to suggest that most RD children are not totally devoid of metacognitive knowledge considering that their intelligence falls in the normal range; it may be somewhat limited and less sophisticated than that of NA readers.

In summary, the correlational data reported provide evidence that metacognitive knowledge might be an important causal factor determining readers' deployment of metacognitive strategies. However, the relationship between reading metacognitive knowledge and cognitive strategy use, particularly in RD children, is unclear as inconsistent results have been found. Further investigation of the relationships between metacognitive processes and reading may help clarify the processes underlying children's reading difficulties.

A related theme emerging from the examination of the cognitive and metacognitive aspects of reading is that learning disabled children are inactive learners. The debate surrounding the issue of "inactive learner" or "competence" is now addressed. The controversy centres on whether LD students' poor performance results from a failure to utilize existing skills (i.e., inactive learner) or from an absence of skill (i.e., competence issue). That is, the inactive learner hypothesis suggests that metacognitive knowledge does not predict performance, while

the competence perspective suggest that it does. A related perspective that has recently emerged, is that of the "inefficient learner" view. These perspectives will be discussed next.

3.4.4.1 Inactive versus inefficient view of cognitive strategy deficits

Students with learning disabilities are often characterized as passive or strategically inactive learners (Torgesen, 1979, 1980, 1982b; Torgesen et al., 1979; Wong, 1980), with a production deficit in the use of relevant cognitive strategies (Flavell, 1970). The characterization of the LD individual as an inactive learner received wide acceptance for a number of decades (Swanson, 1989). Such a description has been based largely on research reporting deficient use of active strategies for storage and recall of information. A suggested revision of this view has been advanced by Swanson (1989), whose investigations have led him to argue "that a more accurate characterization of [learning disabled] children is that they are actively inefficient learners" (p. 10). Swanson (1988, 1989) found that LD children were able to "actively" develop strategic thought patterns, although inefficiently, and suggested that their inefficiencies may be related to preferred use of heuristics. From this perspective, LD children employ active strategic thought processes, but appear constrained in their ability to use more effective strategies with flexibility.

In a related vein, Bray (1985) argued for an examination of strategic flexibility or the point at which learners' strategies break down to gain a more accurate appraisal of strategy use. In other words, the emphasis is on when learners can or cannot use a strategy rather than on the strategy itself. Although LD learners have been assumed to be deficient in their deployment of cognitive strategies across a wide array of tasks, this is not synonymous with a sheer absence of strategic behaviour (i.e., inactive learning). This distinction was made clear by studies which employed the think aloud technique (e.g., Short et al., 1991).

In a review of the cognitive strategy literature, Short, Schatschneider, and Friebert (1992) concluded that while there was considerable support for the "inactive learner" hypothesis (e.g., Short et al., 1991), there was also evidence against the theory, in support of the "inefficient learner" hypothesis. To date, the validity of these competing views have not been tested. According to H. L. Swanson (personal communication, September 20, 1995) the inactive learner has a repertoire of strategies, but cannot use them unless prompted to use them. The inefficient learner also has a repertoire of strategies but chooses an inappropriate one to complete the task at hand.

The present study tested the validity of inactive and inefficient learner views. Both hypotheses would predict some knowledge of strategies given the fact that these children have normal intelligence. However, the inactive strategy learner would have a qualitatively different response from that of the inefficient strategy learner. The inactive strategy user would be able to identify the most appropriate cognitive strategy that could do the task efficiently, but would not use it unless prompted. In contrast, the inefficient strategy user would not be able to identify the most appropriate strategy for efficient performance. Therefore, theorists espousing the inactive learner view are more likely to hypothesize that metacognitive knowledge will not predict cognitive strategy use. Alternatively, theorists espousing the inefficient learner view are likely to hypothesize that metacognitive knowledge will predict cognitive strategy use (H. L. Swanson, personal communication, September 20, 1995). With regard to the relationship between cognitive strategy use and performance, both the inactive and inefficient learner views implicate deficiencies in the use of strategies for learning or reading. Thus, both hypotheses would predict that cognitive strategy use does not predict performance for LD or RD children.

In summary, although the majority of learning disability strategy studies have been conducted within the framework of an "inactive learner" hypothesis, there is growing evidence that LD children may be inefficient learners rather than strategy deficient. A number of important questions have emerged from this review: Does metacognitive knowledge predict cognitive strategy use for RD children? Does cognitive strategy use predict reading comprehension performance for RD children? Are these relationships similar for NA readers?

3.5 Summary and conclusion

Despite the increased popularity of research on metacognition and reading, some fundamental problems persist particularly in relation to the definition of metacognition, and the distinction between metacognition and cognition. It is important to note that the way a researcher defines metacognition and cognition influences the subsequent choices of measures and educational interventions. The different methodological approaches to defining metacognition and cognition were introduced. Given the limitations of singular measures of metacognition, a multimethod assessment approach for measuring metacognition was suggested.

This chapter also reviewed research on the relationships between metacognition, cognition, and academic performance. The application of the construct of metacognition to the study of achievement has provided valuable insights into differences between RD and NA children. The available evidence indicates that there is a moderately strong relation between metacognitive knowledge and the use of metacognitive and cognitive strategies; readers who are more aware are more likely to engage in strategic reading. However, given some of the inconsistent findings, the relationships between metacognitive knowledge, cognitive strategy use, and performance require clarification, particularly for RD children. Clearly, more research is required in order to clarify the relationships among these constructs.

Most of the relationships between metacognition and cognition have been investigated using correlational designs. R. E. Reynolds and Wade (1986) suggest that the link between metacognition and cognitive performance must remain open to question until researchers begin to use designs that will allow them to make *causal* instead of correlational conclusions.

In conclusion, although previous research has made valuable contributions to our understanding that metacognitive knowledge and cognitive/metacognitive strategies are related to reading performance, more research is needed to: (a) define the specific roles of the different aspects of metacognition as they relate to reading performance, (b) determine the relative effects of metacognitive and cognitive strategy use and their direction of causality to performance, and (c) test competing theories (inactive vs inefficient learner hypotheses). The present study aims to address all these questions.

While this account of persistent strategic deficits is plausible and possibly accurate, it cannot be regarded as fully explaining the performance deficits of reading disabled children. The role of "motivational" variables in reading performance has also been forwarded as another explanation. This is due to the evidence that cognitive and metacognitive skills are a necessary but insufficient condition for academic success. In the next chapter, the importance of motivational beliefs, namely attributions and self-efficacy, for effective performance in reading will be elaborated.

CHAPTER 4

MOTIVATIONAL FACTORS IN ACADEMIC PERFORMANCE

According to a number of researchers, the study of motivation is an investigation into the *why* of behaviour (Covington, 1992; Deci & Ryan, 1985). The study of achievement motivation, more specifically, examines the processes that energize and direct behaviour in achievement situations. A significant body of research has emerged specifically related to motivation in school settings (e.g., Ames & Ames, 1984; Fortier, Vallerand, & Guay, 1995; Meece, Wigfield, & Eccles, 1990). In general, such research reveals that academic motivation positively influences academic performance.

A large body of research has revealed that motivational characteristics of reading disabled (RD) children can also undermine their academic performance. RD children often develop motivational problems as a consequence of their reading difficulties, including low self-concept, self-efficacy, and self-esteem. In addition, they tend to attribute their success to external factors and their failure to diminished ability. These motivational problems and beliefs are eventually thought to potentially lead to a debilitating condition referred to as learned helplessness (Licht & Kistner, 1986). Although researchers agree that motivation is important in RD children's achievement, the precise role it plays in achievement is still being actively investigated (Fortier et al., 1995).

Researchers have examined a wide array of motivational belief variables to examine the link between motivation and performance in the domain of reading. These include attributions, expectations, self-esteem, outcome expectancy, intrinsic value, self-efficacy, and self-concept. Two variables which have attracted considerable attention are self-efficacy and causal attributions. Since these variables are important determinants of motivation, they will also be referred to as motivational variables. These motivational beliefs have been shown by research to be important and potent predictors of reading performance and are the focus of discussion in this chapter.

4.1 Self-efficacy

Perceived self-efficacy is defined as the self-perceptions of one's skills and competence to perform the behaviours necessary to attain designated performance levels (Bandura, 1986;

Schunk, 1989b). It is postulated to have diverse effects in achievement settings (Bandura, 1977, 1982; Schunk, 1984, 1989b). Self-efficacy affects behavioural functioning by influencing an individual's choice of activities, effort expenditure, and persistence in the face of difficulties (Bandura, 1986, 1988; Schunk, 1989b). This construct is seen as important in regulating student motivation on a wide variety of academic tasks (Schunk, 1984, 1989c). It has been argued that an individual's efficacy mediates the influence of determinants like gender, prior experience, anxiety, self-concept, and attributions on subsequent performance and is the stronger predictor of that performance when those determinants are controlled (Bandura, 1986; Pajares & Miller, 1994).

Research has indicated that successful performances do not necessarily enhance efficacy perceptions; the impact of this information depends on how it is cognitively appraised and interpreted (Bandura, 1986; Eccles et al., 1983; Schunk 1984). Thus, the appraisal of efficacy is an inferential process in which persons weight and combine contributions of personal and situational factors (Bandura, 1981). In forming efficacy judgments, a number of factors are taken into account. These include, among others, self-perceptions of ability, effort expenditure, task outcomes, and the pattern of successes and failures (Bandura, 1981; Schunk, 1984).

Although it has been established that self-efficacy is a strong predictor of behaviour (Bandura, 1986), research on the relationship between self-efficacy and academic performance is still limited (Bouffard-Bouchard, 1990). Empirical research has demonstrated moderate predictive correlations of self-efficacy with academic performance (Schunk, 1983, 1985). Correlational and path-analytic studies have provided initial support for the direct role of self-efficacy on academic achievement (Horn, Bruning, Schraw, Curry, & Katkanant, 1993; Schunk, 1983).

Considerably more research has been conducted on self-efficacy in the area of mathematics (e.g., Pajares & Miller, 1994; Randhawa, Beamer, & Lundberg, 1993; Schunk, 1984, 1989b; Seegers & Boekaerts, 1993) but relatively little in the domain of reading. The paucity of research in perceived self-efficacy for reading is surprising considering that the construct is intimately tied to reading achievement (Cohen, McDonell, & Osborn, 1989; Henk & Melnick, 1995). It is important to study efficacy within domains, since self-efficacy has been basically conceptualized as a situation-specific belief (Pajares & Miller, 1994), and students' perceptions of self-efficacy can vary across domains (Bandura, Adams, Hardy, & Howells, 1980).

Reading disabled readers have been found to have lower self-efficacy than their normal achieving counterparts (Cohen et al., 1989). However, there are a few studies that have suggested otherwise. For example, Pintrich et al. (1994) showed that RD children did not differ from normal achieving (NA) readers on self-efficacy. In addition, they did not show any signs of learned helplessness, although they did tend to attribute success and failure to external causes more often than their NA peers.

In summary, self-efficacy appears to be an important variable in understanding achievement behaviour. Although researchers have established that self-efficacy is a strong predictor of behaviour, research on the relationship between self-efficacy and academic performance, particularly in the domain of reading, is limited. Moreover, this research has largely been correlational in nature. Thus, there is a need to construct causal models in order to conceptualize and test hypothesized relationships for reading performance (cf. Hackett & Betz, 1989; Meece et al., 1990).

4.2 Causal attributions

While self-efficacy theorists posit that humans are motivated by perceptions of competency and personal control, attribution theorists posit that humans are motivated primarily to understand themselves and the world around them (Weiner, 1979, 1983, 1985a). Perceptions of the causes of achievement outcomes are referred to as causal attributions. Students' perceptions about the causes of their academic success and failure are quite diverse, with the impact of ability, effort, task difficulty, and luck being investigated most often.

Attributional style refers to the predisposition to make certain types of attributions in response to success and failure (Peterson et al., 1982). This construct can be traced to the reformulated learned helplessness model of Abramson, Seligman, and Teasdale (1978) and the locus of control theory of Rotter (1966). The basic premise of this model is that differences in attributional style are related to differences in motivation, performance, and affective reactions (C. Anderson, Jennings, & Arnoult, 1988). Abramson et al. (1978) suggested that some individuals develop a maladaptive attributional style. This becomes manifested as a tendency to attribute negative outcomes to internal, stable, and global causes (e.g., lack of ability) and positive outcomes to external, unstable, and specific causes (e.g., a lucky break). An adaptive attributional style can be achieved from viewing academic successes as personally caused, likely to recur, and under one's control.

Weiner's (1979, 1985a) model of achievement motivation has been widely employed to investigate how causal attributions for success or failure influence achievement-related behaviour. Weiner proposed a logical model of attributional thinking in which the causes that students use to explain academic success and failure are subsumed into dimensions. These dimensions were the locus of cause (internal vs external), the stability of cause (stable/unchanging vs unstable/variable), and the controllability of cause (controllable vs uncontrollable). It should be noted that there are two other attribution dimensions. These include globality versus specificity, and intentionality (Weiner, 1985b). To a large extent the relevant attribution dimensions depend upon the domain of interest. In the present study, discussion is limited to the first three dimensions as they are most relevant in the educational literature. Attribution theorists have emphasized that it is the underlying *dimension* or cognitive structure rather than the *specific causal attribution* (e.g., ability, effort) that is the key to the motivating properties of attributions.

While early attributional models in educational research have been based on attributions of ability, effort, task difficulty, and luck, it has now been confirmed that a more diverse range of possible causes are used by children in judging reading achievement outcomes (Hiebert, Winograd, & Danner, 1984). It has been suggested that a focus on such an ever-expanding list of attributional elements may cause difficulty in building a parsimonious attributional theory (Benson, 1989). However, despite the large number of potential causal attributions, there appear to be certain consistent dimensions which characterize them. Consequently, interest in causal attributions for achievement outcomes has recently gone beyond the perceived causes themselves to the underlying causal dimensions. Such an approach may be useful in enhancing the parsimony and generalizability of attributional theory (Benson, 1989). Considering the greater importance of underlying dimensions of attributional style relative to specific causal attributions as determinants of achievement behaviour (Weiner, 1979, 1985b, 1986), the present research focused on measurement of attributional dimensions.

The underlying causal dimensions are postulated to have differential effects on affect, expectations, perceptions of self or others, and behaviour (Weiner, 1979). Several one-to-one associations between attributional dimensions and the aforementioned performance mediators have been postulated by Weiner (1979). For instance, attributions of locus (internality) were associated with self-esteem, attributions of stability with expectations, and attributions of controllability with motivation and self-esteem. However, subsequent empirical research have failed to support these differential predictions clearly (Benson, 1989).

Weiner (1983) asserted that while the postulated relation between attributions and performance is indirect, the relationship between attributions and performance mediators (e.g., self-esteem, affect, expectations) is direct. Attributing failure to stable-uncontrollable causes (such as ability) reduces persistence as it does not leave the possibility of changing the outcome in the future. Attributing failure to a stable internal cause is likely to produce negative attitudes towards academic tasks (C. G. Rogers, 1987). Considerable evidence has accumulated to indicate that students' attributions for success and failure influences persistence, effort expenditure and achievement task choices (Bar-Tal, 1991).

The influence of attributions on persistence and achievement task choices in turn are manifested as mastery-oriented or learned helplessness behaviour. As Bar-Tal (1982) summarizes, the "attributional model indicates that pupils who tend to attribute success to internal, mainly stable, or controllable causes, and who attribute failure to internal-unstable-controllable causes tend to exhibit adaptive, mastery-oriented achievement behaviour" (p. 190). In contrast, attributing success to external-uncontrollable-unstable causes, and failure to uncontrollable-stable causes tends to lead to maladaptive helpless achievement behaviour. The causal attributions of learning disabled children have been suggested to reflect an attitude of learned helplessness (A. Thomas, 1979).

There is considerable support for the central assertion of Weiner's (1979, 1985b, 1986) theory, that a relationship exists between students' attributions for academic success or failure and achievement (Schunk, 1983). An obvious outgrowth of this reported relation has been the investigation of the effects of causal attributions on achievement motivation and performance in relation to failing students (Ames & Ames, 1984; Covington & Omelich, 1981). It has been proposed that repeated failures cause LD children to view themselves as having no control over achievement outcomes, and to view their efforts as fruitless (Licht, 1983). These beliefs are postulated to result in reduced persistence in mastering schoolwork, which in turn increases the likelihood of continued failures and reinforces the children's perceptions of lack of control. It is expected that over time these beliefs impede academic progress, and, for this reason, interventions to alter LD children's achievement-related beliefs have been recommended (Borkowski, Weyhing, & Carr, 1988; Carr et al., 1991; Licht, 1983).

Numerous studies have shown that LD or RD children are less likely than their NA peers to view their efforts as determinants of achievement outcomes (Butkowsky & Willows, 1980; Licht, Kistner, Ozkaragoz, Shapiro, & Clausen, 1985; Pearl, 1982; Pearl, Bryan, & Donahue, 1980). Learning disabled children are less likely to see their success as reflecting on their

abilities (Butkowsky & Willows, 1980; Pearl, 1982; Pearl, Bryan, & Herzog, 1983), attributing their success, instead, to ease of the task or to luck (Pearl, 1982; A. St. George et al., 1985). Furthermore, LD children are less likely to view their difficulties as changeable through their own efforts (Butkowsky & Willows, 1980; Cooley & Ayres, 1988; Licht et al., 1985; Pearl, 1982; Pearl et al., 1980) and may sometimes generalize these beliefs to tasks on which they have not experienced failure (Butkowsky & Willows, 1980). Such perceptions of powerlessness discourages expectations for future success as well as *responsibility* for one's actions (A. Thomas, 1979).

The research reviewed thus far strongly suggests that LD children are at risk for developing motivational problems that interfere with their work. However, there are individual differences among LD children in terms of how they interpret and respond to their repeated failures (Kistner, Osborne, & LeVerrier, 1988). For instance, some LD children have maintained their confidence in their abilities by attributing their academic difficulties to insufficient effort (Licht et al., 1985) or other external uncontrollable causes, such as an unfair teacher, bad luck, or task difficulty (e.g., Luchow, Crowl, & Kahn, 1985; van Kraayenoord, 1986). Others maintain confidence by not giving up in the face of difficulty (Licht et al., 1985; Speece, McKinney, & Applebaum, 1985).

In general, differences between mastery oriented and learned helpless children have been consistent with respect to the controllability and stability dimensions for success and failure. However, inconsistent results have been found for the dimension of internality-externality for failure attributions of NA and LD children. This may be resolved by distinguishing between internal causes varying in stability and controllability dimensions (e.g., ability and effort). Research has shown that LD as well as NA children make both external and internal attributions for failure (Pearl et al., 1980). Thus, it remains unclear whether an external response to failure is characteristic of LD students, because some studies have indicated that mastery-oriented children were more likely to place blame outside themselves. While the findings for the attributions of LD and NA children are somewhat equivocal, methodological differences such as the sample characteristics (e.g., heterogeneity) and attributional measures are likely to be partly responsible (Licht et al., 1985).

In conclusion, most of the attributional research has been based on specific attributional causes rather than dimensions. Research on the attributional patterns of RD children is limited. Investigation of the attributional style of NA and LD children have yielded inconsistent results, particularly for the dimension of internality-externality. Further

investigation of the attributional patterns of RD children would increase the data base upon which important intervention decisions might be made.

4.3 Relationship between self-efficacy and attributions

The relation between Bandura's (1982) self-efficacy theory and Weiner's (1985a) attribution theory has been explored by Schunk (1983, 1984). Both theoretical paradigms are similar in that they highlight the role of cognitions in (a) appraising environmental information, (b) deriving inferences about factors influencing performance, and (c) engendering expectancies that influence achievement behaviours. However, they differ in the types of judgmental factors that are held to directly influence expectancies (Schunk, 1984).

Despite much theorizing about the relationships between efficacy and attributions, and their relation to performance, no research to date has investigated the joint and separate causal effects of efficacy and attributions on reading comprehension performance. According to most theoretical models, attributions are directly related to expectancies and these expectancies are directly related to achievement (e.g., Bandura, 1986; Weiner, 1985b). Some theorists specify this relationship even more by suggesting the effects of attributions on achievement are mediated by self-efficacy expectancies (Bandura, 1986; Schunk, 1983). Within self-efficacy theory, causal attributions are considered important antecedents of efficacy expectations, and are important sources of efficacy information.

There are several questions that arise from this motivational literature. One question that is of particular importance relates to how self-efficacy and attribution variables vary in their predictive influence on reading performance. Another important question that arises is whether or not the causal patterns of attributions, efficacy, and reading performance differ for reading disabled children, compared to normal achieving children.

4.4 Methodological problems in measuring attributions and self-efficacy

4.4.1 Attributions

Despite the evidence that causal attributions are important determinants of motivation and achievement, relatively little attention has been paid to the reliability and validity of different

methods for assessing attributions. Most methods of assessing attributions involve collecting data on subject's causal explanations for a given outcome. However, in testing predictions from attribution models, it is often necessary to go beyond the specific causal explanation for a given outcome to the underlying causal dimension. In the traditional attribution paradigm, researchers translate the subject's attributions (of ability, effort, task difficulty, etc.) into causal dimensions such as internal-external, stable-unstable, controllable-uncontrollable. However, there is some risk of an invalid assessment because the dimensions are based on the assumption that the researcher's perspective on the underlying meaning of the causes will match the subject's perspective. This is known as the "fundamental attribution researcher error" (Russell, 1982), and evidence of this has been demonstrated in several studies. For example, researchers have traditionally classified ability as a stable, uncontrollable cause, yet research indicates that children may view ability as unstable and controllable in that ability can be developed by learning new skills (Hiebert et al., 1984). Similar findings have been described for task difficulty (Meyer, 1980) and effort (Weiner, 1979).

To overcome the problem of "fundamental attribution researcher error", researchers have suggested that the underlying dimensions be tapped directly (Ronis, Hansen, & O'Leary, 1983; Russell, 1982; Weiner, 1983; Winograd, Witte, & Smith, 1986). A number of studies indicate that direct measures of causal dimensions provide a more accurate understanding of the subjects' attributions, and are better predictors of future performance than indirect or derived measures of dimensions (Ronis et al., 1983; Winograd et al., 1986). With regard to the measurement of attributions for reading, Winograd et al. (1986) concluded that it may be unwise to assume that readers' views of the underlying dimensions can be inferred from the original causal taxonomy, particularly with regard to poorer or younger readers.

Another area of concern with regard to the measurement of attributional style is related to its multidimensional nature. In past research, a number of researchers have used combined scores to assess an individual's attributional style. That is, scores for the distinct dimensions were combined to form a single predictive index. The primary reason for this was the low internal consistency of the separate dimensions (.40 to .70) reported on commonly used measures of attributional style (Peterson et al., 1982). Recently, researchers using structural equation modelling have modelled the multifaceted construct of attributional style for both success and failure outcomes as a single latent variable or theoretical construct (Berndt & Miller, 1990; Carr et al., 1991). This may have been done to simplify data analysis and the conceptual explanation of the findings.

However, Carver (1989) argued that using a composite score when assessing attributional style may result in a loss of information. For example, an obtained group difference could be attributable to one component of the predictor index, to the additive effects of two or more components, or to the interactive effects of all components. Consequently, some researchers have proposed measuring attributional style for positive and negative events separately (Abramson, Metalsky, & Alloy, 1989; Sweeney, Anderson, & Bailey, 1986). Such an approach would enable one to tease out the differential effects of the subcomponents rather than obscuring them, and was therefore adopted in the present study.

Support for an argument that some subcomponents may be more strongly related to behaviour than others can be seen in the depression literature. The findings of a meta-analysis conducted by Sweeney et al. (1986) revealed that attributional style for negative events was more strongly related to depression than attributional style for positive events. An important question that arises from this finding for the achievement literature is: Do subcomponents of attributions, particularly for different outcomes (success and failure), differ in their effect on reading achievement? This question remains uninvestigated to date and addressing it represents one of the goals of the present study.

In the present study, two different measurement or factor models for attributional style were tested (see Chapter 7, section 7.1.1.1). The first measurement model corresponds to the recent theoretical formulation that the components of attributional style are related to a single latent construct (Berndt & Miller, 1990; Carr et al., 1991). The second measurement model, based on the assumption that attributional style for positive and negative events should be considered separately (Abramson et al., 1989; Sweeney et al., 1986), was also tested to investigate the differential effects of attributional style for success and failure outcomes on achievement.

Finally, another area of concern in the measurement of attributions is that researchers infrequently use domain specific measures in examining attributions for reading (Borkowski, Weyhing, et al., 1988; Carr et al., 1991). Eccles (1987) admonished researchers who did not use domain-specific measures of student motivation. Eccles suggested that "the predictive power of particular motivational constructs increases as one makes both the domain of the motivational construct and the achievement outcome being predicted more specific" (p. 2). In light of this, the present study used domain specific measures for all constructs of interest in the study.

4.4.2 Self-efficacy

A review of the literature on reading revealed that there are few instruments developed to measure reading efficacy, with valid and reliable measures being sparse (Henk & Melnick, 1992). In fact, many current self-perception theorists often employ measures of self-concept instead of efficacy to predict performances across situations (Chapman & Tunmer, 1995). These measurement inequities may account for some of the inconsistent results concerning the relationship between motivational beliefs and reading performance in the literature. Bandura (1981) has argued that self-concept of ability represents global estimates of efficacy. It is clear that beliefs regarding confidence are part of an individual's self-concept, but self-efficacy and self-concept represent different phenomena (Bandura, 1986). This distinction is also consistent with expectancy value models of achievement (Eccles et al., 1983).

In addition, the predictive power of self-concept may be less than that of self-efficacy, since self-concept is not measured at the same level of specificity as efficacy (Pajares & Miller, 1994). Self-efficacy involves a context-specific assessment of competence to perform a specific task, while self-concept includes beliefs of self-worth associated with one's perceived competence.

The literature on reading self-efficacy reveals the use of four different methodologies for obtaining reading self-efficacy (see Pereira-Laird & Deane, 1995, Appendix C). However, most existing measures of efficacy measure achievement globally (Horn et al., 1993; Pintrich et al., 1994; Pintrich & De Groot, 1990). Bandura (1986) cautioned that because judgments of self-efficacy are task specific, different ways of assessing confidence will correspond differently to the assessed performance. Thus, he argued that self-efficacy should be specifically rather than globally assessed, and should match the performance task. He claimed that these guidelines were seldom followed in educational research and often led to the mismeasurement of self-efficacy. In light of this, the present study incorporated task and subject specific measures of efficacy.

4.5 Summary

In summary, LD children are at risk for developing motivational problems. As a consequence of their repeated failure, they are likely to possess maladaptive attributional patterns and feel less efficacious. However, not all LD children respond in the same way to failures. Although sometimes confounded by imprecise definitions and varying measurements, findings show that

both attributional style and self-efficacy are related to performance. In fact, theory suggests that self-efficacy mediates the influence of attributions on performance. To date, this relationship has not been investigated empirically in the domain of reading.

Methodological issues regarding attributions and efficacy were discussed and it was concluded that domain specific measures are more predictive of achievement. A number of future research directions were identified. These included: (a) how the subcomponents of attributional style (success and failure) would vary in their effect on achievement, and (b) how self-efficacy and attributions may combine to influence performance.

CHAPTER 5

THE RELATIONSHIPS BETWEEN METACOGNITIVE, COGNITIVE, AND MOTIVATIONAL VARIABLES

Cognitive development in children is beginning to be conceptualized as a product of a self-regulatory system where the focus is on the individual's ability to initiate and direct efforts to acquire knowledge and skill (e.g., McCombs, 1986). Empirical research in this area has stressed the importance of children's selection and use of learning strategies that are well suited to particular learning goals and tasks (Ames & Archer, 1988; Borkowski et al., 1990; Paris & Winograd, 1990b; Zimmerman & Schunk, 1989). In a similar vein, contemporary models of reading specify comprehension as the product of a number of factors, including strategies, metacognitive knowledge, and motivation. These collective measures of cognitive, metacognitive, and motivational variables provide a more comprehensive view of reading achievement and impairment because they combine assessments of cognitive ability, metacognitive skill, and motivational will. Thus, they represent a vast improvement over searches for single factors that distinguish between normal achieving and reading disabled children.

This chapter presents theoretical models and research which have shifted away from a strictly unicomponential approach to cognitive development, and towards an approach which incorporates multiple components. Two theoretical models that propose a multicomponential view of academic achievement are the Triple Alliance model (Short & Weissberg-Benchell, 1989) and Borkowski et al.'s (1989, 1990, 1992) model of metacognition. According to these models a functional cognitive system containing adaptive attributions, high self-efficacy, metacognitive knowledge, and strategy use strengthens performance by promoting children's effort towards the goal of self-determination (McCombs, 1986).

5.1 Triple Alliance Model

Short and Weissberg-Benchell (1989), in their Triple Alliance Model for learning, suggested that the discrepancy between the ability and achievement of learning disabled children can be explained by an analysis of cognitive, metacognitive, and motivational variables. According to this model, a delicate balance must be achieved between cognitive, metacognitive, and

motivational processes for effective learning to occur (Short et al., 1992). This means that metacognitive knowledge is usually insufficient to promote student achievement; students also must be motivated to use the cognitive strategies as well as regulate their cognition and effort (Pintrich, 1988, 1989). Short et al. (1992) suggest that achieving students form "effective alliances" between cognitive, metacognitive, and motivational processes so that metacognitive and motivational processes increase students' cognitive skills. In comparison, unskilled learners (such as learning or reading disabled children) form "faulty alliances" between cognitive, metacognitive, and motivational processes. Hence, the Triple Alliance Model offers a number of explanations for the poor performance of reading disabled children. The model proposes that inactivity in learning may occur because of cognitive deficiencies; because of metacognitive deficiencies which may affect the learner's ability to demonstrate cognitive competence; or because of a defective motivational style (Short et al., 1992).

One limitation of this model is that it does not delineate how metacognitive, cognitive, and motivational factors causally relate to each other. Borkowski et al.'s (1989, 1990, 1992) model does this and thereby provides the conceptual framework of the present study.

5.2 Borkowski's model of metacognition

Recently, Borkowski and his colleagues offered theoretical support for the utility of metacognitive factors in understanding skill-level differences in reading for normal achieving (NA) and reading disabled (RD) children (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Carr et al., 1991; Groteluschen et al., 1990). This general model, in combining several metacognitive, cognitive, and motivational components, may help explain how and why reading difficulties develop in some children. Specifically, the metacognitive model contends that achievement is based on the interaction of metacognitive knowledge (specific strategy knowledge, general strategy knowledge, relational knowledge), higher order executive processes (or metacognitive strategy use), and motivational states related to general strategy knowledge.

There are several descriptions of the model of metacognition proposed by Borkowski and his colleagues (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). Although the descriptions have evolved somewhat over time, they are largely consistent with each other. The apparent variability in descriptions arise because different

parts of the model were emphasized depending on the point the authors wanted to make (M. Carr, personal communication, February 13, 1995). Although the original model was built primarily to substantiate metamemory theory, the components of this model can be applied to a wider range of cognitive activities such as reading comprehension, mathematics, and science.

Borkowski et al.'s (1989, 1990, 1992) model represents an improvement over earlier conceptualizations of metacognition in that it theoretically links components of metacognition with strategies (Schneider & Weinert, 1990). One unique feature of their model is that "it accounts for different declarative knowledge components (i.e., specific strategy knowledge, general strategy knowledge, relational strategy knowledge) which are conceived of as interactive and mutually dependent causes of strategy use and cognitive performance" (Schneider & Weinert, 1990, p. 293). Another unique facet of this model is the motivational nature of one aspect of metacognitive knowledge, that is, general strategy knowledge.

The following is a condensed description of Borkowski and his associates' model (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). However, before proceeding to that description, some explanatory notes are necessary. Specifically, the original term "metamemory acquisition procedures" found in earlier versions (Pressley, Borkowski, & O'Sullivan, 1985) is equivalent to "executive processes" found in later versions (see Borkowski et al., 1989; Groteluschen et al., 1990). In addition, in earlier versions of the models (Borkowski, Johnston, & Reid, 1987), attributions were subsumed under general strategy knowledge, whereas in their more recent papers, attributions are conceptualized as correlates of general strategy knowledge (Groteluschen et al., 1990).

5.2.1 Components in the metacognition model

According to Borkowski et al.'s (1989, 1990, 1992) model, metacognition is comprised of a number of interactive and mutually dependent components (see Figure 2). The four major components proposed in these models include: specific strategy knowledge, relational strategy knowledge, general strategy knowledge, and executive processes or metacognitive acquisition procedures (MAPs). According to Borkowski and Turner (1990), these component processes have independent and unique developmental histories and serve distinctive roles in explaining individual differences in performance among NA and RD children. It should be emphasized

that metacognitive and motivational processes, individually or in combination, do not develop or change rapidly (Kurtz & Borkowski, 1984).

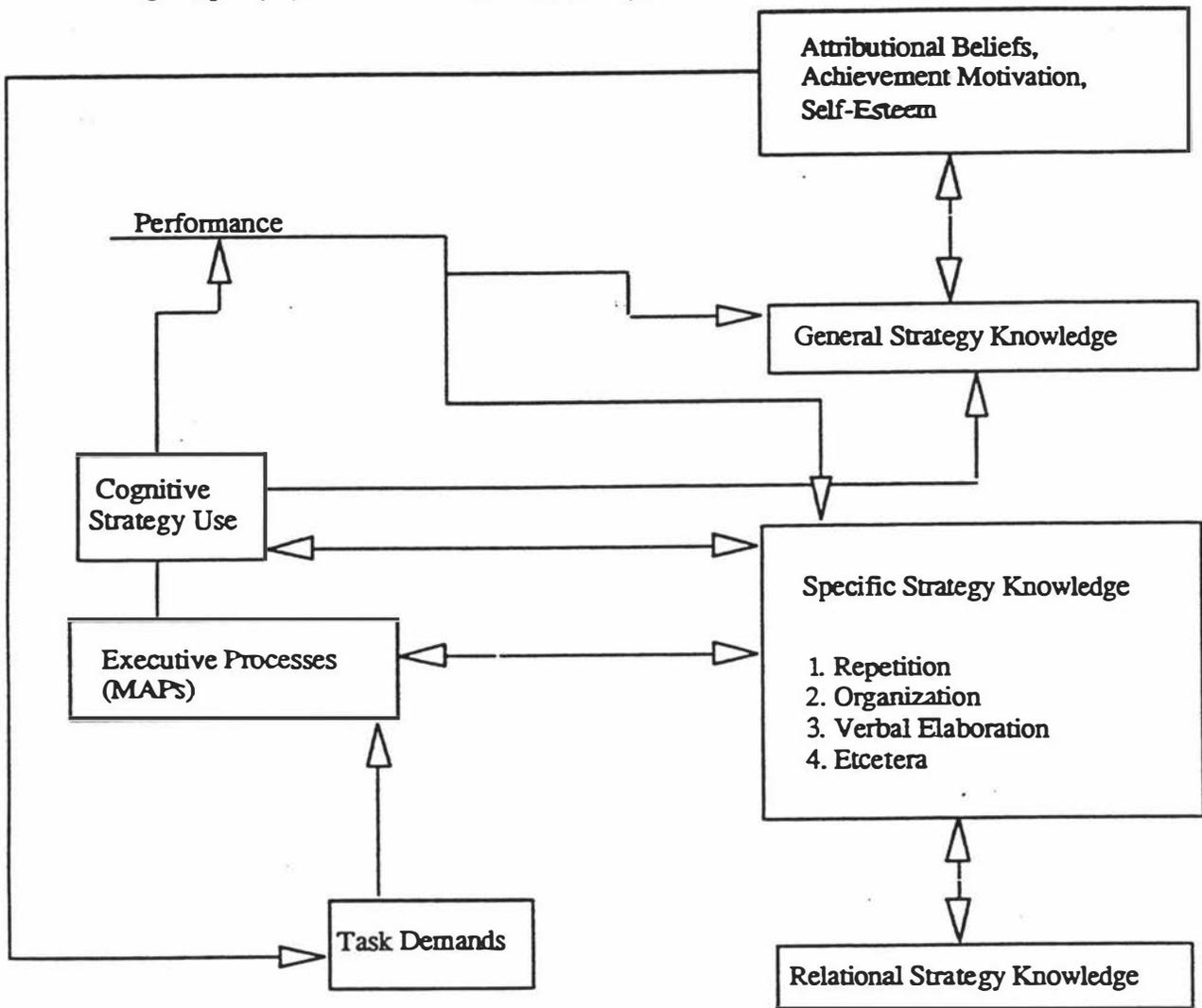


Figure 2. Borkowski's model of metacognition (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990).

Specific strategy knowledge refers to knowing when, why, where, and how to use strategies. It is the equivalent to metacognitive knowledge or awareness, as discussed in Chapter 3. This aspect of metacognition is essential for strategic behaviour and is assumed to be central to generalization effects (Groteluschen et al., 1990). It has been argued that poor strategic behaviour reported in LD students is due in part to inadequate specific strategy knowledge (Borkowski & Kurtz, 1984).

Relational strategy knowledge reflects one's knowledge of the strengths and weaknesses of different kinds of strategies. This knowledge is useful in strategy selection and strategy

revision. Since little is known about relational strategy knowledge in NA children, it will not receive much attention in this thesis.

General strategy knowledge refers to general attitudes and beliefs about the importance of effort in applying strategies, and that effortful strategic approaches can be useful to improve performance. Students with high general strategy knowledge generally believe that effort is useful only when it is channelled through the use of effective strategies, and that effort increases the likelihood of success. Attributions, particularly about effort, are closely related to general strategy knowledge in that a belief in effort gets translated into a belief in being strategic. General strategy knowledge about the efficacy of strategies includes motivational properties (e.g., high self-efficacy, intrinsic motivation, effort-related attributional beliefs for success and failure). Borkowski and his colleagues believe that general strategy knowledge and its associated motivational factors have a reciprocal relationship, each contributing to the development of the other (Borkowski & Turner, 1990; Groteluschen et al., 1990). They specifically highlight attributions as the motivational correlate of general strategy knowledge

Finally, metacognitive acquisition procedures (MAPs) represent higher order metacognitive processes and have also been referred to as self-regulation, executive functioning, or metacognitive strategy use. They account for the mechanisms by which new strategies are acquired and by which existing strategies are selected, implemented, and modified. These procedures help to account for the tailoring of strategies to fit variations in tasks and the self-regulation that is required while cognitive strategies are being used and modified.

5.2.2 Development of metacognitive components in a child

Borkowski and his colleagues (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990) have described how the metacognitive components would develop in a child. This account is paraphrased as follows:

1. The child who is taught to use a strategy will come to learn about the attributes of the strategy with repetition. Attributes include knowledge about the strategy's effectiveness and its range of appropriate applications.
2. Provided the learning environment is stimulating, other strategies will be acquired and repeated, leading to the development and enrichment of specific strategy knowledge.

3. Relational strategy knowledge becomes organized when the teacher shows or the child realizes the similarity and differences of multiple strategies (in a domain), resulting in the organization of skills based on their similarities.
4. The child gradually acquires the capacity to choose and monitor strategies appropriate for some tasks (but not others). When essential strategy components have not been sufficiently taught or learned, the child develops the ability to "fill in gaps" in instructions so that he or she becomes able to fit a strategy to the unique requirements of the task. Executive processes emerge at this stage, and become useful in making decisions about how and when to use a strategy, or to switch to a new strategy when an earlier one proves ineffective.
5. As the use of cognitive and executive strategic processes become refined, the child begins to recognize and believe in the value of being strategic (general strategy knowledge). In addition, the child learns to attribute successful (and unsuccessful) outcomes to effortful strategic behaviour (or lack of it) rather than to luck. At this stage, beliefs of self-efficacy also develop. Self-efficacy beliefs, derived from the use of cognitive strategies, eventually return to energize executive processes. In this way the metacognition model integrates cognitive strategy use with motivational causes and consequences.

The components in the model develop at different times in a child's history and are related to the quality of the child's educational experiences. In addition, each component relies on the quality and form of earlier componential development. For instance, general strategy knowledge emerges only after (a) a proportion of specific strategies have been encountered by the child, and (b) ample understanding about the importance and role of strategies occurs. Executive processes become operative only after detailed presentations in specific, general, and relational knowledge components are accumulated.

5.2.3 Relationships between components of the model

The motivational correlates of general strategy knowledge are components of the self-system. Components of the self-system includes attributional beliefs, self-concept, self-efficacy, locus of control, and achievement motivation. These variables or components provide the critical motivational mechanisms needed to stimulate strategic learning and effective performance (McCombs, 1986). The motivational influence of the self-system determines the seeking of

opportunities to read, and the degree of effort expenditure and persistence in understanding text (Henk & Melnick, 1992). Borkowski et al. (1990, 1992) refers to self-system variables as energisers of the metacognitive system and suggest that they help determine the quality of academic performance.

Recent research and theory suggest that attributional beliefs are the principal catalyst in terms of impacting self-efficacy, strategic learning, and performance (Borkowski et al., 1990; Groteluschen et al., 1990). According to Carr and Borkowski (1989), the development of cognitive strategic behaviour and effort-related attributions is a reciprocal cycle. Attributional beliefs are both the consequences of repeated cognitive strategy use and the causes of strategy selection and monitoring behaviours. In this sense, "attributions arise from lower level skills but fuel higher level executive processes" (Borkowski et al., 1989, p. 66).

Experiences of cognitive strategy use in multiple contexts result in beliefs in the value of strategic behaviour and in the role of effort in producing success. These forms of general strategy knowledge engender increased self-efficacy expectations which in turn may motivate children to confront challenging tasks (Dweck, 1987). According to Groteluschen et al. (1990), positive feelings of self-efficacy generate a greater willingness to experiment with strategies. Such personal experimentation is imperative for the development of executive processes involved in selecting, monitoring, modifying, and generalizing specific cognitive strategies. Thus, the advancement of executive processes requires much effort, and strategic effort is instituted and maintained by beliefs in its utility (Borkowski, Milstead, & Hale, 1988).

Specific strategy knowledge develops spontaneously as a function of the executive processes. There is also a dynamic reciprocal relationship between cognitive strategy use and specific strategy knowledge. Not only does specific strategy knowledge guide the deployment of cognitive strategies, but continued use of these strategies results in an expansion and refinement of specific strategy knowledge.

It is important to note that specific strategy knowledge is insufficient for cognitive strategic behaviour. One must also have confidence in the usefulness of a strategy before expending time and energy in its deliberate application. Furthermore, successful cognitive strategic behaviour requires higher-order executive processes that govern the use of this knowledge. These executive processes inform decisions about how and when to use cognitive strategies.

In summary, it can be seen that not only are specific cognitive strategies necessary for effective learning and reading, they furnish the context for developing executive processes. In addition, they represent the basis for reformulating attributional beliefs and enhancing self-efficacy. The model presented here describes how the interaction among metacognitive components generates a reciprocal chain of events among attributional beliefs, executive processes, and cognitive strategy use. This reciprocal chain determines whether a child learns to engage in strategy use with flexibility and persistence. Effective use of strategies enhances self-efficacy, adaptive attributional beliefs, and the expansion of metacognitive awareness. The self-system variables, in turn, determine readiness to use strategies.

5.2.4 Implications of the metacognition model for RD children

Borkowski et al. (1989, 1990, 1992) suggested that three components may be critical for understanding the differences in performance between RD and NA children. They are attributions (a motivational correlate of general strategy knowledge), specific strategy knowledge, and executive processes (or metacognitive strategy use). More specifically, Borkowski et al. (1989, 1990, 1992) suggested that these components were crucial in explaining the generalization failures of RD children.

In a recent study, Borkowski et al. (1990) concluded that attributional beliefs represented a significant factor in distinguishing underachieving and achieving readers. This is because they play an important role in the development of metacognition and in determining how the child will respond to academic experiences (Borkowski et al., 1990). For instance, effort-related attributions are likely to motivate children to acquire and use new cognitive strategies, resulting in the enhancement of metacognitive knowledge. The development of attributional beliefs is also closely tied to other self-system constructs such as self-efficacy (Eccles et al., 1983; Schunk, 1983), and intrinsic motivation (Watkins, 1984). Self-efficacy and other self-system constructs in turn predict achievement (Oka & Paris, 1987).

Normal achieving children tend to attribute their difficulties to controllable factors and their success to high ability and strategy use. These adaptive attributional beliefs help them maintain effortful, strategic behaviour even in the face of challenging tasks. Thus, the functional attributional patterns of NA children promote the development of motivational, metacognitive, and cognitive processes. In contrast, RD children who are not diagnosed until

after a period of school failure, often develop a negative self-system (e.g., maladaptive attributional beliefs). They are likely to feel less efficacious, and over the long term, fail to develop self-efficacy and adaptive attributions because of recurrent failures in reading. This in turn hinders the development of executive skills which are necessary for generalized problem solving. In addition, their external attributional orientation inhibits the acquisition of strategic behaviour and metacognitive knowledge because they have little reason to learn or to use strategies that they do not feel will help them achieve. Thus, RD children are in jeopardy of developing dysfunctional self- and metacognitive systems.

Metacognitive knowledge and executive processes also distinguish RD and NA children. Skilled readers possess sophisticated metacognitive knowledge about the purpose of reading and reading strategies (Baker & Brown, 1984a). NA children engage in achievement oriented problem solving behaviours because they believe that their difficulties are surmountable through their effortful strategic behaviour. These beliefs foster the development of specific strategy knowledge and executive processes. In contrast, RD children have specific strategy knowledge deficits. These deficits are likely to hinder the development of generalizable strategies. Their inappropriate attributional beliefs make it unlikely for them to take full advantage of new information or to make the best of learning opportunities as they tend to avoid learning challenges (Dweck, 1987). In this way, dysfunctional attributions hinder the development of specific strategy knowledge, and also the emergence of executive processes (Borkowski, Carr, & Pressley, 1987).

Groteluschen et al. (1990) suggest that executive processes are a major factor in accounting for the differences in ability between RD and NA children. Normal achieving children who apparently have mature executive processes possess the capacity to select appropriate strategies for the task at hand. In contrast, RD children have an overall deficit in executive processing as manifested by their ineptitude in selecting, implementing, monitoring, and generalizing available strategies. These executive processes are crucial in the deployment of cognitive strategies. This is because no matter how extensive a child's metacognitive knowledge is, successful deployment of cognitive strategies demands the use of executive processes that govern the use of metacognitive knowledge.

It is argued in the present study that another important factor distinguishing NA and RD children is cognitive strategy use. Deficits in cognitive strategy use have been implicated as reasons for the impoverished performance of RD children. When confronted with difficult

tasks, RD children do not engage in effortful deployment of strategies, often using ineffective or inefficient strategies. This results in a level of performance below their capabilities. Unlike proficient readers who modify or switch strategies in the face of failure, RD children often fail to effectively transform strategies. Strategic behaviour in children creates a sense of control which is reflected in effort-related or strategy based attributional beliefs which, in turn, foster further strategic behaviour.

The model predicts that deficits in metacognitive knowledge, self-efficacy, executive processes, and cognitive strategy use of RD children arise from their maladaptive attributional beliefs (Borkowski et al., 1990, 1992; Groteluschen et al., 1990). Thus, attributional beliefs may be the most important factor in differentiating NA and RD children. It has been suggested that, "although dysfunctional attributional patterns emerge as by-products of poor performance associated with biologically based learning deficits, their subsequent development is hindered by the maladaptive functioning of the entire metacognitive system in which they are imbedded" (Carr et al., 1991, p. 116).

In summary, it seems that a functional/dysfunctional system of cognitive, metacognitive, and motivational processes explains the maintenance of reading achievement/impairment. The dysfunctional attributional patterns of RD students, from the perspective of the model developed here, may hinder the development of important self- and metacognitive components. By espousing dysfunctional attributional belief patterns RD children handicap their development of positive self-efficacy, specific strategy knowledge, cognitive strategy use, and executive processes. Thus, adaptive attributional beliefs are not only necessary for effective learning and reading, but they furnish the context for the development of executive processes and cognitive strategy use, as well as representing the basis for refining and expanding metacognitive knowledge and enhancing self-efficacy. The difficulties of RD children can best be viewed as reflecting interactive problems between and among various metacognitive components. Borkowski et al.'s (1989, 1990, 1992) model provide one of the most comprehensive models to date for explaining and understanding the differences between proficient and learning disabled readers.

5.2.5 Empirical research supporting aspects of Borkowski's model

As Borkowski's model is large and complex, it does not easily render itself to be testable as a whole (Borkowski et al., 1990, 1992; Groteluschen et al., 1990). A limited collection of studies, mostly correlational and others involving intervention, provide support for parts of

Borkowski et al.'s (1990, 1992) theoretical model. This section describes some key studies that have examined the relationships between metacognitive, cognitive, and motivational variables.

Pintrich and De Groot (1990) examined the relationships between motivation (efficacy, intrinsic value, and anxiety), use of cognitive strategies (e.g., elaboration), metacognitive strategies (e.g., planning, monitoring), effort management strategies (e.g., persistence), and academic performance in seventh graders, using multiple regression techniques. Although data were collected on these variables for English and science, no attempt was made to analyze data by content area. Effort management and metacognitive strategies were placed within a generic category of self-regulation learning strategies because there was evidence of a substantial correlation between these strategies. Intrinsic value did not have a direct influence on performance but was strongly related to self-regulation and cognitive strategy use, regardless of prior achievement level. Self-efficacy and intrinsic value were positively related to use of cognitive and self-regulatory strategies and performance. However, efficacy did not predict subsequent performance when strategies were included in the regression analysis. That is, the expectancy component appeared to play a facilitative role in relation to strategy use such that strategy use variables were more directly tied to actual performance. Self-regulation, self-efficacy, and test anxiety emerged as the best predictors of performance, but this was somewhat dependent on the outcome measure.

Pokay and Blumenfeld (1990), using path analytic procedures, investigated how the relationship between motivation and strategy use changed during a semester in a high school geometry class. Their findings suggested that the relationships among motivation, strategy use, and achievement varied with the time in the year. Early in the semester, expectancies and value predicted the use of a variety of strategies: metacognitive, general cognitive, effort-management, and geometry specific strategies. In addition, expectancies, metacognitive strategies, and effort management strategies predicted grades. Late in the semester the pattern of relationships changed; only value continued to predict strategy use (all types), while geometry self-concept and metacognitive strategy use predicted grades. These findings suggest that students achieve early successes by using task specific strategies, but as skills become established, metacognitive strategies become more important.

Meece, Blumenfeld, and Hoyle (1988) conducted a study with a group of fifth and sixth grade science students to assess the relationships between their task-specific goal orientation,

intrinsic value, attitudes towards science, and their cognitive engagement (a measure of their reported cognitive and metacognitive strategy use). Using structural equation modelling techniques, the authors found that students' intrinsic motivation toward school and science attitudes were significant predictors of goal orientation. Moreover, intrinsic motivation and goal orientation were significant and positive predictors of students' cognitive engagement.

Empirical support for Borkowski et al.'s (1989, 1990, 1992) model, particularly with regard to the relationship of motivational processes and strategy use, is also provided by Schunk's (1989b) research. Schunk suggested that students' strategic knowledge, with respect to the cognitive and metacognitive strategies that make for effective reading, was not sufficient for good performance. It was suggested that students must also believe that (a) the strategies will help them to read better, and (b) they are capable of implementing those strategies efficiently and effectively. Hence, students' efficacy related perceptions of the strategies and of themselves as users of those strategies were critically important in predicting who was able to regulate their own learning. Students who believed both that they know how to employ the strategies successfully and that using the strategies is likely to lead to greater academic success were most likely to engage in effortful attempts to use the strategies about which they have learned (Schunk, 1989b). Thus, goal-directed strategic behaviour is a product of *skill and will* (Palincsar & Ransom, 1988; Paris, 1988).

Empirical evidence for the relationship between self-efficacy and metacognitive processes in the domain of reading has also been established. Paris and Oka (1986) found that students' perceptions of competence were positively related to performance on a reading comprehension task, metacognitive knowledge about reading, and reading achievement. In another study, Oka (1985) found that while under- and over-achievers had similar metacognitive knowledge, they differed in terms of self-perceptions. Oka concluded that underachieving readers "may fail to put their awareness about reading to work because they lack a sense of self-efficacy and self-competence" (p. 14).

The relationship between attributions and strategy use has been established in a number of correlational studies (McCombs, 1988; Palmer & Goetz, 1988; Short & Ryan, 1984). For example, Wagner, Spratt, Gal, and Paris (1989) measured metacognitive knowledge, causal attributions, perceptions, and reading performance from one cohort of 350 first-grade children over a 1 year period, and another cohort of 464 fifth-grade children over a 3 year period. They found even when background variables and cognitive skills were controlled, attributions

and metacognitive knowledge still predicted a significant amount of variance of subsequent reading achievement in the Moroccan children.

Probably the most compelling evidence for a linkage between attributional beliefs and strategy use comes from intervention studies in which attributional beliefs were manipulated in combination with other metacognitive processes (Borkowski, Weyhing, et al., 1988; M. K. Reid & Borkowski, 1987). The goal of these combined reattribution and metacognitive training studies was to transform LD students into active learners by restoring responsibility for learning to them.

The repeated emphasis by researchers that there should be an equal focus on executive processes and reattribution training is supported by research. For example, M. K. Reid and Borkowski (1987) investigated the effects of three instructional conditions on strategy maintenance and generalization in LD children who were hyperactive. In the control condition, children received specific strategy instruction. In the self-control (or executive process) condition, children learned self-regulatory procedures in addition to specific strategy training. Children in the self-control plus attributional condition received strategy and self-control instruction as well as attributional retraining. The findings revealed substantially superior performance of those children in the self-control plus attribution retraining condition. These children demonstrated better strategy maintenance and generalization, and gained higher personal causality scores that endorsed effort. In addition, they maintained their persistent use of strategies at a 10-month follow-up. In fact, they seemed to have undergone permanent improvements in their metacognitive knowledge and attributional beliefs. Thus, support was provided for the conclusion that metacognitive knowledge, executive processes, and strategic beliefs of effort and personal causality interact and play an essential role in the generalization of strategic behaviours.

In summary, there is empirical support for parts of Borkowski's model (Borkowski et al., 1989, 1990, 1992). Most of the research has been correlational in nature and has focused on selected sets of metacognitive components. As yet, research has not identified the independent and interactive effects of a number of metacognitive components on reading comprehension within a causal framework. These metacognitive components include metacognitive knowledge, cognitive/metacognitive strategy use, and motivational beliefs. The investigation of these components on achievement represents a potentially important direction for future research.

5.3 The need for causal models

Resolving the question of how metacognitive, cognitive, and motivational variables are related to reading comprehension performance requires a different research approach than that traditionally undertaken by researchers. Historically, investigators have relied on correlational studies (including regression analysis) or intervention studies to justify the links amongst these variables. The problem with correlational studies is that they do not reveal the causal influences amongst these variables. Although intervention studies can establish causal links, they do not allow one to disentangle the functional components of a complex package or determine the relative influence of the active components. For instance, the metacognitive training study of Borkowski, Weyhing, et al. (1988) showed sizable gains in comprehension skills of poor readers, but did not permit a specification of the relative contributions of the key variables in the training package.

Given the complex systems of cognitive, metacognitive, and motivational factors involved in children's reading achievement, cognitive development is best studied by using research designs which examine these constructs simultaneously (Kurtz, 1991). Simultaneous measurement of cognitive, metacognitive, and motivational variables allows the researcher to test competing theories. In addition, it allows the researcher to understand the relative strengths of each component as well as interactions among them as they influence reading performance, thereby enriching theory and guiding instruction.

Structural equation modelling is ideally suited to an area such as this, in which existing theory justifies the inference of a causal pattern of interrelations among several variables. "The availability of confirmatory, structural analyses makes possible the causal analyses of interactive factors that produce, maintain, or intensify learning disabilities" (Borkowski, Johnston, et al., 1987, p. 169). It has also been suggested that structural models form effective research tools that enable researchers to go beyond the surface features of the data and examine the dynamic interactions associated with learning (Fogel & Thelen, 1987; Jöreskog & Sörbom, 1989).

Typically, investigators assume that the relationship between variables is stable across developmental or skill level if the same pattern of relationships is found at each age or skill level. However, one can only be sure that a relationship is stable across skill level, if one directly compares the pattern of relationships in one skill level with the pattern of

relationships at another skill level. Such investigations are relatively rare. The present study applied multisample structural equation modelling techniques to investigate the generality of the causal structure of cognitive, metacognitive, and motivational variables across RD and NA children. A description of structural equation modelling is provided in Chapter 7.

5.3.1 Causal models of reading achievement

A number of causal models have been developed in recent years which show the relationship between metacognitive abilities at the word level (e.g., phonological and syntactic awareness) to reading achievement (Spedding & Chan, 1993; Tunmer & Hoover, 1992). However, these models were limited in that they did not take into account motivational factors. There has been some investigation of the interaction of motivation and metacognitive/cognitive factors related to academic achievement, but the research has largely been correlational in nature (Paris & Oka, 1986; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990; Zimmerman & Martinez-Pons, 1990). Basically this research showed that students high in self-efficacy or perceived ability are more likely to engage in self-regulation and to make use of cognitive strategies for learning. Generally, the discussion of most previous research has focused on linear relations between single motivational variables and cognitive engagement (metacognitive and cognitive strategy use) variables. However, student involvement in learning includes both motivation and cognition, operating together, not in isolation from one another. Hence, it is important to address the complex and interwoven self- and metacognitive systems that seem responsible for differential academic achievement in NA and RD populations. Furthermore, it is important to examine the causal ordering of these variables in future research.

In a recent study, Carr et al. (1991) tested a structural equation model to explain variations in reading achievement. They compared achieving and underachieving readers to determine possible differences in causal relationships among metacognitive, motivational, and cognitive constructs. Their model was derived from Borkowski et al.'s (1989, 1990) general metacognition model. Carr et al.'s model is unique in that it incorporated both motivational and metacognitive/cognitive factors to explain reading achievement. In addition, Carr et al.'s study represents a key, and perhaps the first, study that has tested Borkowski's model within a causal framework using structural equation modelling techniques. The model is presented in Figure 3.

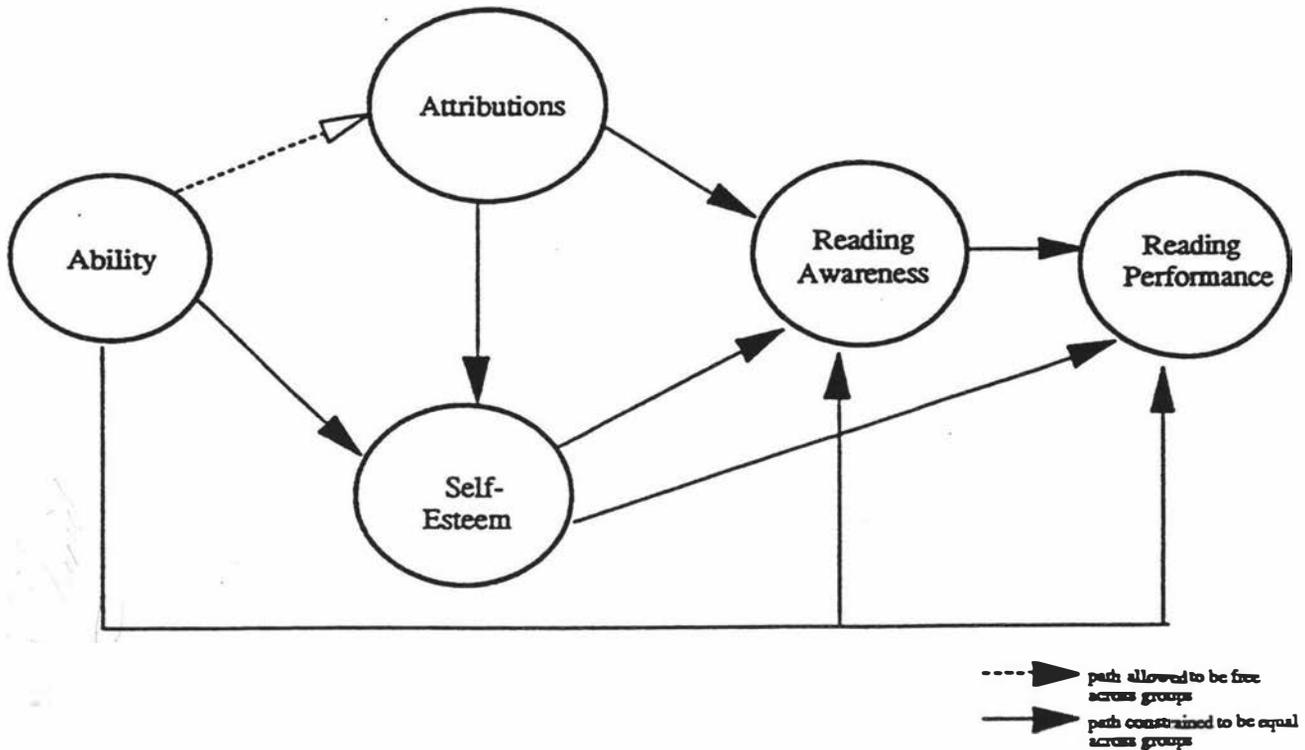


Figure 3. Carr et al.'s (1991) model of reading performance in achieving and underachieving readers.

Carr et al.'s (1991) model differed from Borkowski et al.'s (1989, 1990, 1992) general model in a number of ways. First, Carr et al. explicated more clearly the relationships between motivational beliefs (or self-system constructs) and metacognition. The general metacognition model did not specify how these motivational beliefs interacted with their metacognitive counterparts. In addition, the model did not delineate how the different self-system constructs (attributions, self-esteem, and self-efficacy) were related to each other in their ability to predict metacognitive processes. On the other hand, Carr et al. extended Borkowski et al.'s metacognition model by incorporating attributional beliefs and self-esteem as separate factors and as *antecedents* to metacognitive awareness about reading. These relationships were not delineated in Borkowski's theory; instead the self-system constructs were presented as correlates of general strategy knowledge (cf. Figure 2 in section 5.2.1).

Second, the focus of Carr et al.'s study was on the relationship between ability and attributions in differentiating achieving and underachieving readers (M. Carr, personal communication, February 13, 1995). In the general metacognition model, ability was not

included as a variable. Finally, Carr et al.'s (1991) model was a testable model while the model proposed by Borkowski was too large and complex to be testable as a whole.

The findings of Carr et al. (1991) suggested that for both achievers and underachievers, ability facilitated reading performance as well as the development of reading awareness (metacognitive knowledge). Performance was mediated by attributions, self-esteem, and reading awareness. Although ability has previously been shown to predict self-esteem (Rubin, Chance, & Phares, 1979, cited in Carr et al., 1991), the link between ability and self-esteem was mediated in Carr et al.'s model by attributions about success and failure. Carr and her colleagues argued that positive self-esteem is promoted by adaptive attributional beliefs. In their view, self-esteem and attributional beliefs were expected to enhance reading performance through their impact on reading awareness.

Carr et al.'s (1991) findings also suggested that attributional beliefs were pivotal in determining whether children succeeded in reading. Intellectual ability did not promote the development of adaptive attributions in underachievers. The path from ability to attributions was the only path that discriminated underachievers from achievers. Thus, Carr et al. concluded that "subtle attitudinal differences about learning strategies – as opposed to the complete breakdown of the metacognitive-motivational system – are hypothesized to be an important factor differentiating underachievers and achievers" (p. 109).

In conclusion, Carr et al.'s (1991) model provides a significant advancement in metacognitive theory by testing Borkowski et al.'s (1989, 1990) model of metacognition within a causal framework. In addition, these researchers have extended Borkowski's model by breaking down the general strategy knowledge component and delineating how the different self-system constructs are related to each other and their metacognitive counterparts. Another significant contribution is that they have attempted to test the generalizability of the model across different groups of readers.

5.3.1.1 Limitations of Carr et al.'s (1991) model

(i) Methodological limitations

The study by Carr et al. (1991), while insightful in providing information about the causal relationships between ability, motivation, metacognitive knowledge, and reading performance,

had a number of methodological limitations. First, intellectual ability was measured by the Slosson test of intelligence and the Peabody Picture Vocabulary Test (PPVT). Although these tests were used for pragmatic reasons, they are clearly not ideal measures of ability. For instance, the Slosson test has been criticized for its use of the mental-age concept. It also has limited standardization and reliability (R. L. Taylor, 1984). Furthermore, even though it gives a quick estimate of a person's intelligence, the majority of items are highly verbal and the scores tend to overestimate a person's IQ as determined from more comprehensive measures (R. L. Taylor, 1984). The PPVT, on the other hand, is a measure of receptive vocabulary which presents a limited aspect of general intelligence.

Second, there were a number of issues concerning the adequacy and scoring of the attributional measure. Rather than using a reading specific measure, Carr et al. (1991) used a measure of general achievement attributions. This is questionable as there is evidence for domain specificity of attributional beliefs (Weiner, 1983). Using a global measure of academic attributions impairs the ability to understand and predict behaviour in particular domains and does not take into account the complexity and variation of attributions in that domain.

Furthermore, there may have been a flaw in the way the success and failure attributions were scored. Effort attributions were more heavily weighted than ability (2 Effort + 1/2 Ability) in the scoring procedure and was justified by the authors in that effort was considered to be highly important for strategy use. In spite of this rationalization, this weighted scoring method is considered to be questionable as weighting effort four times that of ability would probably influence the relationships found between attributions and reading awareness in the desired direction. A more equitable approach would have been to give equal weighting to each of the attributions measured.

Third, a general measure of self-esteem was used. The same arguments for attributions apply here. In other words, a global measure of academic self-esteem does not allow clarification of individual differences within the reading domain.

Finally, Carr et al. (1991) used the subscale scores of the Index of Reading Awareness (IRA; Jacobs & Paris, 1987) to form two indicators of metacognitive knowledge. However, the findings of McLain et al. (1991) subsequently suggested that the reliabilities of the subscales of the IRA were too low to support the use of the subscales as separate scores for any

analysis. Hence, the unreliability of the split-half indicators may have confounded the findings of Carr et al.

(ii) *Hypothesized relationships*

The most salient distinguishing feature of RD or underachieving students is that measured intelligence does not predict their reading achievement, whereas intelligence does predict the achievement of their normal achieving peers (Keller & Hallahan, 1987; Short, Feagans, McKinney, & Applebaum, 1984). In other words, for an individual to be called reading disabled, there must be a discrepancy between measured IQ and measured achievement, and reading failure must be "unexpected" on the basis of IQ (Siegel, 1989a).

Although Carr et al. (1991) identified underachievers using a measure of disparity between their achievement and intelligence, they hypothesized that ability would predict achievement and metacognitive knowledge for these children in their model. This hypothesis is not consistent with theoretical and empirical predictions (S. Graham & Harris, 1989; Keller & Hallahan, 1987; Short et al., 1984; Stanovich, 1986a; Swanson, 1991b). The findings of Carr et al. suggested that ability predicted performance and metacognitive knowledge for underachieving readers. However, Carr and her colleagues did not test the competing model that the aforementioned paths may be dissimilar across groups. This competing test would have provided further information as to whether their initial formulated model could be rejected by using incremental-difference statistical tests.

(iii) *Statistical analysis*

In applying structural equation modelling, Carr et al. (1991) constrained variances of their variables to be equal across groups but neither a priori or theoretical justification was provided. In addition, although the authors stated that several alternative models were tested, they did not provide a description of these tests or the findings. Considering differences in Carr et al.'s hypotheses and theoretical predictions concerning the relationship between ability and metacognitive knowledge/performance, the testing of alternative models seemed particularly important as they would have provided some clarification of these relationships. Yi and Nassen (1992) suggested that researchers should compare their proposed model with theoretically relevant, competing models. Only when one can demonstrate that the data are

unlikely to be observed under alternative rival models, can one have confidence in the proposed model. Thus, a consideration of competing models is necessary to evaluate the hypothesized model.

5.3.1.2 Summary

In summary, Carr et al.'s (1991) study, while moving the field forward, had several areas which could be improved with regard to methodology, hypothesized relationships, and the application of structural equation modelling techniques. Clearly, more research is needed to determine the extent to which their model can be replicated. The present study replicated Carr et al.'s study with reading specific measures. It also extended Carr et al.'s study by incorporating metacognitive and cognitive strategy use; variables which have been shown by research to be critical for reading performance.

CHAPTER 6

RATIONALE FOR PRESENT STUDY

6.1 Primary goals of study

The general purpose of the study reported in this dissertation is to investigate how cognitive, metacognitive, and motivational factors causally relate to reading achievement/impairment. This investigation draws on three recent research developments: Borkowski et al.'s (1989, 1990, 1992) model of metacognition; Carr et al.'s (1991) model of reading achievement; and the shift in social science research towards the use of causal models.

In investigating the causal relationships between cognitive, metacognitive, and motivational variables, the present study focuses on a number of specific relationships that requires further clarification in the research literature. These are discussed in section 6.1.2.

6.1.1 Testable interactive or multicomponential causal models

Although several theorists have emphasized the necessity for integrating cognitive, metacognitive, and motivational theories to explain academic performance (Borkowski et al., 1989, 1990, 1992; Short & Weissberg-Benchell, 1989), very few researchers have actually attempted to do this, particularly within a *causal* framework. Moreover, such multicomponential research has yet to be conducted with different populations (e.g., reading disabled), even though researchers have emphasized the need for this type of research (Pokay & Blumenfeld, 1990; Zimmerman & Martinez-Pons, 1990). Reliable and theoretically meaningful models that are generally invariant across populations or subpopulations would contribute much towards theoretical parsimony and progress of educational research. Consequently, the main aim of the study was to generate causal models of cognitive, metacognitive, and motivational variables and to explore the differences in the relationships amongst these variables between reading disabled (RD) and normal achieving (NA) children. Causal models were developed based on an integration of Borkowski's (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990) and Carr et al.'s (1991) models.

As mentioned earlier, previous studies have mostly relied on correlation or simple path analysis. Both of these research techniques have been criticized as being simplistic in theory and design (Byrne, 1989a; Marsh, 1990). Other studies have used intervention studies which reveal to some extent the causal influences among these variables. However, even if researchers show that attributional training and instruction on metacognitive strategies improves reading performance, the question of "How do these components vary in their impact on reading comprehension achievement?" still remains. Hence, a scientific exploration of the interrelationships of cognitive, metacognitive, and motivational variables using a testable causal model is now needed. Structural equation modelling techniques (e.g., LISREL) provide one tool to assist in understanding the dynamics of complex models such as the ones tested in the present study. The pivotal characteristic of LISREL is its ability to probe for multiple influences, as variables can have direct, indirect, or mediated influences on outcomes (A. R. Reynolds & Walberg, 1992).

6.1.2 Issues in need of investigation

From the discussion of issues related to the metacognitive, cognitive, and motivational variables related to reading achievement in Chapters 3, 4, and 5, there are number of controversial, uninvestigated, or unclear aspects of the model in need of clarification. Six of these were selected for examination in the present study.

6.1.2.1 Replication of Carr et al.'s (1991) model

In light of the methodological limitations of Carr et al.'s (1991) study (see Chapter 5, section 5.3.1.1), there is a need for reinvestigation of Carr et al.'s findings using valid domain specific measures and improved methods of defining RD children. In replicating constructs or concepts rather than procedural details of previous studies, the present study provides a "conceptual replication" of previous constructs or ideas (Kidder, 1981).

As pointed out in Chapter 2, the validity of the discrepancy approach as used by Carr et al., (1991) and others in defining RD children has been called into question (see Chapter 2). In particular, the heavy reliance on discrepancy as an identification variable for reading disabilities is neither conceptually nor psychometrically justifiable. In the present study, a multifaceted approach was considered appropriate in defining RD children as it is likely to

facilitate a more accurate identification of the RD population. It should be noted that Carr et al. maintained that their sample consisted of underachieving readers and not reading disabled readers (M. Carr, personal communication, February 13, 1995). However, this can be questioned on two grounds. First, the NJCLD definition and some researchers (e.g., Shaw, Cullen, McGuire, & Brinckerhoff, 1995) suggest that children who have average levels of performance but who show a discrepancy between their intellectual ability and reading achievement may be considered to have a reading disability. Second, the discrepancy method employed by Carr et al. in obtaining their sample of underachieving readers is similar to that used by most researchers in selecting a reading disabled sample. In fact, to most researchers the concept of reading disability has become synonymous with the existence of a discrepancy between intelligence and reading achievement (Mather & Roberts, 1994).

It was argued earlier that one methodological weakness of previous research centred on the lack of specificity in measuring motivational variables. In view of the multidimensional nature of motivational beliefs like self-efficacy, self-concept, and attributional style, it seems legitimate to conclude that general motivational beliefs (e.g., academic self-efficacy, self-esteem) are insufficient to study the effect of self-efficacy or attributions in a specific subject area. Thus, the present study employed reading specific measures in preference to omnibus measures. These domain linked or task-specific appraisals of self-efficacy and attributions generally have higher explanatory and predictive power because of their greater relevance (Assor & Connell, 1992; Bandura, 1990).

6.1.2.2 The mediating role of metacognitive knowledge and strategy use on the relationship between motivational beliefs and performance

Research has demonstrated that metacognitive knowledge is positively related to students' motivational beliefs regarding their efficacy and attributions for performance (Borkowski et al., 1990; Pintrich, 1989). In addition, a sizable body of research has demonstrated that students' use of learning strategies (Pressley et al., 1987; Zimmerman & Martinez-Pons, 1986, 1988, 1990) and metacognitive knowledge (Cross & Paris, 1988; Paris & Winograd, 1990b; Schneider & Pressley, 1989) promote academic achievement.

While metacognitive knowledge and strategy use are important for academic performance, there is also evidence that metacognitive knowledge does not ensure their effective and consistent use (Borkowski et al., 1990; Garner, 1990; Paris et al., 1991). Knowledge of

strategies is necessary for actual strategy use, but it may not be sufficient (Schneider & Pressley, 1989). Students must be motivated to use this knowledge in the form of strategies (Garner, 1990; Paris et al., 1991). Accordingly, it is hypothesized that both self-efficacy and metacognitive knowledge directly predict metacognitive and cognitive strategy use. It is also hypothesized that both self-efficacy and metacognitive knowledge would also directly predict combined strategy use. Combined strategy use represents a global measure of strategy use and includes a combination of metacognitive and cognitive strategies. Such a single factor formulation of strategy use has been conceptualized by previous researchers who refer to it as "cognitive engagement" (Meece et al., 1988). Both cognitive and metacognitive strategies have also been considered to be important components of self-regulated learning by a number of theorists (e.g., Zimmerman, 1989a). It is important to note that in the present study the generic term "strategy use" refers to strategy use in general irrespective of its category (cognitive, metacognitive).

Given the conflicting empirical findings, the relationship between metacognitive knowledge and cognitive strategy use is unclear, particularly with regard to RD children. There is a lack of agreement amongst reading disability researchers as to whether RD children are "inactive" or "inefficient" learners (see Chapter 3, section 3.4.4.1). The present study uses LISREL to test the competing positions of the "inactive" and "inefficient" learner views. In addition, the nature of the relationship between metacognitive knowledge and strategy use across NA and RD readers was also examined.

When causal modelling has been used in previous research, key variables identified as influencing reading performance, most notably metacognitive knowledge, and/or metacognitive/cognitive strategy use, have been excluded. The following relationships remain uninvestigated: (a) the causal impact of students' use of strategies on reading achievement, (b) the mediating role of metacognitive knowledge on motivational beliefs and strategy use, and (c) the mediating role of strategy use on metacognitive knowledge and academic performance. Furthermore, it is unclear whether the nature of the aforementioned relationships is similar for both NA and RD readers.

Carr et al. (1991), in their causal model of reading achievement, used two indicators or measures of reading performance: cognitive strategy use and reading grades. For the present study it is argued that it may be preferable to consider strategy use and reading performance as separate components in the model. This would provide a clearer understanding of how

strategy use affects performance. Given that there is a strong relationship between strategy use and reading performance, and that metacognitive knowledge is necessary for strategy use, it is considered important to examine the causal ordering of these variables. It is hypothesized that metacognitive knowledge will mediate the relationship between motivational beliefs (efficacy and attributions) and strategy use. It is also predicted that strategy use mediates the relationship between metacognitive knowledge and reading comprehension. Support for these predictions comes from the models of Carr et al. (1991) and Borkowski (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990).

6.1.2.3 The relative impact of self-efficacy and attributional style on metacognitive knowledge

Although sometimes confounded by imprecise definitions and varying methods of measurement, motivational beliefs such as causal attributions and self-efficacy have each been linked to students' metacognitive knowledge. However, it is not clear how these processes may work together to predict metacognitive knowledge.

Correlational data on efficacy and attributions have supported the positive relation between these constructs and metacognitive knowledge. Following Carr et al.'s (1991) findings, it was posited that both attributions and self-efficacy would have a direct effect on metacognitive knowledge. The relative influence of attributions and self-efficacy on students' metacognitive knowledge about reading was investigated in the present study, as this aspect of Borkowski et al.'s (1990, 1992) theory has yet to be tested. On the basis of Borkowski's theory, it was hypothesised that attributional style would have a stronger positive impact on metacognitive knowledge than would efficacy.

6.1.2.4 The effect of self-efficacy on strategy use and performance

According to the theoretical formulation of Borkowski's model, self-efficacy is expected to have a direct effect on metacognitive strategy use but not on cognitive strategy use (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). However, empirical findings do not support this prediction. For instance, research has shown self-efficacy to be a significant predictor of cognitive strategy use (Pintrich & De Groot, 1990), metacognitive strategy use (Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990), and general strategy use (Ames & Archer, 1988) among adolescent students.

These competing empirical and theoretical predictions suggest a further need for research in this area to clarify the relationship between efficacy and cognitive/metacognitive strategy use. In accordance with empirical findings, it is hypothesized that efficacy would predict both cognitive and metacognitive strategy use. This would also mean that efficacy is a direct predictor of combined strategy use.

The experimental and correlational studies of Schunk (1985, 1989b) with elementary and junior high students indicate that self-efficacy is consistently related to academic performance. The findings of Pintrich and De Groot (1990) indicate that self-efficacy has an indirect effect on performance, suggesting that self-efficacy may play a facilitative role when cognitive and metacognitive strategy variables are included. However, this may depend on the outcome measures used. Studies that used domain specific measures of cognitive skills showed a direct relationship between self-efficacy and performance (Schunk, 1989b; Shell, Murphy, & Bruning, 1989). In contrast, those that used global measures of performance showed self-efficacy to have an indirect effect on performance when strategy use variables are included (Pintrich & Schrauben, 1992). In the present study, domain specific measures of academic performance were used; consequently it was hypothesized that self-efficacy would also have a direct effect on performance.

6.1.2.5 The influence and relative impact of cognitive and metacognitive strategy use on performance

It was previously argued that the use of cognitive strategies should be distinguished from the use of metacognitive strategies as distinct components of strategy use (see Chapter 3, section 3.2.3). Evidence of such a distinction is supported by both theory and research. A large number of studies attest to the strong relationship between cognitive/metacognitive strategies and performance. The present study investigated the influence of these strategies on reading comprehension.

Contrary to Borkowski et al.'s (1989, 1990, 1992) model, it is predicted that metacognitive strategy use is directly linked to reading comprehension performance. Although Borkowski's theoretical model specifies that the relationship between metacognitive strategy use and performance is mediated by cognitive strategy use (Borkowski & Turner, 1990; Groteluschen et al., 1990), empirical findings based on path analysis and multiple regression suggest

metacognitive strategy use may have a direct effect on performance (Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990).

Empirical research also suggests that cognitive strategy use is a direct predictor of performance, particularly for NA children (Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990). However, with regard to RD readers, there has been much speculation on the nature of the relationship between cognitive strategy use and performance. It is unclear whether cognitive strategy use directly predicts reading achievement for RD readers. Research to date has yet to compare the nature of the relationship between cognitive strategy use and performance for NA and RD readers. With regard to RD readers, both the "inactive learner" and the "inefficient learner" perspectives implicate deficiencies in the use of cognitive strategies for learning. Thus, both views would suggest that cognitive strategies would not predict performance for RD students. In the present study, it is hypothesized that the relationship between cognitive strategy use and performance would differ between NA and RD readers.

The relative impact of the distinct components of strategy use on reading performance is generally an under-researched question and would have important implications for educational planning. This question of relative impact is examined in the present study. It is predicted that cognitive strategy use will be more strongly related to performance than metacognitive strategy use for NA readers. This prediction is based on Pokay and Blumenfeld's (1990) study, which indicated that geometry specific and general cognitive strategies were more strongly correlated with mathematical performance than metacognitive strategy use.

6.1.2.6 The relative impact of success and failure attributions on efficacy and metacognitive knowledge

According to Borkowski's theory, attributional beliefs represent a key variable in the development of metacognitive processes (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990). However, achievement theories to date have not discussed the differential effects of subcomponents of attributions on metacognitive development. Evidence that differential effects may occur amongst subcomponents of attributions is provided in the depression literature (Sweeney et al., 1986). Furthermore, Heider (1958) suggested that different attributions for success and failure have distinct consequences for the individual's affective reaction, expectancy of future success, and subsequent behaviour.

There is indirect evidence in the achievement literature that different elements of attributions (which vary on the three causal dimensions) have a differential impact on achievement. In the attribution literature, attributions have generally been measured as specific factors or elements (e.g., ability, effort). An investigation of the relative impact of success and failure attributions on metacognitive knowledge would extend previous research and theory. The significance of this model rests on the premise that attributions for different outcomes (success and failure) will have distinct consequences for the individual's metacognitive knowledge and efficacy. The present study tests the validity of this assumption.

6.2 Hypothesized Models

A theoretically parsimonious model is one that would be invariant, to a large degree, across populations. This consideration dictated that few differences in the relationships amongst the variables be specified between groups. In doing so, a conservative model will be generated which explains the differences between NA and RD readers. Four models were generated to investigate the six issues discussed above and to investigate how relationships among motivational, cognitive, and metacognitive variables varied between NA and RD readers.

The four models depicted in Figures 4 through 7 specify the causal relationships amongst the variables, and how they vary between NA and RD readers. A large number of hypotheses can be generated from these models. However for practical reasons, the present study limits the number of hypotheses by focusing on specific aspects that are unique and of particular interest in the present study.

The models in the present investigation included two multifaceted constructs; namely attributional style (attributions) and combined strategy use. Each of these multifaceted constructs was composed of two auxiliary concepts, each of which could be distinguished conceptually from the other and measured separately, despite being related to each other both theoretically and empirically. Attributional style was considered as a multifaceted construct comprising two major subcomponents of achievement outcome: attributions for success and attributions for failure. Combined strategy use also consisted of two subcomponents: metacognitive strategy use and cognitive strategy use. One important advantage of combining facets of a multidimensional construct, is that it simplifies the data analysis and the conceptual explanation of the findings (Carver, 1989). However, summing measures that are conceptually distinct is likely to lead to a loss of information and erroneous conclusions. Carver (1989) suggested that without testing each component separately, in addition to the test

of the composite index, one cannot be sure of the basis by which the composite was associated with the outcome measure. Hence, these multifaceted constructs were initially tested as one factor models by combining the subcomponents into a single predictive index. This was also done on the basis of parsimony and its conceptual consistency with theoretical formulations of previous researchers who have used such constructs (Carr et al., 1991; Meece et al., 1988). The separate effects of the subcomponents on the rest of the model were also examined.

The first model is tested to replicate some of Carr et al.'s (1991) findings. The second model extends previous research by testing the multicomponential influence of attributions, self-efficacy, metacognitive knowledge, and combined strategy use in predicting performance. The third and fourth models examine in detail particular latent constructs that are of theoretical interest in Model 2. More specifically, the differential effects of the subcomponents of combined strategy use and attributional style are investigated. According to Borkowski et al. (1989, 1990, 1992), attributional style and components of combined strategy use are considered important in explaining differences between NA and RD readers.

Note that the four models depicted in Figures 4 through 7 have been modified a posteriori and have been presented here in order to clarify the subsequent analyses and results. These a posteriori modifications are mentioned as footnotes under the figure. For example, one of the footnotes suggests that the path from ability to metacognitive knowledge was fixed to zero. This was because the preliminary results had indicated that this relationship was nonsignificant for both groups.

6.2.1 Proposed General Model 1

Figure 4 presents a graphic representation of the first proposed general structural model. The labelled large circles represent the latent (unmeasured) constructs and the rectangles represent observed (or measured) variables. The unfilled arrows from the large circles to the rectangles indicate that the observed variables are theorized to be generated by the latent construct. Relationships that are hypothesized to be similar between groups have their causal paths constrained to be equal (solid line), whilst those that are hypothesized to differ between groups have their causal paths freed between groups (dotted line).

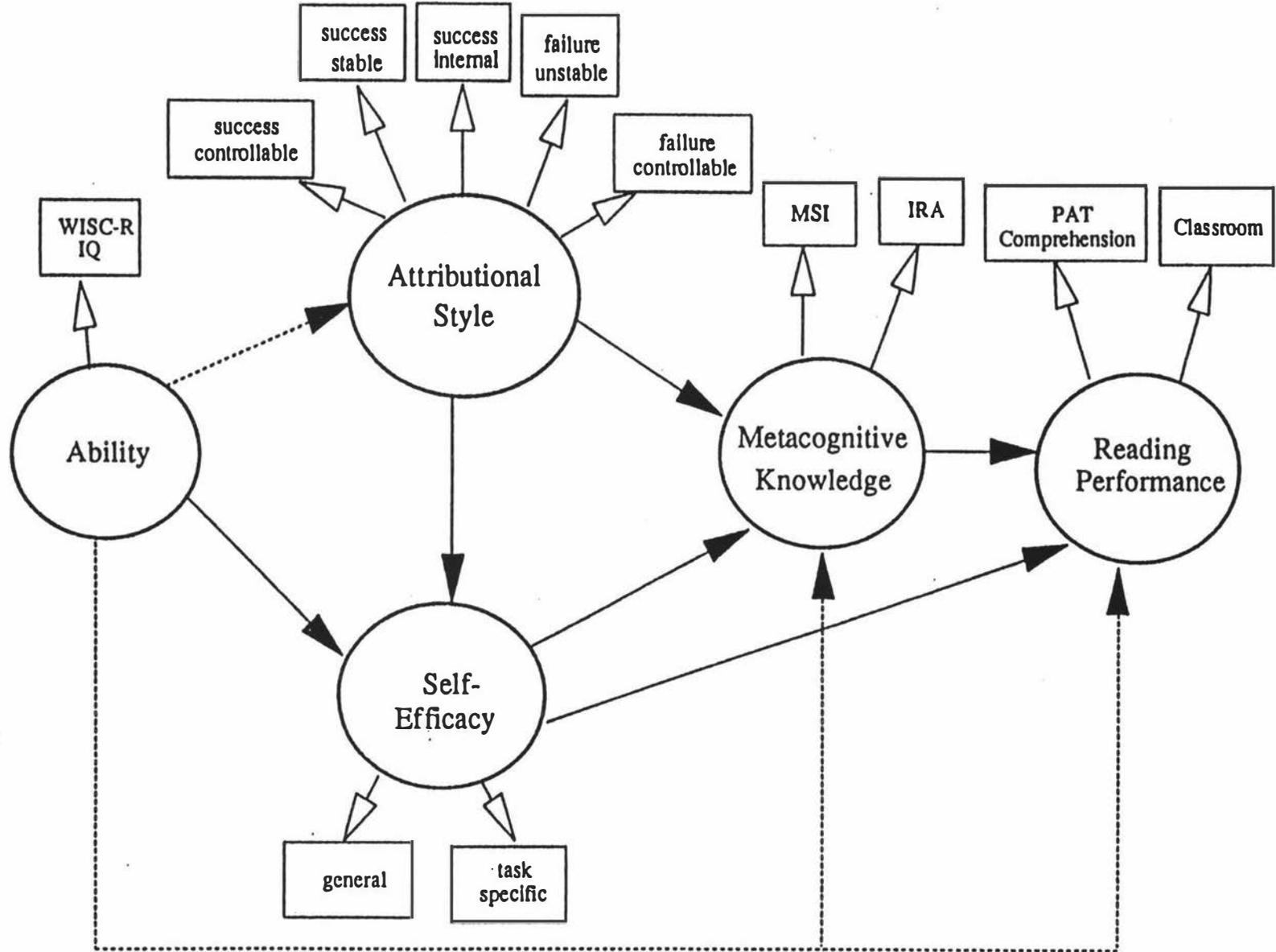


Figure 4. Hypothesized theoretical Model 1 for NA and RD readers.

Note: Failure Internal attributions were dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero for all competing models except Model 1.1.

.....▶ path allowed to be free across groups
 —▶ path constrained to be equal across groups

Only five indicators are used to represent the construct of attributional style due to a modification to this model post facto (see Chapter 8, section 8.3.4). Students with an *adaptive* attributional style ascribe the causes of their success to internal, stable, and controllable factors, and their failure to controllable, *unstable* factors (e.g., Weiner, 1985a). Hence, the latent construct of attributional style is represented by manifest variables of success attributional dimensions of locus of causality (internality), controllability, and stability for success, and failure attributional dimensions of controllability and *instability*. In other words, all measures were scored so that higher scores implied adaptive attributions. General and task specific measures of confidence in reading represent indicators for self-efficacy. The metacognitive knowledge construct is represented by two metacognitive knowledge questionnaires, the MSI and the IRA. Standardized reading comprehension scores (PAT) and classroom scores are manifest indicators of the reading achievement construct. The latent construct of ability has one indicator, intellectual ability as measured by the WISC-R.

The components of this model are based on Carr et al.'s (1991) study. However, the measures used are not identical. Model 1 tested the plausibility of a postulated causal system comprising the latent variables of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance. The relations postulated in this model differed slightly from that proposed by Carr et al. (1991). They are as follows:

1. In this study, a direct relationship between ability and self-efficacy is specified in accordance with Bandura's (1986) theory and previous empirical findings. In Carr et al.'s (1991) model, the link between ability and self-esteem was mediated by attributions about success and failure.
2. In the present study it is hypothesized that the direct relations between ability and metacognitive knowledge, and between ability and reading comprehension performance would differ across groups. Carr et al. (1991) suggested the converse. Further, in contrast to Carr et al.'s model, the present study also postulates that three paths will discriminate RD children from NA children: the paths from ability to attributional style, ability to metacognitive knowledge, and ability to performance.
3. Compared to Carr et al. (1991), the present study used a more domain specific measure of expectancy, namely reading self-efficacy. In accordance with empirical research findings, it is hypothesized that self-efficacy will have a direct effect on reading comprehension performance. Such a relationship was not specified for self-esteem in Carr et al.'s (1991) model.

Apart from the aforementioned paths, the present study aims to replicate the causal ordering of the variables found in the model tested by Carr et al. (1991).

6.2.1.1 Hypotheses

It is hypothesized that:

1. Ability directly predicts self-efficacy.
2. Attributional style directly predicts self-efficacy.
3. Efficacy mediates the relationship between attributions and metacognitive knowledge.
4. Attributional style would have a stronger direct effect on metacognitive knowledge than efficacy.
5. Self-efficacy directly predicts reading comprehension performance.
6. Not all the relationships between motivational, metacognitive, and cognitive variables would vary in magnitude across NA and RD readers.
7. The magnitude of the relationship between ability and attributions would differ for NA and RD readers. More specifically, the relationship between ability and attributions is not significant for RD readers, but significant for NA readers.
8. The relationship between ability and performance would differ for NA and RD readers. More specifically, the relationship between ability and performance is significant for NA readers but not for RD readers.
9. The relationship between efficacy and performance would be similar for NA and RD readers.
10. The relationship between efficacy and metacognitive knowledge would be similar for NA and RD readers.
11. The relationship between metacognitive knowledge and performance would be similar for NA and RD readers.
12. The relationship between attributions and efficacy would be similar for NA and RD readers.

6.2.2 Proposed General Model 2

The second model (see Figure 5) represents an extension of Model 1 by incorporating the latent construct of "combined strategy use". The manifest variables of this construct include a self-report measure and a behavioural measure of global strategy use. Each indicator represents a global measure of strategy use which included a combination of cognitive and

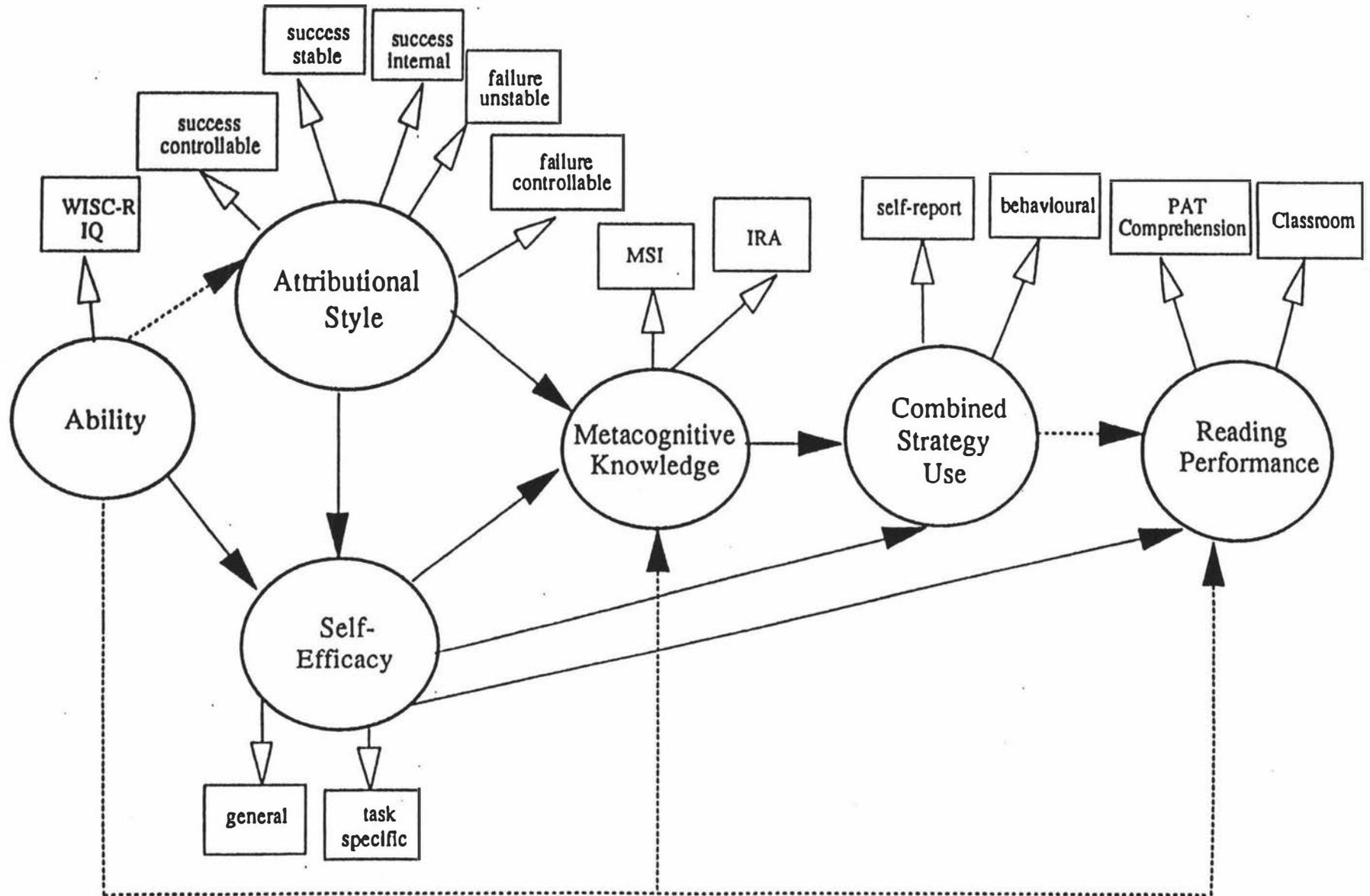


Figure 5. Hypothesized theoretical Model 2 for NA and RD readers.

Note: Failure internal attributions were dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero for all competing models.

.....▶ path allowed to be free across groups
 —▶ path constrained to be equal across groups

metacognitive strategies. A one factor representation of strategy use is consistent with formulations of previous researchers (Meece et al., 1988) and self-regulated learning theorists (e.g., Pintrich & De Groot, 1990).

The proposed model specifies that attributional style predicts metacognitive knowledge directly or indirectly through self-efficacy and that combined strategy use mediates the effect of metacognitive knowledge on performance. The model also specifies a direct relationship between efficacy and combined strategy use. In this model, it is hypothesized that the following four paths will discriminate RD children from NA children: ability to attributions, ability to metacognitive knowledge, ability to performance, and combined strategy use to performance. All relationships postulated are considered to be in the positive direction.

6.2.2.1 Hypotheses

It is hypothesized that:

1. Metacognitive knowledge mediates the relationship between motivational beliefs (attributions, efficacy) and combined strategy use.
2. Combined strategy use mediates the relationship between metacognitive knowledge and reading comprehension performance.
3. Self-efficacy directly predicts combined strategy use.
4. The relationship between metacognitive knowledge and combined strategy use would be similar for NA and RD readers.
5. The magnitude of the relationship between combined strategy use and performance would differ for NA and RD readers.

6.2.3 Proposed General Model 3

In Model 3, combined strategy use was respecified to form two latent variables. An implicit assumption underlying strategy use in Model 2 is that the responses to items used to define each measured variable are reasonably unidimensional. In this way the multiple indicators of a particular latent construct can be used to test the underlying structure and to assess measurement error. However, this approach may not be appropriate when the formation of

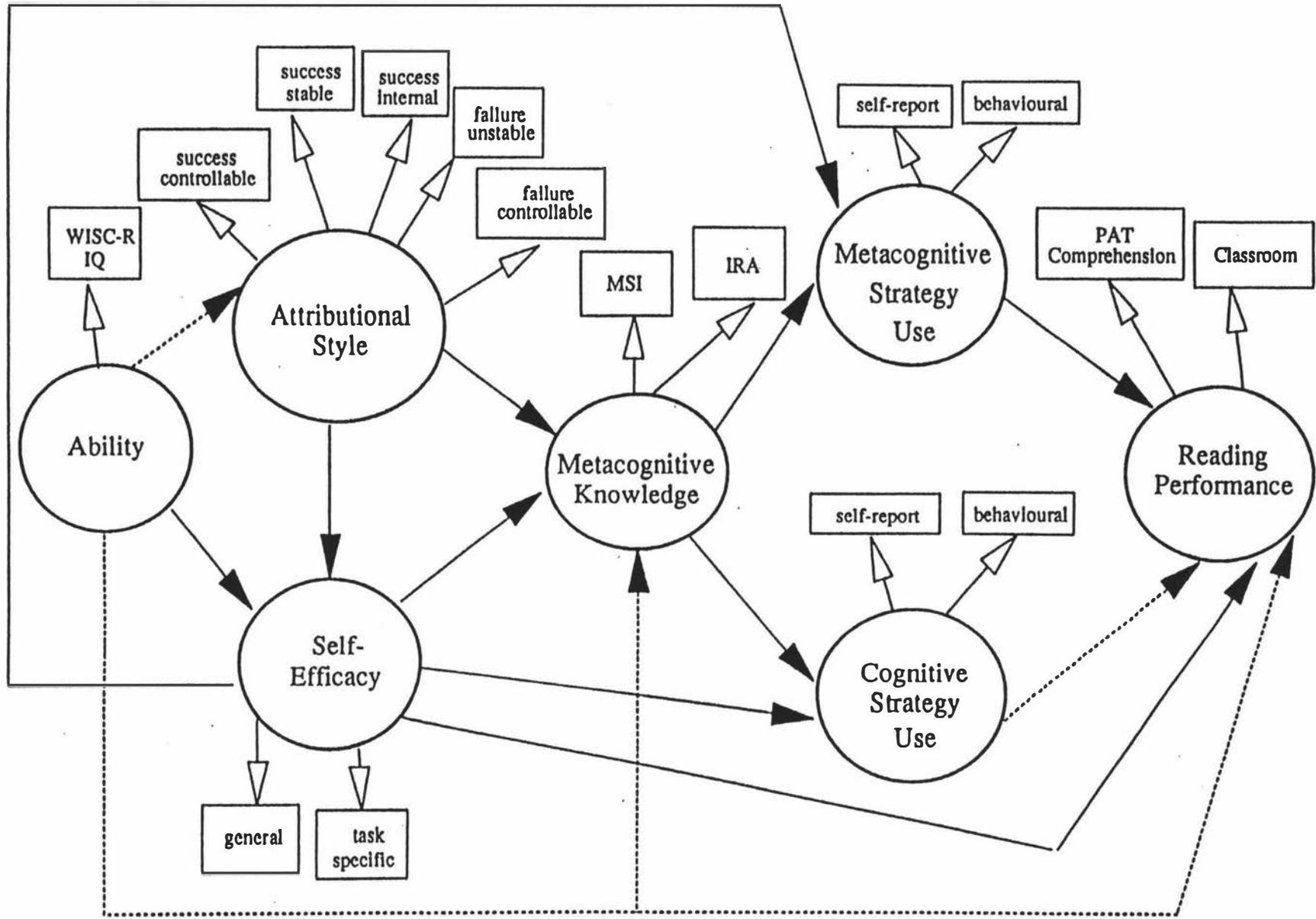


Figure 6. Hypothesized theoretical Model 3 for NA and RD readers.

Note: Failure Internal attributions were dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero for all competing models.

-----> path allowed to be free across groups
 —————> path constrained to be equal across groups

item clusters confounds information from different domains (e.g., cognitive versus metacognitive) which are substantively important and differentially related to other latent constructs of interest. According to Borkowski's theory, it seems likely that responses to different clusters of items of strategy use would be differentially related to performance (Borkowski et al., 1989, 1990, 1992). In other words, distinct components of strategy use are likely to have differential effects on the other latent constructs of interest in this study. In addition, empirical research in the domain of mathematics, has demonstrated that cognitive strategy use is more strongly related to performance than metacognitive strategy use for NA readers. Consequently, by averaging responses to these different components of strategy use, such potentially important distinctions cannot be evaluated.

Model 3 (see Figure 6), an extension of Model 2, investigates in more detail the latent construct of combined strategy use by examining the differential effects of its metacognitive and cognitive subcomponents. The latent construct of metacognitive strategy use is specified by two indicators: a self report measure and a behavioural measure of metacognitive strategy use. Similarly, a self report and behavioural measure of cognitive strategy use represent the indicators of the latent construct of cognitive strategy use.

Model 3 specifies that both cognitive and metacognitive strategy use mediate the relationship between metacognitive knowledge and reading performance, particularly for NA readers. Metacognitive knowledge mediates the relationship between motivational beliefs (attributions, self-efficacy) and metacognitive/cognitive strategy use. It is hypothesized that four paths will discriminate RD children from NA children; the paths from ability to attributional style, ability to metacognitive knowledge, ability to performance, and cognitive strategy use to performance. All relationships postulated are considered to be in the positive direction.

The paths specified in Model 3 are in accordance with theoretical and/or empirical predictions. Two paths that were specified in Borkowski's theory are excluded from the proposed model (Borkowski et al., 1992; Groteluschen et al., 1990); instead they are tested as competing models. The paths include those between attributional style and metacognitive strategy use, and between metacognitive strategy use and cognitive strategy use.

6.2.3.1 Hypotheses

It is hypothesized that:

1. Metacognitive knowledge directly predicts both cognitive and metacognitive strategy use.
2. Self-efficacy directly predicts both cognitive and metacognitive strategy use.
3. Metacognitive strategy use directly predicts reading comprehension performance.
4. Both metacognitive and cognitive strategy use would mediate the relationship between metacognitive knowledge and reading comprehension performance for NA readers. Only metacognitive strategy use would mediate the relationship between metacognitive knowledge and reading comprehension performance for RD readers.
5. Cognitive strategy use is more strongly related to performance than metacognitive strategy use for NA readers.
6. The relationship between cognitive strategy use and reading comprehension performance would differ for NA and RD readers. More specifically, the relationship between cognitive strategy use and performance is significant for NA readers but not for RD readers.
7. The relationship between metacognitive strategy use and reading comprehension performance would be similar for NA and RD readers.
8. The relationship between metacognitive knowledge and metacognitive strategy use would be similar for NA and RD readers.
9. The relationship between metacognitive knowledge and cognitive strategy use would be similar for NA and RD readers.
10. The relationship between self-efficacy and metacognitive strategy use would be similar for NA and RD readers.
11. The relationship between efficacy and cognitive strategy use would be similar for NA and RD readers.

6.2.4 Proposed General Model 4

Theoretical Model 4 (see Figure 7) represents an extension of Model 3, and explores the differential effects of the attributional subcomponents. No empirical evidence or theory exists to explain how these subcomponents would differ in their impact on self-efficacy and metacognitive knowledge. As such, no hypotheses were formulated regarding their differential impact. However, hypotheses were formulated regarding how these relationships would vary across groups. This model specifies that both success and failure components of attributional style directly predict self-efficacy and metacognitive knowledge.

6.2.4.1 Hypotheses

It is hypothesized that:

1. Ability significantly predicts success and failure attributional style for NA readers.
2. Ability does not significantly predict success and failure attributional style for RD readers.
3. The relationship between success attributions and metacognitive knowledge is similar for NA and RD readers.
4. The relationship between failure attributions and metacognitive knowledge is similar for NA and RD readers.
5. The relationship between success attributions and efficacy is similar for NA and RD readers.
6. The relationship between failure attributions and efficacy is similar for NA and RD readers.

6.3 Secondary goals of study

6.3.1 Differences in motivational, cognitive, and metacognitive variables across groups

The secondary focus of this study involved examining how attributions, efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use would vary between NA and RD readers. This is investigated for two reasons:

1. It would provide validation of the multifaceted selection criteria used in the present study for selecting RD readers.
2. It would clarify how motivational, cognitive, and metacognitive variables vary across NA and RD readers. Research has sometimes provided conflicting empirical findings regarding the attributional beliefs, efficacy, metacognitive knowledge, and cognitive strategy use of RD readers.

6.3.1.1 Hypotheses

It is hypothesized that compared to NA readers:

1. RD readers will have lower self-efficacy and possess less adaptive attributions.
2. RD readers will have less metacognitive knowledge for reading.

3. RD readers will engage in less metacognitive and cognitive strategy use.

6.4 Summary

In summary, the present study draws on self-efficacy, attribution, and metacognitive theories of achievement to identify predictors of reading achievement/impairment. It examines the interactive influence of ability, attributional style, self-efficacy, metacognitive knowledge, cognitive strategy use, and metacognitive strategy use on young adolescents' reading comprehension performance. It also examines how these relationships vary across NA and RD readers. This research extends previous research in several important ways. First, earlier findings are replicated with a different sample of adolescents. Second, the role played by metacognitive knowledge and strategy use in mediating the relationship between motivational beliefs and reading comprehension performance is examined. Third, the differential effects of cognitive and metacognitive strategy use on reading achievement are investigated. In addition, the differential effects of attributions for success and failure on metacognitive knowledge and self-efficacy are explored. To date, these relations have not been investigated. Fourth, the relations between predictor variables and their causal links to outcome variables are tested with more sophisticated structural equation modelling techniques than were used in previous studies. This study also uses sophisticated group comparison procedures to assess possible population related differences in the predictive influence of cognitive, metacognitive, and motivational variables on reading comprehension performance. Finally, the present study limits the analyses to a sample of junior high school (intermediate) students to control for possible developmental effects.

CHAPTER 7

STRUCTURAL EQUATION MODELLING

Structural equation models have been referred to variously as covariance structure models, analysis of covariance structures, linear structural relations models, latent variable structural equation modelling, and perhaps most commonly as the LISREL models (Fergusson, 1995; Jöreskog & Sörbom, 1989, 1993; Long, 1983b). They have been useful in solving many substantive problems in the social and behavioural sciences. Such models are now being used in psychology and education, in addition to areas of sociology, econometrics, and marketing. However, the use of structural equation modelling (SEM) in special education research is relatively rare and is only in its infancy (A. D. Moore, 1995). The purposes of this chapter are to: (a) provide a general introduction to, and explanation of, structural equation models, and (b) describe how the hypothesized models that were formulated earlier (see Chapter 6, section 6.2) could be represented as structural equation models.

Although various forms of the general structural equation model have been advanced in the literature (e.g., Bentler, 1982; Jöreskog & Wold, 1982), the most popular and arguably the most useful formulation of structural equation modelling is the LISREL model described by Jöreskog and Sörbom (1989, 1993). The widespread use of LISREL model specifications and accompanying computer software has resulted in the term "LISREL" becoming virtually synonymous with structural equation modelling. The term has also been used to refer to the computer software used to fit structural equation models. The form of the covariance structure model that is presented in this thesis is essentially that incorporated in LISREL and will be referred to as the covariance structure model or the structural equation model. The LISREL formulation is used for two reasons. First, it provides a simple and readily accessible framework for the specification and testing of a vast range of models. Second, it is currently the most widely used general model for the analysis of structural equation models. A brief account of its terminology and methodology is given below. This is largely based on the accounts of structural equation modelling given by Bollen (1989), Bollen and Long (1993), Hayduk (1987), Long (1983a, 1983b), Fergusson (1995), and Fergusson and Horwood (1988).

7.1 An introduction to structural equation modelling and the LISREL model

Structural equation modelling (SEM), also known as covariance structure analysis, represents a set of techniques for theory testing with correlational data. The theories that can be tested are those that can be represented as a system of equations that describe unidirectional and/or bidirectional influences of several variables on each other. Typically, a structural equation model is specified via a simultaneous set of structural linear regressions of particular variables on other variables. In other words, these equations describe some postulated model of the *causal* relationships between certain variables of theoretical interest. The method is called covariance structure analysis because the implications of the simultaneous regressions are studied primarily at the level of correlations or covariances.

SEM combines the advantages of factor analysis with path analysis. First, latent factors (theoretical constructs) are derived by factor analysis from the multiple observed measures of each construct. Second, path coefficients are estimated for the paths that presumably relate one set of latent factors to another. Since latent factors are by definition error free, some of the dangers of path analysis with manifest variables are avoided. Hence, the correlation between the latent factors provides a more accurate estimation of the relationship between the constructs they represent than the correlations between pairs of observed variables.

Structural equation models have been defined by Fergusson (1995) as describing

the measurement and structure of a set of observed variables in terms of simultaneous linear equations which summarise and encapsulate the assumptions of some prespecified conceptual model of the measurement and causal relationships of the observed variables. (p. 122)

Although it is important to recognize that structural equation models seldom provide any direct test of the causal assumptions on which the models are based, such models can be thought of as "a formal statement of the investigator's beliefs about the causal structures that underlie a set of observed variables" (Fergusson & Horwood, 1988, p. 325).

Most theories in the social and behavioural sciences are formulated in terms of theoretical hypothetical constructs that cannot be measured directly (e.g., self-esteem, motivation, metacognition, ability, school performance). Such hypothetical, non-observed constructs are

commonly termed latent variables. The measurement of the hypothetical constructs is done indirectly by measuring multiple observable variables which are designated as indicators of the underlying theoretical construct (see also section 7.3).

The investigator defines the hypothetical constructs by specifying the dimensions of each construct (see Bollen, 1989, pp. 179-181). For example, the different dimensions of attitudes, persistence, and effort-expenditure may be united to form the construct or concept of motivation. The theory specifies further how the various hypothetical theoretical constructs are postulated to be interrelated. This includes: (a) classification of the constructs into dependent and independent constructs; (b) specifying for each dependent construct, which of the other constructs it is postulated to depend on; and (c) specifying the sign and/or relative size of the direct effect of one construct on another.

The theory cannot be tested directly since the theoretical latent constructs are not observable. This means that the researcher can only examine the *theoretical validity* of the theorized relationships in a given context. A set of operationable variables must be defined for each dimension of each construct before the theory can be tested empirically. There must be clear rules of correspondence between the measured variables and the theoretical constructs such that each dimension and construct is distinct.

The theoretical postulations between the latent constructs form the *structural* part of the model, and the relationships between the observable indicators and the theoretical latent constructs form the *measurement* part of the model. The testing of the structural model may be meaningless unless it is first verified that the measurement model holds. If the chosen measures or indicators for a construct do not measure that construct, then it may not be possible to provide an adequate test of the structural model. Therefore, it may be useful to test the measurement model first. In order to test the model, each of the structural and measurement parts must be formulated as a statistical model. This will be discussed in the following section.

The analysis is predominantly *confirmatory* in nature, that is, it seeks to determine the extent to which the postulated structure (as described by the linkages among the latent variables and among the latter and their indicators) is actually consistent with the empirical data at hand. This is done by computing the variance-covariance matrix of the observed indicator measures as implied by the specified model (Σ) and comparing it to the actual variance-covariance

matrix based on the empirical data (S). The variance-covariance matrix of the set of observed variables are assumed to be measured on interval scales and to have a mean of zero.

In the remaining sections of this chapter, the statistical theory and methodology for testing covariance structure models will be developed within the framework of the LISREL model as described by Jöreskog and Sörbom (1993). This is done only for models that meet the following assumptions: (a) linearity of structural relationships, (b) interval scaling of all variables, whether latent or observed, and (c) multivariate normality of the distribution of the observed variables. For simplicity, this theoretical development focuses on a consideration of models for a single population initially. A brief description of how the theory can be applied to the models for two populations or more (referred to as multisample analysis), and the justification for the use of LISREL-type models will follow. The process of testing a given model involves four related aspects: model specification, identification, parameter estimation, and model evaluation.

7.1.1 Model specification

Latent (unobserved) variables can be classified as endogenous or exogenous. If a latent variable is directly caused or influenced by any of the other latent variables, it is classified as endogenous. However, if a latent variable acts as a "cause" and not as an "effect," then it is exogenous. Therefore, two types of observed variables are distinguished by the LISREL model:

1. the x variables, which are indicators of latent independent or exogenous variables.
2. the y variables, which are indicators of latent dependent or endogenous variables.

The general structural equation model consists of two complementary models:

1. *the measurement model*, which provides a set of equations that describe how each of the latent constructs is operationalized via the observed or manifest variables (x , y).
2. *the structural model*, which provides a set of equations that describe the relationships between the latent constructs, and reflects substantive hypotheses based on theoretical considerations.

The general LISREL procedure can be used to assess either measurement or structural equation models and when appropriate, both simultaneously.

7.1.1.1 Measurement model

The measurement model for the x variables is

$$X = \Lambda_x \xi + \delta$$

- where:
- X is a $q \times 1$ vector of observed predictor variables, x .
 - ξ is an $n \times 1$ vector of latent independent or exogenous variables.
 - Λ_x is a $q \times n$ parameter matrix of regression coefficients that describe the relationships between the observed variables x and the latent independent variables, ξ . In other words, it is the matrix of the loadings of the x 's on the ξ 's.
 - δ is a $q \times 1$ vector of errors of measurement (or disturbance terms) in x .

Similarly, the measurement model for the y variables is

$$Y = \Lambda_y \eta + \epsilon$$

- where:
- Y is a $p \times 1$ vector of observed responses on outcome variables, y .
 - η is an $m \times 1$ vector of latent dependent or endogenous variables.
 - Λ_y is a $p \times m$ parameter matrix of regression coefficients that describe the relationships between the observed variables y and the latent dependent variables, η . In other words, it is a the matrix of the loadings of the y 's on the η 's.
 - ϵ is a $p \times 1$ vector of errors of measurement (or disturbance terms) in y .

As can be seen from the equations, the variances of the observed x and y variables can be partitioned into components attributable to the latent variables of the model (a true score) and to errors of measurement (also known as unique factors). The error variables (δ , ϵ) specify the cumulative effects of excluded variables and purely random errors on the observed x and y variables.

Measurement error (i.e., errors in recorded responses) may be interpreted as a "residual" covering a variety of influences, in addition to that of the true score, which produce the observed score. According to Groves (1989), a variety of factors which vary across individuals can influence subjects' responses. These include (a) questionnaire method effects,

(b) subjects' inability or lack of requisite effort to answer questions or obtain correct answers, (c) deficiencies in the wording of questionnaires, or (d) effects due to mode of data collection (e.g., written questionnaires, test administration, use of face-to-face versus telephone communication). Groves suggested that all these factors may be included as that part of the measurement which is assigned to measurement error. The measurement model thus points to the existence of a true score for each of the particular theoretical variables of interest. However, the existence of measurement error means that it is not possible to examine these true scores directly. Consequently, the concept of a latent (unobserved, unmeasured) variable is introduced to represent the true score.

7.1.1.2 Structural model

The relationships between the latent variables are represented by the structural model and can be expressed as:

$$\eta = B \eta + \Gamma \xi + \zeta$$

where:

- B is a $m \times m$ matrix of regression coefficients which describe the relationships between the endogenous variables, η .
- η is an $m \times 1$ vector of latent dependent or endogenous variables.
- Γ is an $m \times n$ matrix of regression coefficients that describe the relationships between the latent independent or exogenous variables, ξ , and the latent dependent or endogenous variables, η .
- ξ is a $n \times 1$ vector of latent independent, or exogenous variables.
- ζ is an $m \times 1$ vector of random disturbances in the structural equations, indicating that the endogenous variables are not perfectly predicted by the structural equations.

In this general form the structural model permits all exogenous variables to influence all endogenous variables, and all endogenous to be related to each other, rendering it liable to problems in identification (see section 7.1.4 on identification). Thus, in practice, to obtain an identified model, it is necessary to set some of the elements of the parameter matrices, Γ and B , to zero.

There are two types of structural models: recursive and nonrecursive. Recursive models have two defining characteristics (Carmines, 1986). First, the causal paths among the latent

variables run in only one direction and no feedback loop exists, and second, the disturbances (ζ) in recursive models are uncorrelated. These constraints mean that no reciprocal effects are possible in recursive models. In nonrecursive models, reciprocal relationships may exist among the latent variables and disturbances may be correlated.

A number of assumptions need to be met for the general LISREL model. These include:

1. Variables are assumed to be measured as deviations from their means.
2. Common and unique factors are uncorrelated. In other words, ϵ is uncorrelated with η ; δ is uncorrelated with ξ ; ζ is uncorrelated with η .
3. Unique factors are uncorrelated across equations. In other words, δ , ϵ , and ζ are mutually uncorrelated.
4. Latent variables and errors are uncorrelated. In other words, ξ and η are uncorrelated with the disturbance terms (ϵ , δ , ζ).
5. None of the structural equations is redundant.

7.1.2 Other parameter matrices of LISREL

So far, four matrices (B , Γ , Λ_x , Λ_y) have been introduced. Four other matrices of coefficients to be estimated are required to complete the specification of the general model. The following matrices are permitted to be non-zero:

1. the matrix, Φ , which is an $n \times n$ matrix representing the variances and covariances of the unobserved latent variables ξ .
2. the matrix, Θ_ϵ , which is a $p \times p$ matrix representing the variances and covariances of the measurement errors ϵ for the measurement model for the y variables.
3. the matrix, Θ_δ , which is a $q \times q$ matrix representing the variances and covariances of the measurement errors δ for the measurement model for the x variables.

4. the matrix, Ψ , which is an $m \times m$ matrix representing the variances and covariances of the disturbance terms ζ .

As can be noted three of the four remaining matrices (Θ_ϵ , Θ_δ , Ψ) are variance-covariance matrices of the error or disturbance variables. When the structural and measurement equations are estimated, the error terms become part of these matrices. For example, the variances of ϵ_i and δ_i are actually estimated as the diagonal values of the Θ_ϵ and Θ_δ matrices, respectively. Likewise, the variances of the structural equation disturbance terms (ζ_i) are the diagonals of the Ψ matrix. Several useful observations can be made of the abovementioned four matrices. First, they are square matrices. Second, the diagonal elements should always be positive since they represent variances and variances must be positive. Third, the matrices are symmetric, so only the lower (or upper) triangular elements of these matrices need be written. Fourth, if the variables are standardized, these matrices become correlation matrices.

7.1.3 Types of model parameters

A wide range of structural equation models can be specified in terms of the parameter matrices Λ_x , Λ_y , Γ , B , Φ , Ψ , Θ_ϵ , and Θ_δ , by imposing various constraints on the parameter matrices to satisfy a given model specification. There are three types of model parameters: fixed, constrained, and free parameters. Fixed parameters are assigned given values, so they are not estimated as part of the model. The absence of causal paths (represented by zero values for some of the β and γ parameters) and assignment of units of measurements to latent variables (represented by λ 's set equal to 1) are typical examples of fixed parameters. Constrained parameters are parameters whose values are unknown but which are specified to be equal to other model parameters. For example, theory may suggest that two independent variables (ξ_1 and ξ_2) have the same impact on a dependent variable (e.g., η_3) which could be incorporated in the model by specifying $\gamma_{31} = \gamma_{32}$ (elements of the Γ matrix). Note that with constrained parameters, it is only necessary to estimate one of them. Finally, free parameters are those which have unknown values and which need to be estimated by the programme. The basis for restricting parameters to be fixed or constrained or for allowing them to be free depends on prior theoretical knowledge, logical criteria, empirical evidence, or methodological design considerations (Bagozzi, 1980).

Table 1. Summary of variables, matrices, and terms used in the analysis of covariance structures

		Symbol	Definition
I.	Observed variables	y	observed endogenous variable/ indicators of unobserved endogenous variable
		x	observed exogenous variable/ indicators of unobserved exogenous variable
II.	Unobserved variables	η (eta)	unobserved endogenous variable
		ξ (ksi)	unobserved exogenous variable
		ζ (zeta)	disturbances in structural equation
		ϵ (epsilon)	errors of measurement of endogenous variable
		δ (delta)	errors of measurement of exogenous variable
III.	Coefficient Matrices	B (beta)	coefficient matrix of the relationships among the unobserved endogenous variables with elements β representing causal parameters among endogenous dimensions
		Γ (gamma)	coefficient matrix of causal effects of the unobserved exogenous variables on the unobserved endogenous variables with elements γ representing causal parameters from exogenous to endogenous dimensions
		Λ_y (lambda y)	factor matrix of y on η with elements denoted λ^y
		Λ_x (lambda x)	factor matrix of x on ξ with elements denoted λ^x
		Φ (phi)	variance-covariance matrix of unobserved exogenous variables with elements denoted ϕ
		Ψ (psi)	variance-covariance matrix of structural disturbances with elements denoted ψ ; diagonal elements are the amount of error variance in the endogenous dimensions
		Θ_ϵ (theta epsilon)	variance-covariance matrix of errors of measurement of observed endogenous variables with elements denoted θ_ϵ ; diagonal elements are the error variances of the y indicators
		Θ_δ (theta delta)	variance-covariance matrix of errors of measurement of observed exogenous variables with elements denoted θ_δ ; diagonal elements are error variances of the x indicators

Note: Lower case Greek letters indicate the parameters of the model and denote that the system is specified in equation form. Upper case Greek letters denote the corresponding specification of the equations in matrix algebra format. The general model is usually discussed in the literature in terms of its matrix format.

Table 1 presents a summary of the variables and matrices used in SEM. As was seen in section 7.1.1.1, any causal model can be described in matrix equation form. A causal model

can also be expressed graphically. To illustrate, a graphic representation (using standard Greek notation) of a simple model is presented in Figure 8. The labelled large circles represent the latent (unmeasured) factors and the rectangles represent observed (or measured) variables. The symbols used to represent latent constructs (ξ , η), factor loadings (λ), measurement error (δ , ϵ), path coefficients (γ , β), and structural disturbances (ζ) reflect the standard notation used in the LISREL literature. The single headed arrows from the large circle to the rectangles indicate that the observed variables are theorized to be generated by the latent construct. The variables (δ , ϵ , ζ) beside the detached arrows represent the disturbance terms. The δ , ϵ error variables specify the cumulative effects of excluded variables and purely random measurement errors on the observed x and y variables, respectively. The variances of the measurement error variables (ϵ , δ) represent the unique portion of residual variance (uniqueness) that is not accounted for by the common factor. The variance of the disturbance term, ζ , represents unexplained variance in the prediction of the endogenous constructs by the conceptual level variables. The obtained parameter values of the measurement model can be interpreted as factor loadings, and the path coefficients of the structural model as being analogous to the regression weights known from simple path analysis.

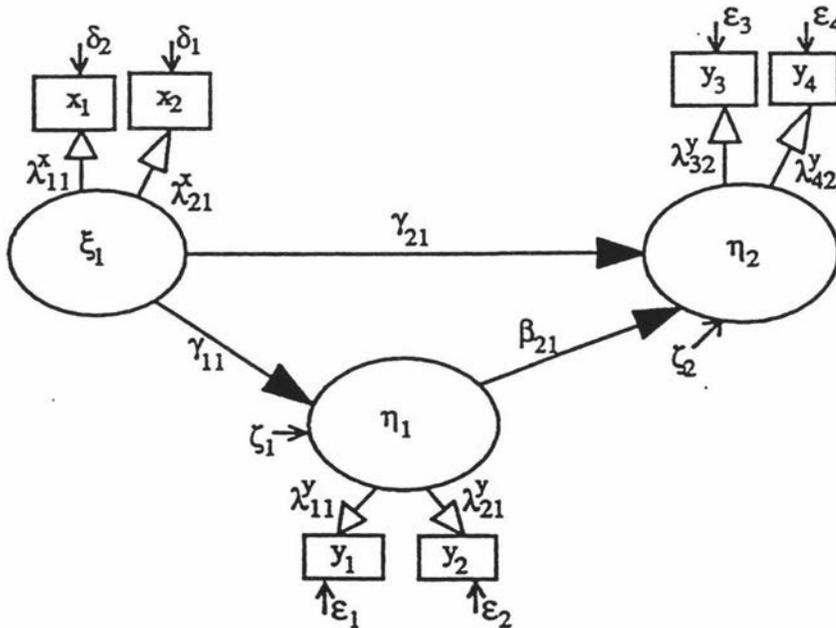


Figure 8. Schematic representation of the general LISREL model.

While the model in Figure 8 is clearly very simple, the general principles of formulating more complex models follow a similar logic. More specifically, a theoretical model is expressed as a path diagram of relationships between variables of interest, and this path diagram is represented by a system of simultaneous linear equations.

7.1.4 Model identification

Once the model is specified, it is important to determine whether it is possible to obtain unique solutions for the estimated parameters, an issue termed *identification*. The identification status of any covariance structure model is determined by the correspondence between the amount of information provided by the empirical data (e.g., variances and covariances of the observable variables) and the total number of model parameters that need to be estimated. Problems with identification are often the result of a problem in the theory or the data.

A unique solution for all parameters exists, if a model is identified. In fact, alternative solutions for parameter estimates exist in overidentified models but for a given fitting procedure these will converge to a unique set of parameter values. When a unique solution for a set of parameters does not exist, they are said to be not identified or underidentified and the model is indeterminate. In such cases the model fitting procedure may not produce a solution or may converge to an improper solution. A system of equations can only be solved if the number of equations is at least equal to the number of unknowns. Hence, to obtain a unique solution of the parameters in a structural equation model, it is necessary to have at least as many unique (i.e., linearly independent) equations relating the variances-covariances of the observed variables to the model parameters as there are parameters. Consequently, "a necessary condition for identification is that the number of independent parameters being estimated is less than or equal to the number of non-redundant elements of S , the sample matrix of covariances among the observed variables" (Long, 1983b, p. 66). Hence, if t parameters are to be estimated, the *minimum* condition for identification is: $t \leq s$ where $s = \frac{1}{2}(p+q)(p+q+1)$, p = number of y variables, and q = number of x variables. If $t > s$, the model is underidentified, resulting in an infinite number of values of the parameters that would satisfy the equations (i.e., there are more unknowns than equations).

An exactly identified model (where $t = s$) is one that produces only one estimate for each parameter because there are as many equations as there are unknowns. In other words, the number of parameters to be estimated is equivalent to the number of observed variances and

covariances. Although such a model can be estimated (i.e., parameters have a unique solution), it is not amenable to falsification through goodness-of-fit methods since there is no information left that can be used to test the model (i.e., there are zero degrees of freedom).

A desirable feature for models is that they be overidentified ($t < s$), where it is possible to obtain several estimates of the same parameter, since there are more equations than unknowns. This means that the number of observed variances and covariances is larger than the number of parameters to be estimated. It should be emphasized that it is not always the case that models in which there are more observed variances and covariances than model parameters to be estimated will be fully identified. In other words, the condition $t \leq s$ is *necessary* for identification of a LISREL model, but is not a *sufficient* condition.

Therefore, it is imperative to conduct further tests to establish that a given model is identified. One way would be to establish identification through algebraic manipulations of the model's variance-covariance equations, so that each of the model parameters may be solved in terms of the variances and covariances of the observed variables. Solving these algebraic equations by hand may prove to be difficult, particularly if the model is complex. This is because the number of parameters to be estimated is often too large for manual checks and the possibility of making errors may arise. Fortunately, empirical tests of model identification are now routinely included in most computer packages used for conducting SEM, thus making it possible to examine the identification status of the model during estimation. The LISREL programme has a number of diagnostics that provide warnings about identification problems. For example, if any of the parameter matrices in Table 1 cannot be inverted (i.e., are not positive-definite) an appropriate message is printed and the same happens if a particular parameter is found to be underidentified. Other indications of identification problems may be obtained from the LISREL output. These include: (a) large standard errors for the estimated parameter, (b) negative error variances, (c) standardized coefficients exceeding 1.0, and/or (d) correlations between coefficient estimates exceeding about $\pm .9$.

7.1.5 Parameter estimation

Estimation involves finding values for the eight parameter or coefficient matrices in Table 1 (i.e., Λ_x , Λ_y , Γ , B , Φ , Ψ , Θ_e , and Θ_b) that are consistent with the constraints imposed on the model (fixed, free, and constrained parameters) and "generate an estimated covariance matrix Σ that is close as possible to the sample covariance matrix S " (Long, 1983b, p. 44). A function that measures how closely Σ is to S is known as a fitting function and designated as

$F(S; \Sigma)$. Different estimation procedures have different fitting functions and the LISREL 8 programme offers a variety of estimation procedures, three of which are commonly used: unweighted least squares (ULS), generalized least squares (GLS), and maximum likelihood (ML). The following discussion is confined to the maximum likelihood (ML) approach as it is the most widely used estimation method in SEM and LISREL modelling (Huba & Harlow, 1987).

ML is an iterative procedure whereby final parameter estimates are obtained through numerical algorithms which successively improve initial estimates (starting values); the latter supplied by the user or automatically estimated by the LISREL programme. The values of the parameters are estimated by maximising the likelihood of observing a given matrix Σ conditional on a set of parameter values and the assumption of multivariate normality. Although parameter estimates based on the maximum likelihood approach are usually robust to violation of the multivariate normality assumption, this violation may influence the standard error estimates and model fit (Jöreskog & Sörbom, 1989).

7.1.6 Model evaluation

One of the most significant aspects of the estimation of any model is an assessment of the quality of fit of the model to the observed data. The mere fact that a model fits the data does not imply that the model is indeed the "correct" or "best" causal model (Jöreskog & Sörbom, 1993, p. 114). It is always possible to propose various models which will fit a particular set of data.

A confirmatory (or theory testing) approach is usually taken by researchers to evaluate structural equation models. Such an approach entails testing the model via some goodness-of-fit measures to confirm that the proposed model accounts for the observed data. Assessment of fit is determined by examining the extent to which the parameters of a given model reproduce the observed variance-covariance matrix of the observed measures. If the fit is satisfactory, the model is accepted; if the fit is unsatisfactory, the model is rejected. Such an approach may be misleading since it does not consider how effective alternative models are in explaining the same data (e.g., Sternthal, Tybout, & Calder, 1987). This problem has been recognized by a number of users of structural equation modelling (Bentler, 1990; Cliff, 1983; Jöreskog, 1993; Yi & Nassen, 1992).

It must be noted that rarely is the investigator's causal theory so precise that he or she can specify the exact structure of the causal model. Horwood (1987) suggests that the process of model construction should be inspected within the context of a hypothesis or theory testing framework. Hence, it may be appropriate for the investigator to specify a number of competing representations, to compare the adequacy of these models in explaining the observed data, and to be able to distinguish ways to respecify models to overcome their imperfections. This means that the evaluation of a model may require two quite distinct processes. First, it will be necessary to assess the goodness-of-fit of the model. Second, it may be necessary to compare a number of competing models with the proposed model so as to find a model among a set of alternatives which provides the best approximation to the observed data.

7.1.6.1 Goodness-of-fit indices

There are a variety of indices of fit which can be obtained for a given model. These indices may be used to determine the adequacy of fit of a single model or to compare the relative fit of a number of alternative competing models of the data. In general, no single index has been endorsed the "best index" by the majority of researchers (Gerbing & Anderson, 1993), so adequacy of fit is usually assessed on the basis of observation of several indices together. A variety of different goodness-of-fit indices have been proposed in the literature (see Bollen, 1989; Browne & Cudeck, 1993), and a brief summary of the more commonly used indices will be given below. It may be useful to classify the information for model evaluation and assessment of fit into two groups: measures of overall fit and detailed assessment of fit.

(i) *Measures of overall fit*

1. *Chi-square index*

The LISREL programme estimates values for the nonzero parameters specified in the model; the maximum likelihood criterion is used to yield parameter estimates that provide the best fit to the observed covariances among measured variables. Even the best-fitting solution of the parameters specified in a model can amount to a poor fit with the observed data. In that event the hypothesized model is rejected as an inaccurate description of the true relations in the population that gave rise to the data. Statistically, the decision regarding the general goodness-of-fit of the model is made on the basis of the log likelihood chi-square statistic,

which is analogous to the conventional chi-square statistic. The chi-square statistic provides a direct test of the hypothesis that the model fits perfectly in the population, not a test of a competing null hypothesis. That is, the null hypothesis initially tested is that a sample covariance matrix is obtained from a population that has the proposed causal structure. Thus, the significance test works paradoxically here: If a chi-square is large in relation to its degrees of freedom (i.e., statistically significant), then the model is rejected. If chi-square is statistically nonsignificant, then the model may be retained as a plausible description of the relations that exist in the population. The degrees for freedom for chi-square are $\{ \frac{1}{2} [(p + q) (p + q + 1)] - t \}$ where $(p + q)$ is the number of observed variables analyzed and t is the total number of independent parameters estimated.

Jöreskog and Sörbom (1993) have suggested that the chi-square statistic should be regarded as a "badness of fit measure" in the sense that large chi-square values correspond to bad fit and small chi-square values to good fit. For the chi-square test to be valid, it is assumed that the sample size is sufficiently large. Boomsa (1987) recommended that a sample size of at least 200 be used in confirmatory factor analytic studies. The p value reported by the programme is the probability level of chi-square (χ^2), that is, the probability of obtaining a chi-square value larger than the value actually obtained, given that the model is correct.

The conventional probability level for the chi-square test in covariance structure analysis appears to have been set at $\geq .05$. This value refers to the probability of obtaining an equal or worse fit (as measured by chi-square) than what was actually obtained, given that the null hypothesis is true. Obviously, in probability terms, this does not constitute strong confirmation of the research hypothesis. Since accepting the null hypothesis amounts to accepting one's theory, it has been suggested that levels of .1 or .2 be exceeded before nonsignificance is confirmed (Fornell, 1983; Hayduk, 1987). The conservative criterion of .2 was adopted in the present investigation, considering its modest sample size.

The chi-square index has been criticised on several grounds (see Fornell & Larcker, 1981). For instance, the chi-square statistic computed via LISREL is very sensitive to departures from multivariate normality of the observed variables and increases as a direct function of sample size. Therefore, given departures from multivariate normality or large sample sizes, a proposed causal model can easily not statistically fit the data even though the discrepancy between the sample covariance matrix and that reproduced by the parameter estimates of the

causal model may be insignificant from a practical point of view. Hence, it is considered important to supplement the chi-square index with other measures of goodness-of-fit.

2. Goodness-of-fit index (GFI)

This is a measure of the relative amount of observed variances and covariances accounted for by the model. In other words, it assesses how well the covariances predicted from the parameter estimates reproduce the sample covariances. A value in excess of .90 is generally thought to be indicative of a reasonable fit to the data and indicates that the model accounts for most of the joint variances and covariances among observed variables in the model.

3. Normed fit index (NFI)

Bentler and Bonett's (1980) normed fit index (NFI) assesses fit by the degree to which the model accounts for the sample covariances relative to a more restricted nested model – usually the null model in which all indicators are specified as uncorrelated. This index ranges from 0 (a fit that is no better than the null model) to 1 (perfect fit).

4. Parsimonious normed fit index (PNFI)

Mulaik et al. (1989) argued that goodness-of-fit alone is insufficient to determine the adequacy of a model because it capitalizes on the number of parameters freed for estimation. As a result, a model may fit a set of data quite well but not be parsimonious. Hence, a parsimonious normed fit index (PNFI) was developed by Mulaik et al., to gauge model parsimony. The PNFI takes into account the ratio of degrees of freedom to the NFI. This leads to a significant reduction (of about half) in PNFI values to NFI values so that a high NFI is associated with a moderate PNFI. The PNFI value is used as a standard comparison of a series of models, whereas NFI and GFI indices are more model specific (see Mulaik et al., 1989, for a review). Higher values of PNFI are better, and its principal use is for the comparison of models with differing degrees of freedom. It is used to compare alternative models, and there are no recommended levels of acceptable fit (Hair, Anderson, Tatham, & Black, 1992). However, when comparing between models, differences of .06 to .09 are proposed to be indicative of substantial model differences (Williams & Hazer, 1986).

5. Root mean square residual (RMR)

Another measure that has been used frequently in the research literature is the root mean square residual (RMR). It is a measure of the average magnitude of the residual correlations, that is, an index of the extent to which the "fitted" matrix fails to reproduce the original

correlation matrix and thus does not fully account for the observed data. When a covariance matrix is analyzed, RMR is not readily interpretable, because its magnitude is affected by the units of the measures of the indicators. Under such circumstances, the standardized RMR can be used (as was done in the present study) which is provided by LISREL 8 as part of its goodness-of-fit output. Models with a standardized RMR value of less than .1 can be considered to be of good fit.

(ii) *Detailed measures of fit*

In addition to overall measures of fit, much detailed information about fit can also be used to further evaluate the model. While the measures of fit described above give some indication of the overall quality of the fit of a model to the data, they give no indication as to how a model may be respecified to improve fit. The detailed measures of fit provide information by which specification errors or possible sources of lack of fit may be detected in the model.

1. *Standardized residuals*

One detailed measure of fit is the distribution of standardized residuals, which is output in the form of a graphic quantile - quantile (Q-Q) plot, thereby providing a visual representation of overall model fit. Standardized residuals are estimates of the number of standard deviations the observed residuals are away from the zero residuals that would exist if the model were a perfectly fitting one (Hayduk, 1987). Values greater than 2.58 for any element usually provide a clue as to possible model misspecifications (Jöreskog & Sörbom, 1989). Overall, for a well fitting model, if the errors are random, 95% of the residuals would fall within 2 standard deviations of zero (Hayduk, 1987). Residuals that follow the dotted line rising at a 45-degree angle in the Q-Q plot are indicative of a well fitting model, with steeper slopes representing better fits. Those that deviate widely from the 45-degree line in a nonlinear fashion indicate that the model is in some way misspecified.

2. *Modification index*

An insight into which parameters to add to the model can be gained by looking at the modification indices. Modification indices are computed for each fixed or constrained parameter in the model (i.e., a value set at zero implying the absence of that parameter in the model or parameters constrained to be equal). The value of the index indicates the minimum reduction in the chi-square statistic that could be obtained if a parameter was freed for estimation, the equivalent of adding a path to the model. A modification index of about 9.0

or more would be considered worthy of attention (Fassinger, 1987), although most researchers use a value of 5.0 as recommended by Jöreskog and Sörbom (1989). It is important that theoretical considerations guide the relaxation of a parameter because "even if the model initially suggested by substantive theory is rejected, there are generally some parameters that are definitely required on the basis of past research, and some parameters that make no sense to include" (Long, 1983b, pp. 76-77). An excellent discussion on model modification is contained in Bollen (1989).

3. Correlation of parameter estimates

The correlation of parameter estimates provides another measure of detailed goodness-of-fit. Correlations between coefficient estimates exceeding about $\pm .9$ may signify identification problems or collinearity problems (Hayduk, 1987).

4. Squared multiple correlation

Although the squared multiple correlation provided for each equation in the model is considered by some as a detailed measure of fit (Byrne, 1989a), L. J. Horwood (personal communication, March 27, 1995) considers it to be a measure of explanatory power which does not necessarily have anything to do with fit. For instance, it is possible to have a large R^2 but a chi-square that indicates an unacceptable level of fit (Long, 1983b). Nonetheless, this measure can provide an assessment of the acceptability of the model as an adequate theoretical representation. In other words, this measure helps establish whether the relationships in the model are poorly determined.

For the measured variables, the squared multiple correlation is an indication of the proportion of variance accounted for by the latent variable and shows the extent to which the individual measured variables are free from measurement error. Therefore, R^2 provides an indication of the reliability of each observed measure with respect to its underlying latent construct; values close to 1.00 represent good construct validity/reliability. The coefficient of determination for the x and y variables indicates how well the observed variables as a group serve as measures for the latent exogenous and endogenous variables, respectively. For the latent variables in the structural model, the squared multiple correlation is an indication of the amount of variance in each endogenous latent variable accounted for by the independent latent variables in the model. The coefficient of determination shows the overall strength of the structural relationships taken together.

5. *Parameter values and standard errors*

The quality of the model may also be assessed by looking at the parameter values and the standard errors. The parameter values can be examined for unreasonable values. For instance, negative variances of parameters or correlations greater than 1.0 in magnitude suggest misspecification errors. Another indication of poor quality of the model is indicated by standard errors which have abnormally large values. This may be an indication that the model is not identified.

To conclude, it is important to note that there is no single measure which provides a definitive answer to the problem of assessing model fit. Often, it may therefore be better to base the assessment of model fit on the use of a number of alternative indices and measures of fit.

7.1.6.2 **Comparison of nested models**

The previous section has focused on examining the fit of a single specified model. This section considers the alternative models situation, in which the researcher wishes to compare the fit of a number of competing models for the same data set. The models can range from ones with extremely different structures to those containing minor differences in one or two parameters.

Researchers agree on the importance of nested model comparisons in structural equation modelling (e.g., Bentler, 1980). A model presented with no consideration of competing models is difficult to evaluate. In building a series of nested models, one can use incremental-difference statistical tests to evaluate the model. This type of test evaluates the contribution of additional elements over the simpler model and forces the inclusion of the principle of parsimony in the model evaluation process.

Consider the case when the models are nested so that they can be ordered in some sort of hierarchy of complexity by increasing the number of parameters (decreasing degrees of freedom). Let M_1, M_2, \dots, M_k , be a sequence of nested models and let d_1, d_2, \dots, d_k represent the degrees of freedom for each model respectively. M_1 is the most parsimonious model, that is, the one with fewest parameters (most degrees of freedom). M_k is the most complex model, that is, the one with most parameters (fewest degrees of freedom). If the models are nested, each model is a special case of the preceding model; that is, M_i is obtained from M_{i-1} by putting restrictions on the parameters of M_{i-1} .

It is assumed that model M_1 has been tested by a valid chi-square test and found to fit the data well and to be interpretable in a meaningful way. The nested models can be tested sequentially by testing M_2 against M_1 , M_3 against M_2 , and so on. Chi-square difference tests are commonly used for comparing the fit of nested models. The difference between the two chi-squares is also distributed as a χ^2 with degrees of freedom equal to the difference between the degrees of freedom for the two models. The only requirement is that the number of constructs and indicators remain the same, such that the null model is the same for both models; that is, they are nested models. The effect of adding or deleting one or more causal relationships can also be tested in this manner by making comparisons between models with and without the relationships.

Consider estimating two arbitrary models, M_1 and M_e , from the above sequence, one which is nested within the other in that it can be created from the other model by freeing additional paths to be estimated. The difference between the chi-square goodness-of-fit statistics for the two models can be interpreted as a chi-square test of the improvement in fit of the more complex model M_e over the simpler model M_1 , with a significant chi-square difference indicating a substantial improvement. When the difference in goodness-of-fit between two models is found to be nonsignificant, the model that provides the more parsimonious solution is accepted.

One limitation of the chi-square difference statistic is that it is still subject to the same liabilities of interpretation as the general chi-square goodness-of-fit statistic. Hence, it is customary to adopt a relatively free interpretation of this statistic: if the chi-square difference is large relative to the difference in degrees of freedom between the two models, then the changes can be said to represent a real improvement. On the other hand, if the improvement is small relative to degrees of freedom then it is likely that the improvement in fit has been gained by capitalising on chance variation in the data.

Another approach is to compare the models on the basis of some criteria that take parsimony as well as fit into account. This is especially useful for non-nested models. In this case the chi-square difference test is no longer appropriate and the researcher must rely on the parsimonious normed fit index described earlier. Chi-square comparisons of nested models can also be supplemented by other indices of fit which are independent of sample size, such as the NFI described previously.

7.2 Multisample model

The LISREL methodology has been frequently described and applied in the single sample situation as described above. In this section it will be demonstrated how the general LISREL model can be extended to an analysis across samples, which is otherwise known as multisample analysis. In this circumstance, the question "Does the same causal model apply to two or more groups or samples?" can be asked. For example, in the present study the question "Is the model of reading comprehension similar for reading disabled and normal achieving readers?" can be asked. Such model comparisons across populations are not prevalent in the educational and psychological literature.

In the general LISREL model discussed earlier, the elements of Σ can be constructed from the eight matrices, Λ_x , Λ_y , Γ , B , Φ , Ψ , Θ_e , and Θ_δ . As for the single sample situation, a population variance-covariance matrix Σ can be constructed from the eight matrices, $\Lambda_x^{[g]}$, $\Lambda_y^{[g]}$, $\Gamma^{[g]}$, $B^{[g]}$, $\Phi^{[g]}$, $\Psi^{[g]}$, $\Theta_e^{[g]}$, and $\Theta_\delta^{[g]}$, for the multisample situation, where 'g' denotes group. Elements of these eight matrices may consist of fixed, constrained, and free parameters, as in the single-sample case.

Usually in multisample analyses, equality constraints (e.g., equality constraints for factor loadings) are placed between the samples. For example, it is possible to test whether differences exist between elements of these matrices for one group compared to another. One may be interested whether $\Lambda_y^{[1]} = \Lambda_y^{[2]} = \dots = \Lambda_y^{[k]}$ or $B^{[1]} = B^{[2]} = \dots = B^{[k]}$, where the superscript represents a distinct matrix for each group or sample, with k equal to the number of groups. The researcher may impose various restrictions on the similarity of the parameters across groups. For example, in the most restrictive situation one might require that all elements in each of these matrices be equal across groups. Less restrictive would be the requirement that selected elements in these matrices be equal for all groups while other elements are free to vary from one group to another. The least restrictive model would allow all the parameters to vary across the samples. In such a situation, each sample can be analyzed separately. The procedure also allows for some causal paths to be fixed at zero in one sample, while in the other sample the path is estimated. One assumption of the multisample analysis is that the samples are independently drawn from their respective populations and the subgroups are mutually exclusive.

The measurement and structural equations for each group can be expressed as follows:

$$X^{[g]} = \Lambda_x^{[g]} \xi^{[g]} + \delta^{[g]}$$

$$Y^{[g]} = \Lambda_y^{[g]} \eta^{[g]} + \epsilon^{[g]}$$

$$\eta^{[g]} = B^{[g]} \eta^{[g]} + \Gamma^{[g]} \xi^{[g]} + \zeta^{[g]}$$

where $g = 1, 2, \dots, k$ with k equal to the number of groups.

Identification and estimation of the parameters are essentially similar to the single-sample case, except they are at a more complex level. A chi-square test is still available to test the adequacy of the model to fit the data (i.e., to fit the two or more observed variance-covariance matrices). A chi-square test is calculated as before with degrees of freedom equal to $\{ \frac{1}{2} k [(p + q)(p + q + 1)] - t \}$ where t is the total number of estimated coefficients, $(p + q)$ is the total number of observed indicators (p endogenous, q exogenous), and k is the number of groups.

The parameter estimates are neither averages across the groups nor based upon separate calculations for each group. Rather, the estimation procedure simultaneously calculates values for the elements that minimize the fitting function. In other words, the estimation procedure calculates the elements of the matrices, Λ_x , Λ_y , Γ , B , Φ , Ψ , Θ_e , and Θ_b , that will make the estimated Σ 's (i.e., $\Sigma^1, \Sigma^2, \dots, \Sigma^k$) as nearly as identical to their respective observed S 's (i.e., S^1, S^2, \dots, S^k) as they can possibly be, given the fixed, constrained, and free elements of the model as specified by the researcher.

7.3 Justification for LISREL-type models

The above highly abstract account of LISREL does not provide obvious reasons why one would choose to analyze data using the LISREL approach in preference to less complex means of data analysis. There are three major justifications that may be proposed for choosing a LISREL approach.

First, the majority of social science observations contain substantial errors of measurement, arising mainly from imperfections in the measurement instruments (Horwood, 1987). The effects of measurement error and the repercussions of failure to account for such errors are widely recognised and documented (e.g., Aigner & Goldberger, 1977; Blalock, 1985). In particular, these measurement errors, if ignored, can seriously bias the estimated parameters

of the structural model. Thus, an important and realistic assumption underlying the hypothesized models in this dissertation is that each of the measures is likely to be measured with error. Parkes (1987) suggested that psychometric scales are likely to contain sources of error variance which would attenuate the correlations among the latent constructs. These sources of error variation are likely to arise from both random errors of measurement and the presence of method, response, or situation specific variance in the scores. In psychometric scales, the self-report scale scores and behavioural (interview) scores obtained on psychological measures such as those used in the present study are seen as the sum of two components. The first component represents the individual's true score on the characteristic of interest, and the second component is due to measurement error. True scores reflect real characteristics of the individual but they cannot be directly assessed, since observed scores are always to some extent contaminated by measurement error. Furthermore, because measurement error has an attenuating effect on measures of association, the magnitudes of the correlations among true scores tend to be underestimated by observed score correlations (cf. Huba & Bentler, 1982; Huba, Wingard, & Bentler, 1981; Stacy, Widaman, Hays, & DiMatteo, 1985). LISREL overcomes these problems of unreliability in self-report measures.

A second justification concerns the distinction between observed and non-observed or latent variables. Many of the variables the social scientist may wish to examine are not in fact directly observable or measurable; rather they are hypothetical constructs or latent variables which depict the dispositional characteristics of observations. The properties of the latent variable must be deduced on the basis of observations on fallible indicator variables purportedly representing the variables. These fallible indicator variables can be observed and may be assumed to measure some aspects of the underlying latent variable with varying degrees of accuracy.

The distinction between observed and latent variables raises the important issue of establishing rules of correspondence between the latent variables and the observed variables which supposedly represent these variables. In this respect, LISREL makes a major contribution by drawing a clear distinction between what the researchers have observed (i.e., fallible indicator variables), and what they would like to observe (i.e., errorless measures of latent constructs). In addition, LISREL provides a technology to estimate the relationships between these variables subject to a given model specification. The distinction between observed and latent variables also has important implications for the estimation of the coefficients of a causal model. For the most part, social scientists have tended to resolve the distinction between

latent and observed variables by assuming that the structural relationships between the observed variables will in some way reflect those that exist between the latent variables which are represented by the observed variables. This premise has been proven to be suspect (Costner, 1985). However, LISREL provides a subsidiary theory (i.e., the measurement model component) that links the observed variables to the latent variables they represent, thus enabling the estimation of the structural relationships between latent variables.

The third justification is that model formulation can be made explicit. To impose a given causal interpretation on a set of data, the LISREL model requires the researcher to make explicit all of the assumptions that are required. In contrast, in most other data analytic procedures, these assumptions are often left implicit, opening the way for them to be ignored or overlooked. In turn, this lack of an explicit model can lead to faulty inferences.

7.4 LISREL formulation of reading performance models

While the above discussion has provided a basic introduction to the techniques and rationale for using methods of SEM, the purpose of this section is to demonstrate how the models developed in Chapter 6 (see section 6.2), can be represented as LISREL models. In the first part of this section, the models will be presented based on the assumption that there is only one group. Then the model will be elaborated to take account of correlated errors of measurement, and the two-group situation. This development is limited to a specification for Model 1. Since Models 2 through 4 represent simple extensions of this model to include additional observed and latent variables the LISREL specifications for these models are not presented here. However, detailed LISREL specifications for all four models as they are finally reported in the Results (Chapter 9) can be found in Appendix G.

7.4.1 Specification of Model 1 for the single sample situation

The structure of the relationships between intellectual ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance is shown diagrammatically in Figure 9. The model shows that there are four endogenous variables and one exogenous variable. Let ξ_1 denote the ability factor, η_1 denote the attributional style factor, η_2 denote the self-efficacy factor, η_3 denote the metacognitive knowledge factor, and η_4 denote the performance factor. Assume that self-efficacy, metacognitive knowledge, and reading performance have each been

measured using two indicators whilst attributional style has been measured by five indicators and intellectual ability by one indicator. Thus, there are a total of 12 (1 exogenous and 11 endogenous) measures or indicators to be analyzed. The measure of intellectual ability is based on the WISC-R (x_1); measures of attributional style are based on controllability (y_1), stability (y_2), internality (y_3) dimensions of success attributions, and stability (y_4), controllability (y_5) dimensions of failure attributions; measures of self-efficacy are based on general (y_6) and task specific (y_7) measures; measures of metacognitive knowledge are based on two questionnaires (y_8, y_9); and measures of reading performance are based on a standardized test (y_{10}) and classroom performance (y_{11}). Further, assume that all measures are measured as continuous variables. The relationships between the 12 measures are thus described by the matrix of variances and covariances between these measures. It is convenient to regard each variable as being measured as deviations from its mean so that mean values can be excluded in the equations.

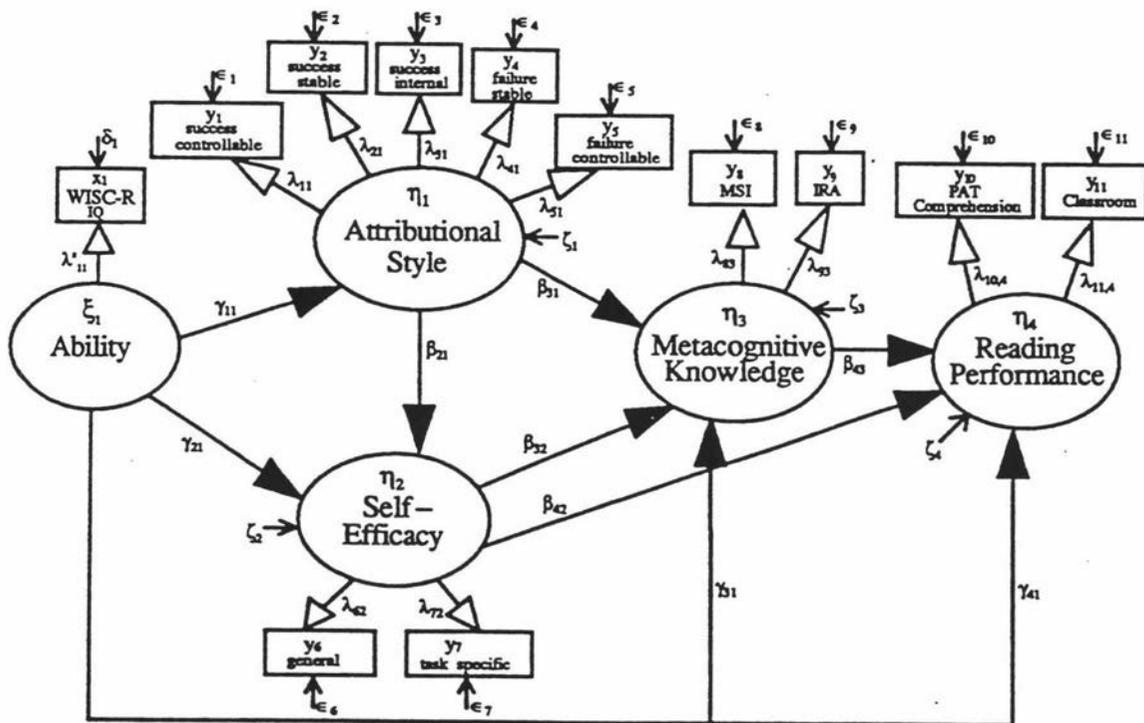


Figure 9. Hypothesized relationships between cognitive, metacognitive, and motivational variables in a single sample.

As discussed earlier, an important and realistic assumption underlying Model 1 is that each of the measures is likely to be measured with error. One way of overcoming problems of

unreliability in self-report measures involves the use of multiple indicators. As is seen above and in Figure 9, multiple indicators are used for each latent construct except for ability.

The variables and parameters of this model may be interpreted as follows:

1. The factors $\eta_1, \eta_2, \eta_3, \eta_4$, which represent attributional style, self-efficacy, metacognitive knowledge, and performance, respectively, are latent unobservable endogenous variables which are assumed to be measured without error and to have a mean of zero. Similarly, ξ , the latent exogenous ability variable, is assumed to be measured without error and to have a mean of zero.
2. The terms $\zeta_1, \zeta_2, \zeta_3$, and ζ_4 are disturbance terms which represent an aggregate of all known and unknown influences on young adolescents' attributional style, efficacy, metacognitive knowledge, and reading performance, respectively, which are not measured in the model. Ψ is assumed to be diagonal. When B is triangular (as is the case in this model) and Ψ is diagonal, the model is said to be recursive.
- 3a. Restricting elements of B and Γ to a value of zero indicates the absence of a causal relationship between the appropriate variables. Fixing the $(i, j)^{\text{th}}$ element of Γ to zero (γ_{ij}) implies that the latent exogenous variable ξ_j does not have a causal effect on the latent endogenous variable η_i . Similarly, fixing the $(i, j)^{\text{th}}$ element of B to zero (β_{ij}) implies that the absence of a relationship between the latent endogenous variables η_i and η_j . In the proposed model, a causal path is specified between the latent exogenous variable (ξ_1) and each of the latent endogenous variables ($\eta_1, \eta_2, \eta_3, \eta_4$) and these relationships are unidirectional (recursive). In the case of the relationships between the latent endogenous variables, the specification of $\beta_{41} = 0$ implies an absence of a causal path from the latent variable of attributional style to the latent variable of performance. As the model is recursive, B is triangular and if η_i affects η_j ($\beta_{ji} \neq 0$), η_j does not affect η_i ($\beta_{ij} = 0$). The diagonal elements of B are assumed to equal zero, since an endogenous variable does not cause itself.
- 3b. The parameter β_{21} is the regression coefficient that represents the association between the endogenous latent factors of attributional style (η_1) and self-efficacy (η_2). The parameters β_{31} and β_{32} are the regression coefficients that represent the association of the latent factor of metacognitive knowledge (η_3) with attributional style (η_1) and efficacy (η_2), respectively. The parameters β_{43} , and β_{42} are the regression coefficients that represent the association of the latent factor of performance (η_4) with metacognitive knowledge (η_3)

and efficacy (η_2), respectively. The interpretation of these coefficients is dependent on the scale of measurement used. In particular, the interpretation of these coefficients will vary depending on whether the indicator variables are measured in raw score form or in standardized form. If the variables are measured in raw score form, then β_{21} units represents the change in η_2 which would be associated with a unit change in η_1 , all other variables being held constant. If the variables are expressed in standardized form, in which they are scaled to have a mean of zero and variance of one, β_{ij} represents the standard deviation change in η_i which would be associated with a one standard deviation change in η_j , all other variables being held constant.

4. The parameters γ_{11} , γ_{21} , γ_{31} , and γ_{41} are the regression coefficients that represent the association between the latent exogenous factor of ability (ξ_1) and the endogenous factors of attributional style (η_1), efficacy (η_2), metacognitive knowledge (η_3), and performance (η_4), respectively. Interpretations of the coefficients as raw scores or standardized scores are similar to the β parameters. For example, in raw score form, γ_{21} indicates that a unit change in the exogenous variable ξ_1 results in a change in η_2 of γ_{21} units, holding all other variables constant. In standardized form, a standard deviation change in ξ_1 results in a γ_{ij} standard deviation change in η_i , all other variables being held constant.

Assessment of the relationships between ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance amounts to: (a) developing models of the relationships between the 12 observed, fallible measures of x_i , $y_i - y_{1i}$, and corresponding errorless latent variables representing the true (but non-observed) levels of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance, and (b) developing models of the relationships between the latent variables of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance.

7.4.1.1 Measurement model

The measurement model describing the relationships between the measures and the latent constructs can be expressed as a series of linear equations. The general form of these equations for the endogenous variables is:

$$y_i = \lambda_{ij} \eta_j + \epsilon_i$$

where η_j ($j = 1$ to 4) denotes the non-observed attributional style, efficacy, metacognitive knowledge, and performance factors; the coefficients λ_{ij} describe the regressions of the psychometric measures y_i ($i = 1$ to 11) on the latent attributional style, efficacy, metacognitive

knowledge, and performance factors, η_j ($j = 1$ to 4); and ϵ_i ($i = 1$ to 11) denotes random errors of measurement in the psychometric or self-report data.

As there is one exogenous variable, the measurement model for the exogenous variable reduces to:

$$x_1 = \lambda_{11} \xi_1 + \delta_1$$

Since model specification in LISREL is done by means of parameter matrices, the measurement and structural equations must be transformed into matrix form. Note that the mathematical specification of Model 1 is directly derived from its path diagram. This is reflected in the *subscripts* of the various path coefficients which correspond to the locations of row/column positions of the coefficients in the matrices. The absence of a path is indicated by entering a zero in the corresponding matrix location, while for path coefficients fixed a priori to a given value (e.g., the λ 's used to fix the units of measurement of the latent variables), the given values are entered directly in the relevant matrices (e.g., the 1's in the Λ_y and Λ_x matrices).

In LISREL notation the above measurement model may be expressed in matrix notation as follows:

$$\begin{array}{c}
 \mathbf{y} \\
 \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \end{bmatrix}
 \end{array}
 =
 \begin{array}{c}
 \Lambda_y \\
 \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ \lambda_{41} & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \lambda_{93} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda_{11,4} \end{bmatrix}
 \end{array}
 \begin{array}{c}
 \boldsymbol{\eta} \\
 \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}
 \end{array}
 +
 \begin{array}{c}
 \boldsymbol{\epsilon} \\
 \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \end{bmatrix}
 \end{array}$$

$$\begin{array}{c}
 \mathbf{x} \\
 \begin{bmatrix} x_1 \end{bmatrix}
 \end{array}
 =
 \begin{array}{c}
 \Lambda_x \\
 \begin{bmatrix} \lambda_{11} \end{bmatrix}
 \end{array}
 \begin{array}{c}
 \boldsymbol{\xi} \\
 \begin{bmatrix} \xi_1 \end{bmatrix}
 \end{array}
 +
 \begin{array}{c}
 \boldsymbol{\delta} \\
 \begin{bmatrix} \delta_1 \end{bmatrix}
 \end{array}$$

where $\lambda_{11} = 1$.

Θ_δ , the matrix representing the variances and covariances of the measurement errors δ_i , is given by:

$$\Theta_\delta = [\theta_{11}]$$

where θ_{11} is set to a prespecified value.

The variance-covariance matrix of the structural disturbances is given by:

$$\Psi = \begin{bmatrix} \psi_{11} & & & \\ 0 & \psi_{22} & & \\ 0 & 0 & \psi_{33} & \\ 0 & 0 & 0 & \psi_{44} \end{bmatrix}$$

As Θ_ϵ , Θ_δ , and Ψ are variance-covariance matrices, they are symmetric and therefore the elements in the upper diagonal, which are identical to those in the lower diagonal, are redundant. By convention, only the elements in the lower diagonal are shown.

7.4.1.4 Model assumptions

There are nine assumptions underlying this model:

1. The exogenous x_i and endogenous y_i indicator variables are scaled to have a mean of zero.
2. All relationships implied by the model are linear.
3. As the latent factors are not directly observable, it is necessary to establish a scale of measurement for these factors. Hence, the scale of measurement of the latent endogenous factors (η_j) is set by fixing the elements λ_{11} , λ_{72} , λ_{83} , and $\lambda_{10,4}$ of the Λ_y matrix to a pre-specified value. In addition, the measurement scale of the latent exogenous factor is set by fixing the parameter λ_{11} of the Λ_x matrix. By convention, this value is 1.0, indicating that the latent variables are scored in the same units as the observed indicators.
4. With only a single measure (indicator) of the latent construct of "ability", it was necessary to make assumptions about the values of the measurement parameters. Ability is assumed to be a fallible measure, and hence a perfect reliability of this measure cannot

be accepted. However, the LISREL programme cannot compute the residual of a single indicator (as it does when two or more variables are specified to load on an underlying construct). Consequently, the variance (θ_{11}) of the measurement error, δ_{11} , is fixed to a specified value. This value is calculated by using the following formula: (1 - reliability of indicator) x variance of indicator. In fixing this variance, one is fixing the portion of variance in an indicator that is thought to arise from sources other than the construct the indicator is supposed to indicate.

5. The variance-covariance matrix Θ_ϵ of the measurement errors of the endogenous variables ϵ_i is diagonal, indicating that the measurement errors are uncorrelated with each other. The error variances in the diagonal indicate that these variances are "free parameters" to be estimated by the programme. Fixing an error variance at a nonzero value implies that entities other than the underlying concept can influence the indicator and hence acknowledges some unreliability in the measurement of the latent factor.
6. The measurement error terms ϵ_i are uncorrelated with the latent endogenous ($\eta_1, \eta_2, \eta_3, \eta_4$) and exogenous factors (ξ_1).
7. The measurement error terms δ_i are uncorrelated with the latent endogenous ($\eta_1, \eta_2, \eta_3, \eta_4$) and exogenous factors (ξ_1).
8. The model assumes that the structural disturbance terms $\zeta_1, \zeta_2, \zeta_3$, and ζ_4 , are uncorrelated to exogenous and endogenous latent factors and to the measurement error terms ϵ_i and δ_i .
9. The model assumes that the disturbance terms $\zeta_1, \zeta_2, \zeta_3$, and ζ_4 , are mutually uncorrelated with each other. This means that the variance-covariance matrix Ψ of the structural disturbances is diagonal, indicating that the errors in equations are uncorrelated across each other.

The interpretation of the above model parameters is clearest in the case in which the model is completely standardized so that all observed and latent variables have a mean of zero and a variance of one. Under this assumption, the following statements can be made:

1. The factor loadings λ_{ij} represent the correlations between the observed self-report scores, y_i , and the latent endogenous (η_j) or exogenous (ξ_j) factors. The squares of the factor loadings represent the amount of variance in the psychometric data that is common with

the latent construct and can be used as measures of the reliability/validity of y_i as an indicator of η_j , or the reliability/validity of x_i as an indicator of ξ .

2. The error variances θ_{ii} represent the proportion of variance in y_i that is not explained by the latent factor. This unexplained variance is likely to arise from two sources of unreliability and invalidity in the psychometric data. First, there may be random errors of measurement which reflect test unreliability. These errors are uncorrelated with each other and with the latent factors in the model. Second, there may be systematic errors of measurement which reflect test invalidity. These errors may be correlated with each other or with the latent exogenous factor but are not correlated with the latent endogenous factors.
3. The covariances Ψ_{jk} represent the associations of the disturbance terms of the latent constructs η_j, η_k . Since the total variance in each η is 1.0, the ψ_{ii} variance corresponding to each η gives the proportion of error variance in the prediction of that η . Equivalently, $1 - \psi_{ii}$ gives the proportion of variance in the η explained by the whole model.

7.4.1.5 Identification of model

The model of the relationships between the 12 observed variables can be shown to be overidentified. This can be seen from the fact that there are 78 observed variances-covariances and 32 model parameters to be estimated, leaving a model with 46 degrees of freedom. The 32 model parameters to be estimated include: (a) 7 coefficients λ_i^y , (b) 4 coefficients γ_i , (c) 5 coefficients β_i , (d) 1 coefficient ϕ , (e) 11 variances of the error terms ϵ_i , and (f) 4 variances of the disturbances ζ_i . Evidence also comes from the literature that this type of model is identified. The measurement component of this model is of the form of a simple confirmatory factor model; in the absence of correlated errors, such models have been shown to be identified by Long (1983a). The structural model component is in the form of a recursive path model and such models are known to be identified (Duncan, 1975).

7.4.2 Method effects in self-report data

The model proposed in Figure 9 provides a basic account of the relationships between motivational, cognitive, and metacognitive factors. However, it is necessary to elaborate this model further, as one of the key assumptions listed previously, that the disturbance terms, ϵ_i ,

are uncorrelated with each other, is likely to be unrealistic. This is because it is probable that common response tendencies influence the observed measures in a similar way within groups. Psychometric data obtained from self-report questionnaires are usually prone to systematic distortion by generalized response biases (Parkes, 1987). These response biases include: the tendency to agree with items regardless of content, the tendency to respond consistently at either end of the scale rather than in the centre, and the tendency to respond in a socially desirable manner. There is also evidence to suggest that errors of measurement in reported data collected from the same source (e.g., use of questionnaires, use of psychometric scales, use of teacher-ratings) may be correlated (Fergusson & Horwood, 1987, 1989). Thus, the validity of the assumption of uncorrelated errors of measurement is open to doubt.

It is possible to incorporate method or response specific errors of measurement into the general equation as follows:

$$y_i = \lambda_{ij} \eta_j + \epsilon u_i + \epsilon u_{ih}$$

where η_j denotes the latent endogenous factor, ϵu_i denotes random errors of measurement in the i^{th} measure y_i , and ϵu_{ih} denotes systematic sources of variance in y_i from source h which are shared or in common across measures. Thus, variation in the self-report data is assumed to be a function of three non-observed sources of variation:

1. Variation attributable to true variation in y_i .
2. Variation attributable to random errors of measurement (unreliability) in the reporting of y_i .
3. Variation attributable to common response tendencies or common method effects which reflect situational factors or response tendencies which exert a consistent influence on self-report data or psychometric instruments.

It is possible to rewrite the above equation as:

$$y_i = \lambda_{ij} \eta_j + \epsilon_i$$

(where $\epsilon_i = \epsilon u_i + \epsilon u_{ih}$)

where the disturbance term ϵ_i combines variation from two sources: (a) random errors of measurement that are uncorrelated with each other, and (b) variation due to response, method, or situational factors in reporting, which exert a constant influence on psychometric instruments. In such models it is immediately apparent that for any measures y_i, y_j sharing the same systematic or common sources of error variance, the assumption of uncorrelated errors of measurement becomes invalid, that is covariance $(\epsilon_i, \epsilon_j) \neq 0$.

Hypothesized correlations between measurement errors are represented by two-headed, curved arrows. The absence of such curved arrows between ϵ 's in Figure 9 (see p. 121) indicates that all errors of measurement correlations were constrained to a value of zero. The addition of correlated measurement errors to a model will nearly always improve goodness-of-fit; however, it can also "mask a true underlying structure" (Gerbing & Anderson, 1984). Furthermore, it is inherently susceptible to capitalization on chance characteristics of the data, thus raising the question of whether these model modifications generalize to other samples or to the population (MacCallum, Roznowski, & Necowitz, 1992). For these reasons, correlated measurement residual assumptions should be based on substantive considerations. Fornell (1983) suggests that allowing for correlated measurement errors should not be motivated by goodness-of-fit improvement, unless (a) it is warranted on theoretical or methodological grounds, or (b) it does not significantly modify the structural parameter estimates. In addition, it has been suggested that the sensitivity of parameter estimates to the specification of correlated errors of measurement should be investigated (Hughes, Price, & Marrs, 1986).

7.4.2.1 Respecification of Model 1 to include correlated errors

Figure 10 presents an extension of Model 1 proposed earlier, to allow for the influence of correlated errors of measurement. While the structural component of this model is identical to the structural model described previously, it is now necessary to elaborate the measurement component of the model presented previously to allow for the presence of response tendencies in the self-report data. More specifically, an important elaboration of the measurement model is that the variance-covariance matrix Θ_ϵ of the disturbance terms ϵ_i is no longer assumed to be diagonal. Instead, the disturbance terms of the psychometric indicators are allowed to be correlated.

The above model, which incorporates response tendency or method effects (i.e., correlated errors of measurement), can be shown to be overidentified. This can be seen from the fact that there are 78 observed variances-covariances and 35 model parameters to be estimated, leaving a model with 43 degrees of freedom. The 35 model parameters to be estimated include: (a) 7 coefficients λ_{γ} , (b) 4 coefficients γ_{β} , (c) 5 coefficients β_{ϕ} , (d) 1 coefficient ϕ , (e) 11 variances and 3 covariances of the error terms ϵ_i , and (f) 4 variances of the disturbances ζ_i . Evidence also comes from the literature that this type of model is identified. The measurement component of this model is of the form of a simple confirmatory factor model; in the presence of correlated errors, such models have been shown to be identified by Long (1983a) and Horwood (1987). The structural model component is in the form of a recursive path model and such models have been shown to be identified by Duncan (1975).

7.4.3 The multisample situation

The models described in sections 7.4.1 and 7.4.2 provide a basic account of the relationships between motivational, cognitive, and metacognitive factors for a single group. However, it is necessary to elaborate the model further to include a multisample situation, since this was the focus of the present study. Assume that there are two groups of readers: reading disabled (RD) and normal achieving (NA). In conducting multisample analyses, one can test the hypothesis that the proposed model in Figure 11 (see p. 133) holds in each group simultaneously. By imposing between-sample constraints on selected coefficients (structural path estimates, factor loadings) or sets of coefficients, it is possible to test whether these coefficients are equal across groups. In general, any degree of invariance can be tested, from the one extreme where all parameters are assumed to be equal across groups, to the other extreme where only the pattern of fixed and free parameters is the same (Jöreskog & Sörbom, 1989). In the present case, it seemed particularly interesting to examine whether the causal paths among the latent variables, the factor loadings, or both were equal across groups for the model. Inasmuch as the latent variables were assessed via the same indicators in both groups, an identical factor loading matrix might reasonably be expected. However, according to theory, one may expect some of the causal relationships between the latent variables to vary in magnitude across groups.

It has been recommended that the general form of the measurement model must hold for all groups for which meaningful comparisons are to be made (Bollen, 1989). Conceptually,

measurement invariance refers to the extent to which a construct or a measure of a construct retains its meaning across groups or over time (Byrne, 1989a, 1989b; Byrne, Shavelson, & Muthén, 1989). In the measurement model, the factor loadings are of central importance in the definition of each latent variable and the validation of the underlying factor structure. For comparisons across multiple groups, the minimal condition of *factorial invariance* is typically taken to be the invariance of factor loadings (Bollen, 1989). However, Byrne et al. (1989) maintain that comparisons between groups can be undertaken if at least one factor loading per factor is statistically equivalent across the groups, a condition referred to as *partial invariance*.

7.4.3.1 Respecification of Model 1 to include the multisample situation

Figure 11 shows that for each group there are four endogenous variables and one exogenous variable. The measurement (except for the errors of measurement) and structural components have similar patterns of relationships across groups. The specification of the structural relationships is similar across groups, except that the magnitude of some of the causal relationships are hypothesized to differ across groups. Interpretation of variables and parameters is similar to that outlined for the model in section 7.4.1 except that it now applies to two reader groups.

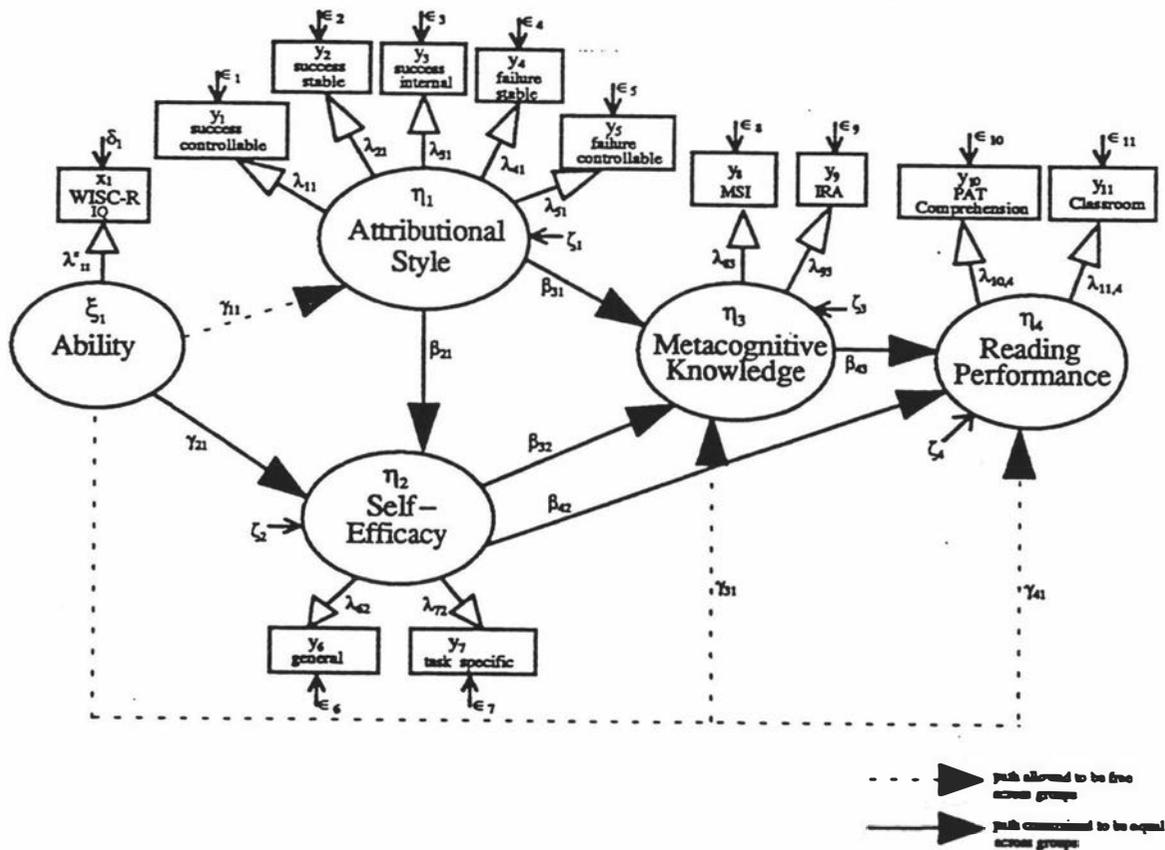


Figure 11. Hypothesized Model 1 respecified for the multisample situation.

The measurement model described in section 7.4.2.1 with the model specifications given previously for the model can be extended to the two group situation to produce the LISREL structural equation specifications described below. It is important to note that the assumptions underlying this model (except for model assumption no. 5 concerning uncorrelated measurement errors, see section 7.4.1.4), and the interpretation of the model parameters are identical to those described previously for that model. As these assumptions and parameter interpretations have been described previously, they are not repeated here in the interests of brevity.

(i) Measurement model

The measurement model describing the relationship between the measures and the latent constructs can be expressed as a series of linear equations. The general form of these equations for the endogenous variables is

$$y_i^{[g]} = \lambda_{ij}^{[g]} \eta_j^{[g]} + \epsilon_i^{[g]}$$

where η_j ($j = 1$ to 4) denotes the non-observed true attributional style, self-efficacy, metacognitive knowledge, and performance factors; the coefficients λ_{ij} describe the regressions of the psychometric measures y_i ($i = 1$ to 11) on the latent attributional style, efficacy, metacognitive knowledge, and performance factors η_j ($j = 1$ to 4); ϵ_i ($i = 1$ to 11) denotes random and response/method specific errors of measurement in the psychometric or self-report data; and where the superscript g represents a distinct matrix for the NA group ($g = g1$) or RD group ($g = g2$).

As there is one exogenous variable the measurement model for the exogenous variable reduces to

$$x_1^{[g]} = \lambda_{11}^{[g]} \xi_1^{[g]} + \delta_1^{[g]}$$

In LISREL notation the above measurement model may be expressed in matrix notation as follows:

$$y^{[g]} = \Lambda_y^{[g]} \eta^{[g]} + \epsilon^{[g]}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \end{bmatrix}^{[g]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ \lambda_{41} & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \lambda_{93} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \lambda_{11,4} \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}^{[g]} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \end{bmatrix}^{[g]}$$

$$\begin{aligned}
 \mathbf{x}^{[g]} &= \Lambda_x^{[g]} \xi^{[g]} + \delta^{[g]} \\
 [\mathbf{x}_1]^{[g]} &= [\Lambda_{11}]^{[g]} [\xi_1]^{[g]} + [\delta_1]^{[g]} \\
 \text{where } \lambda_{11} &= 1.
 \end{aligned}$$

In this model, assume that equality constraints of the factor loadings across the NA and RD reader groups are made for the measurement parameters. This is specified as $\Lambda_x^{[g1]} = \Lambda_x^{[g2]}$ and $\Lambda_y^{[g1]} = \Lambda_y^{[g2]}$ where g1 represents the NA group and g2 represents the RD group.

(ii) *Structural equation model*

The structural relationships between the latent factors of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance implied by Figure 11 are given by the following equations:

$$\begin{aligned}
 \eta_1^{[g]} &= \gamma_{11}^{[g]} \xi_1^{[g]} + \zeta_1^{[g]} \\
 \eta_2^{[g]} &= \gamma_{21}^{[g]} \xi_1^{[g]} + \beta_{21}^{[g]} \eta_1^{[g]} + \zeta_2^{[g]} \\
 \eta_3^{[g]} &= \gamma_{31}^{[g]} \xi_1^{[g]} + \beta_{31}^{[g]} \eta_1^{[g]} + \beta_{32}^{[g]} \eta_2^{[g]} + \zeta_3^{[g]} \\
 \eta_4^{[g]} &= \gamma_{41}^{[g]} \xi_1^{[g]} + \beta_{42}^{[g]} \eta_2^{[g]} + \beta_{43}^{[g]} \eta_3^{[g]} + \zeta_4^{[g]}
 \end{aligned}$$

In LISREL notation these equations may be written in matrix form as follows:

$$\begin{aligned}
 \eta^{[g]} &= \Gamma^{[g]} \xi^{[g]} + \mathbf{B}^{[g]} \eta^{[g]} + \zeta^{[g]} \\
 \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}^{[g]} &= \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \\ \gamma_{41} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 \\ 0 & \beta_{42} & \beta_{43} & 0 \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}^{[g]} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}^{[g]}
 \end{aligned}$$

7.5 Advantages and limitations of using SEM techniques

The preceding discussion has outlined the statistical theory and rationale behind structural equation modelling and the specification of structural equation models within a LISREL framework. It is clear that SEM has much to offer in the study of the models examined in this thesis. Specifically the advantages are:

1. SEM permits directional predictions amongst variables and modelling of indirect effects (Bollen, 1989).
2. SEM allows one to test models in which the latent factors are derived from multiple measures of the same construct rather than from a single measure; in using multiple rather than single indicators, it directly deals with the inherent unreliability of scales used in social science research (Hoyle & Smith, 1994).
3. SEM enables one to test the overall adequacy of an entire model as well as its subcomponents (Hull & Mendolia, 1991).
4. SEM identifies sources of poor model fit (Hayduk, 1987).
5. SEM allows one to go beyond testing hypotheses between means of groups to testing hypotheses about similarities and differences in structural relationships among latent factors across populations (Jöreskog & Sörbom, 1989).
6. SEM permits a test of competing views or theoretical models; for example, one can evaluate whether the fit of model improves by dropping, adding, or constraining specific paths, which have been hypothesized by competing theories (Bollen, 1989).

These features provide substantial advantages over more traditional methods in which data are analyzed using significance testing and other procedures. Despite these advantages, there are a number of shortcomings in using SEM techniques which deserve mention. First, in applying SEM techniques, there is the requirement that one has a clearly formulated theoretical model in mind to test. However, most existing social science theories are not precisely formulated. Consequently, the process of model building requires the investigator to make various assumptions and "inductive leaps" from existing theory regarding the causal structure of the variables (Fergusson, 1995; Fergusson & Horwood, 1988). The extent to which these assumptions may be justified depends on the state of the existing knowledge and a priori theory within a given area. Thus, the use of SEM is effective in the context of confirmatory data analysis (Fergusson & Horwood, 1987), that is, where there is substantial a priori theory

about the processes under study. It is likely that an investigator using SEM techniques in an exploratory framework to strengthen a weak theoretical model could encounter problems. The present study illustrates this point clearly to the extent that the models proposed draw heavily from a body of a priori theory. In the absence of this prior theoretical and empirical knowledge, it would have been difficult to develop clearly formulated models of the relationships between cognitive, metacognitive, and motivational variables.

A second problem is that SEM techniques assume interval scaled data, yet variables of interest are often only nominal or ordinal (Bollen, 1989; Fergusson, 1995). Although LISREL can accommodate discrete exogenous variables, because of the violation of the assumption of homoscedasticity in error terms, there are problems of estimation and prediction when either an endogenous variable or operationalization is dichotomous. In general, parameter estimates will not be efficient, predictions will be imprecise, tests of significance will not apply, and estimated standard errors will not be consistent (Hayduk, 1987).

Another assumption of SEM techniques that poses a problem is the assumption of linear and additive functional relationships among variables. Since some phenomena in social sciences are governed by nonlinear processes, it may be inappropriate and misleading to use linear structural equations to represent them. While SEM methods usually do not allow for the presence of non-linear relationships between variables, a number of techniques have been developed recently to accommodate non-linear or interactive effects within structural equation models (see Bollen, 1989; Hayduk, 1987). A simple interactive structure is investigated in the present study. This involves studying the implications of group status (reading disabled vs normal achieving), a categorical variable, on the relationships of the cognitive, metacognitive, and motivational variables. One way this simple interactive structure is modelled is via an extension of the theory for single population models to a multisample context (Jöreskog & Sörbom, 1989). With a multisample analysis it is possible to test whether different covariance structures hold for different subgroups of the population.

SEM techniques have a number of other shortcomings. It assumes that observations are obtained from a multivariate normal distribution. The estimates of parameters are efficient only for large samples. The chi-square goodness-of-fit test, which is a large sample approximation, is directly sensitive to sample size. Another important limitation is the identifiability of the model. It is not necessarily the case that formulated models will be identified and therefore be amenable to solution using SEM. To obtain identifiability, the investigator may be forced to simplify the model by assuming that a number of potential causal relationships do not exist. These simplifying assumptions may "make structural

equation models approximations, perhaps caricatures, of the complex reality they represent" (Fergusson, 1995, p. 140).

A final limitation concerns the extent to which all variables of theoretical relevance to the processes under study have been considered and included (Fergusson, 1995). While the omission of variables does not necessarily impugn the validity of a model, it is always preferable that the investigator make allowances for such difficulties in the model specification and data collection. In cases where it may be thought that one or more theoretically important variables have been omitted, it is important to recognise that it is unrealistic to expect a single model to provide a full account of all the variables that may be of potential theoretical relevance to the processes under study.

7.6 Conclusion

It is apparent that SEM techniques provide the social scientist with a potentially powerful tool for testing the adequacy of causal hypotheses. The logic of SEM is that it forces researchers to clarify the conceptual reasons for interrelating measures. That is, in specifying a causal model, a researcher is forced to explain in detail why measures should be interrelated. The present review suggests that the process of model building and testing requires an interplay between the researcher, theoretical considerations, and the application of appropriate statistical methodology. The problems in specifying, fitting, testing, and interpreting the covariance structure models may sometimes lead to failure in one or more aspects of this process.

In conclusion, it can be argued that there are several advantages to using these techniques in the present study. More specifically, the use of SEM allows the investigator to specify and test a range of clearly formulated hypotheses, in addition to competing views and theories. It also allows an examination of the similarities and differences in the structural relationships across subpopulations. Unlike approaches that simply sum subcomponents to form a composite scale, SEM enables researchers to specify the theorized relations among the subcomponents in terms of latent variables. In addition, it provides a test of these theorized measurement and structural systems, while retaining information about the unique aspects of the subcomponents. For all these reasons, structural modelling was considered to be particularly suited to testing the theorized relations among ability, attributional style, self-efficacy, metacognitive knowledge, strategy use, and reading comprehension.

CHAPTER 8

METHOD

8.1 Overview

The role of intellectual ability, motivational beliefs, metacognitive knowledge, and cognitive/metacognitive strategy use in reading comprehension was studied in a sample of reading disabled and achieving readers from nine intermediate schools in the Manawatu and Wanganui regions. Prior to conducting the present study, several pilot tests were conducted in order to refine measures that were developed for this study. These included the efficacy and strategy use measures.

There were three phases involved in the present study. Phase 1 concerned sample selection. Selection of reading disabled adolescents was based on a six-stage multidefinitonal approach. A sample of normal achieving readers with achievement in reading consistent with age was also identified.

The data were collected in early-mid 1994 in two phases (Phases 2 and 3). Phase 2 involved the administration of two self-report questionnaires (Reading Questionnaire I and II) which examined students' attributional style, self-efficacy, metacognitive knowledge, and use of strategies in the domain of reading. Testing of groups of students took place during two 45 minute sessions with both questionnaires administered within a period of one week.

Phase 3 involved reading interviews, conducted individually. These interviews lasted 25 minutes and were conducted within two weeks of Phase 2. The taped interviews provided further data on students' cognitive and metacognitive strategy use. Following Phase 3, data coding was undertaken and reliability of the interview data was established. Thereafter, analyses of the data proceeded.

This chapter therefore describes the identification of subjects, the IQ and achievement tests used, the cognitive, metacognitive, and motivational measures used, and how they were applied. In addition, the statistical methods that were employed to analyze the data are also described.

8.2 Sample

Parental and child consent was obtained for those students who were selected to participate in the study (see Appendix A). The sample was selected from the 1993 cohort of Form 2 students enrolled at nine different intermediate schools (which served a wide range of socio-economic backgrounds) in the Manawatu and Wanganui regions. This Form 2 cohort comprised approximately 2000 students. The schools varied in size and type of location (rural vs urban), internal organisation, and methods of teaching reading.

Initially, students who gave consent to participate in the study were assessed via a battery of tests and screens involving standardized tests of achievement and a test of intelligence. The aim of this battery was to appraise how the students were functioning in terms of their academic competence. Students' achievement in a variety of subjects was initially assessed using all available school records of the seven tests in the Progressive Achievement Test Series (PAT; Elley & Reid, 1971; N. A. Reid, 1993; N. A. Reid, Croft, & Jackson, 1978; N. A. Reid & Elley, 1991). A short form of the Wechsler Intelligence Scale for Children - Revised (WISC-R; Wechsler, 1974) was used as a measure of intellectual functioning. Only those who met the selection criteria continued in the study. The procedures for selecting the reading disabled (RD) and normal achieving (NA) groups are summarized below.

8.2.1 Selection of reading disabled (RD) sample

In Chapter 2, it was argued that a multidefinitonal assessment procedure may provide an improvement over traditional discrepancy-based methods in identifying RD children. Such an approach requires the determination of a set of selection criteria that would constitute key defining characteristics (or operational interpretations) of the RD child. In the present study, a number of key defining features that have been commonly employed across definitions of reading or learning disability were extracted from the literature. These key features were based on the best available knowledge of the learning disability concept.

A six-stage assessment model was used where: (a) RD readers were recommended by teachers (based on absolute criterion of low reading achievement in the classroom), (b) students achieved below or equal to the 25th percentile in one of the reading PAT tests (comprehension or vocabulary), (c) students had an inconsistent achievement profile based on their standardized achievement test scores, (d) students had no evidence or history of

neurological abnormality, emotional disturbance, or cultural deprivation as was apparent from school records, (e) students did not experience difficulties with English-as-a-second-language, and (f) students had prorated IQ scores of 85 or more. A more detailed description of each of these steps and the rationale for including them follows.

The first step involved teacher recommendation and has been included by some researchers as one step in identifying RD children (Berninger et al., 1992; Carr et al., 1991; Leong, 1985, 1987). The second step of the screening procedure identified students who had at least one reading PAT score below or equal to the 25th percentile and at least one PAT score above or equal to the 30th percentile. The 30th percentile represents one-half a standard deviation below the mean. Students who had at least one PAT result equal to or greater than the 30th percentile were postulated to have an inconsistent achievement profile. Furthermore, it was believed that such achievement may indicate average intelligence. Students showing consistent low levels of academic achievement across a variety of domains (indicated by *all* PAT scores below the 30th percentile) were excluded from subsequent screening with the WISC-R because of the likelihood that they would be slow learners. This screen provided students who were achieving at a level approximately 1-2 years below the average Form 2 age range (12-13 years) in reading but who gave some indication of being in the normal IQ range. The reading achievement criterion (at or below the 25th percentile) adopted in this study has commonly been used in the reading disability literature (Bowers, Steffy, & Tate, 1988; Fletcher et al., 1994; B. A. Shaywitz et al., 1992; Siegel, 1992; Swanson & Trahan, 1995).

In the second step of this six-stage assessment model, it was decided to dismiss the notion of a severe discrepancy definition *based on a difference score* as commonly used (i.e., a calculation of discrepancy scores between intelligence and reading achievement). Instead, a low achievement definition of reading disability was adopted. In this definition the notion of discrepancy is still maintained as the finally selected students are required to have normal intelligence (see step 6). Such an approach is increasingly being used in studies of RD and LD children (Szuskiewicz & Symons, 1993). In the low achievement definition, the "unexpectedness" would still occur because these adolescents are not intellectually deficient and possess adequate intelligence to master reading skills (Siegel, 1988, 1989a). Fletcher et al. (1994) suggest that this approach simplifies the definition of reading disability at a single point based on arbitrary decisions about severity by avoiding problems associated with calculating a discrepancy score. Such an approach provides a dimensional view of reading

disability and is comparable to those used in defining mental retardation and childhood psychopathology (Fletcher et al., 1993).

The third step entailed looking at the "intraindividual" differences in academic achievement of the RD student. This criterion was included as it has been suggested that differences between performance on different tasks or between academic areas is characteristic of the RD child whilst fairly even, low levels of achievement is characteristic of the mentally retarded child or slow learner (Chalfant, 1989; Keogh 1988a; Kirk & Chalfant, 1984).

The employment of steps 2 and 3 as key defining features of learning disabilities is supported by a number of researchers. For example, Epps, Ysseldyke, and Algozzine (1985) noted three diagnostic characteristics of learning disabilities. These were a discrepancy between ability and achievement, low achievement, and a "scattered" assessment profile indicating variable performance in a number of areas. Valus (1986) surveyed the teachers of students with learning disabilities who reported that intraindividual differences, achievement-potential discrepancy, and below average academic work were the major identification criteria.

In the fourth step, school records were consulted. All students with any known history or evidence of neurological abnormalities, behavioural, emotional and attentional problems, hearing and vision abnormalities, cultural deprivation or any other major handicapping condition were excluded. This was done in accordance with the conventional exclusion criteria for reading disabilities. Students experiencing English-as-a-second-language difficulties were also excluded (step 5). Exclusion of children with difficulties in English has been used in a number of studies (Fletcher et al., 1994; Horrex, 1992; Swanson & Ramalgia, 1992). A total of 228 students met the above 5-step criteria.

The last step involved selecting students with normal intelligence. Several researchers have suggested that IQ has limited value for the differential diagnosis of reading disability (e.g., Siegel, 1988). However, it is argued here that IQ tests should not be rejected completely; in the present study, it is included as a threshold concept where an IQ of 85 is conceptualized as the lower bound in defining reading disabilities (Leong, 1989). It is reasonable to include IQ so long as it is not employed as the sole criterion for defining reading disabilities. In addition, normal intelligence was considered necessary to control for general academic aptitude so as to exclude mental deficiency as an explanation for students' reading difficulties.

Level of intelligence was determined by testing the students with a short form of the WISC-R (comprising of Information, Vocabulary, Picture Completion, and Block Design subtests). Testing was conducted in the schools during the early part of 1993 by the researcher and five senior postgraduate psychology students who had advanced testing qualifications and practical experience. The next procedure involved deleting from the target group those students whose prorated intelligence scores were less than 85. Full Scale IQ scores were prorated from the summed scaled scores, following the procedure suggested by Tellegen and Briggs (1967). All subjects in this final group ($n = 204$) obtained a full scale IQ score above 85, thus classifying them within the normal IQ range which is in accordance with current RD definitions and research (Bowers et al., 1988; Eden, Stein, Wood, & Wood, 1995; Stanovich, 1991a, 1991b; Swanson & Ramalgia, 1992).

8.2.2 Selection of normal achieving (NA) sample

Normal achieving readers were randomly selected from the nine intermediate schools. These students had been administered the same ability and achievement tests as the RD group. Students who scored less than the 30th percentile on either the PAT Reading Comprehension or PAT Vocabulary test were dropped from further study. Students included in the final NA readers group ($n = 203$) met the following criteria:

1. PAT Reading Comprehension and Reading Vocabulary scores greater than or equal to the 30th percentile.
2. a prorated IQ of 85 or above on the WISC-R.

The cut-off point of the 30th percentile in both the PAT reading tests indicates that these students were performing at a level no less than one-half a standard deviation below the mean. The 30th percentile cut-off for identifying NA children has been used by a number of researchers (e.g., Siegel, 1992). This criterion was somewhat lower than is typical in the literature. A number of studies adopt a criterion of the 40th percentile or above (Share, Moffitt, & Silva, 1988; B. A. Shaywitz et al., 1992). This would mean that 40% of the population are not normal achieving. The use of such a conservative criterion in the literature is open to criticism given that normal performance is commonly defined at plus and minus one standard deviation from the mean (equivalent to a percentile rank range of 16 to 84). Given this definition, the current study used a more stringent definition for the lower end of

the range and a more lenient one at the higher end. The adoption of such a criterion for this study allowed a wide range of students to be examined and was defensible, given the goal of studying within group differences in a model of reading achievement. It should be noted that although the NA group consisted mainly of average achieving readers, some high achieving readers were also included.

8.2.3 Demographic information

Table 2 presents a summary of the demographic characteristics of the sample in each group. The sample consisted of 51.6% males and 48.4% females. The majority of the students were of European descent (68.1%), 14.3% were of mixed European and Maori descent, 13.3% were of Maori descent and 4.3% belonged to the "other" ethnic categories. The NA sample ethnicity and gender proportions are consistent with those reported by the Ministry of Education for the Manawatu-Wanganui region (Ministry of Education, personal communication, October 1994).

The RD group comprised 120 boys and 84 girls. The proportion of males to females in the RD sample is consistent with previous research indicating that boys make up a greater proportion of the RD population (Chapman et al., 1984; S. E. Shaywitz & Shaywitz, 1988). Significant differences for age ($t = -4.07, p \leq .001$), gender ($t = 2.95, p \leq .003$), and ethnicity ($t = 6.92, p \leq .001$) were found between the RD and NA groups.

Table 2. Demographic characteristics of reading disabled and normal achieving readers

Group	Number of cases	Age (months)		Gender (%)		Ethnicity (%)			
		<i>M</i>	<i>SD</i>	Male	Female	European	European and Maori	Maori	Other
NA	203	149.6	3.8	44.3	55.7	83.3	7.9	6.4	2.4
RD	204	151.2	4.1	58.8	41.2	52.9	20.6	20.1	6.4

The gender and ethnic differences between groups are a predictable consequence of separation into NA and RD readers, that is, one can expect more boys and non-Europeans to be poorer readers (e.g., Fergusson, Lloyd, & Hørwood, 1991; Marsh, Smith, & Barnes, 1985). However, these differences may have an impact on the structural relationships of the hypothesized models investigated in the present study. This issue was investigated later in the Results chapter in the form of supplementary analyses (see also Appendix J).

Table 3. Ability and achievement scores of reading disabled and normal achieving readers

Group	Number of cases	WISC-R Full Scale IQ		PAT Reading Comprehension (Age percentile)		PAT Reading Vocabulary (Age percentile)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
NA	203	102.9	9.1	64.0	18.5	62.3	19.5
RD	204	95.6	8.1	20.4	12.7	20.9	13.4

Table 3 presents the IQ and Reading PAT data for the two sample groups. The mean prorated IQ of the RD and NA groups were 95.6 and 102.9, respectively. The difference in IQ scores between the groups is statistically significant ($t = 8.58, p \leq .001$). This finding is consistent with previous research which has shown the RD group, often even after matching, to have IQ scores which, though falling in the normal range, are somewhat lower than those of the NA group (Stanovich, 1986a; Torgesen & Dice, 1980; van Kraayenoord, 1986). The mean difference in IQ between groups of 7.3 is consistent with that found in a formal survey conducted by Torgesen and Dice (1980). They found the mean IQ of the learning-disabled group to be 6 points lower than that of the control group.

As expected, significant group differences were also found for the PAT Reading Comprehension ($t = 27.7, p \leq .001$) and Vocabulary tests ($t = 24.9, p \leq .001$). According to national norms for Form 2 students, the PAT Reading Comprehension figures represent approximately the 64th percentile for NA readers and the 20th percentile for RD readers.

8.3 Measures

The selection of the research instruments was dictated by a number of requirements. First, the instruments chosen were required to have good reliability and validity. Second, instruments that were commonly used in the literature were employed to enable comparison with similar research that investigated predictors of reading comprehension. When reading specific measures were unavailable from the research literature, they were developed by the author. Finally, measures that were brief and easy to administer were employed due to time constraints for testing in schools, and also to avoid encroaching too much into students' classroom time.

Nearly all of the measures used in the present study involved self-report. In order to optimize the validity of the students' self-reports, the collection of the self-report data conformed to the recommendations set forth by Assor and Connell (1992). First, students were made to feel valued as reporters by informing them why they were being asked the questions and what the investigator planned to do with their answers. Second, the investigator ensured that subjects understood the questions and made it clear what information was required from them without using deception. Third, students were informed that their responses were confidential. In addition they were informed of the process by which confidentiality would be upheld. Fourth, the investigator reassured the subjects that any answer given to the question was acceptable provided it was what they really believed. This was done by designing sets of instructions to put students at ease and encouraging them to use the full range of responses on the questions. Fifth, the questionnaires and interviews were administered without a "known" adult present (no teachers, parents, or principals). Sixth, the questionnaires were administered to manageable sized groups of 15 to 20. Seventh, sessions lasted about 45 minutes (as students disengage when the assessment is too long). In addition, each questionnaire session was divided into intervals with a 3 minute break between sections. Lastly, the investigator read the questions to the students and ensured they were on task, offering individual help when deemed necessary (e.g., if students required help in understanding the question).

Table 4 shows the constructs that were examined in the present study and the measures that were used to assess these constructs. A more detailed description of the measures is provided in sections 8.3.2 through 8.3.7. The measures reported contain items that were originally intended to be used in the study. However, preliminary factor and reliability analyses

conducted on the questionnaire items led to some minor modifications of the intended measuring scales.

Table 4. Constructs of interest and their corresponding indicators

Construct	Instruments
Ability	Short form of the Wechsler Intelligence Scale for Children - Revised (WISC-R; Wechsler, 1974)
Academic achievement	Progressive Achievement Test Series (PAT; Elley & Reid, 1971; N. A. Reid, 1993; N. A. Reid et al. 1978; N. A. Reid & Elley, 1991)
Reading comprehension	1. Progressive Achievement Test for reading comprehension (N. A. Reid & Elley, 1991) 2. Classroom performance
Attributional style	Causal Dimension Scale (CDS; Russell, 1982)
Self-efficacy	1. General measure of reading self-efficacy (GRSE; Pereira-Laird & Deane, 1995) 2. Task-specific measure of reading self-efficacy (TRSE; Pereira-Laird & Deane, 1995)
Metacognitive knowledge	1. Metacomprehension Strategy Index (MSI; Schmitt, 1988, cited in Schmitt, 1990) 2. Index of Reading Awareness (IRA; Jacobs & Paris, 1987)
Metacognitive strategy use	1. Self-report measure: Reading Strategy Use scale (RSU; Pereira-Laird, 1996), RSU-meta 2. Behavioural measure: Think-Aloud Reading Strategy Use Measure (TARSUM; Pereira-Laird & Deane, 1996), TARSUM-meta
Cognitive strategy use	1. Self-report measure: Reading Strategy Use scale (RSU; Pereira-Laird, 1996), RSU-cog 2. Behavioural measure: Think-Aloud Reading Strategy Use Measure (TARSUM; Pereira-Laird & Deane, 1996), TARSUM-cog

Although the questionnaires were designed to measure the latent variables of interest in this research, some preliminary factor analyses were necessary to ensure that the variables measured the constructs they were designed to measure (see Appendix B, sections B.1 and B.2). Even though items were drawn from prior studies, exploratory factor analyses were considered necessary because the wording was changed slightly and reading specific items were added. A series of principal components analyses were conducted on items of strategy

use (self-report), self-efficacy, metacognitive knowledge, and attributions, for each group and for the whole sample using a varimax and then an oblique rotation. A second series of principal components analyses was conducted using strategy-use items only. Factor analysis resulted in the exclusion of some items because of a lack of correlation or stable factor structure across groups. Based on the results of factor analyses, composite scores were formed of those questionnaire items that theoretical and statistical analysis suggested as valid measures of the constructs of interest.

8.3.1 Pilot tests

Several pilot tests were conducted in order to refine the TRSE, RSU-meta, RSU-cog, TARSUM-meta, and TARSUM-cog measures, to establish the amount of time taken to administer the measures, to trial instructions given to the students, to determine the appropriateness and clarity of the language of the measures, and to determine the suitability of the passages. Details of these pilot tests are provided in Appendixes C through E.

8.3.2 Ability

In terms of measured intelligence, individual tests are preferred over group tests with the Wechsler scales being the most frequently used (German, Johnson, & Schneider, 1985). Accordingly, a short form of the Wechsler Intelligence Scale for Children - Revised (WISC-R: Wechsler, 1974) was used as the measure of general intellectual functioning. The WISC-R, which correlates highly with other measures of intelligence, has been described as the best standardized and most widely used test for school-age students (R. L. Taylor, 1984). The subscales are moderately reliable and highly intercorrelated, and the derived verbal and performance scales are even more reliable and intercorrelated (Mitchell, 1985). The WISC-R has also been shown to predict academic achievement equally well for blacks and whites (in the United States) and to be relatively free from item bias (C. R. Reynolds, 1982).

Some of the items in the information subtest were modified to suit New Zealand students. Following Tuck, Hanson, and Zimmerman (1975), the changes made were as follows: question (Q) 5 was changed from "pennies" and "nickel" to "cents" and "dollar"; Q12 was changed from "America" to "New Zealand", Q19 from "two countries that border the United states" to "two Australian states", Q20 from "pounds" and "ton" to "kilograms" and "tonne",

Q24 "American man" to "New Zealand man", and Q27 "New York to Los Angeles" to "Auckland to Sydney". These modified items have been used previously by other New Zealand researchers, and reliability and validity data of the modified version were comparable to that of the original scale (R. St. George & Chapman, 1987).

Since it was not feasible to administer the full WISC-R test to a large number of students, a short form comprising four subtests was used. These were Information, Vocabulary, Picture Completion, and Block Design. They were chosen because of their relatively high correlations with Full Scale IQ, and because of their ease of administration. As shown by Sattler (1992), this particular combination of subtests as a four-subtest short form does not alter greatly the correlation of the short form with the Full Scale IQ -- all coefficients were about .94 for the ten combinations of four subtests presented by Sattler. Full Scale IQ scores were prorated from the summed scaled scores, using the procedure recommended by Tellegen and Briggs (1967).

8.3.3 Achievement

8.3.3.1 Academic achievement

Initial screening of subjects involved examination of students' academic achievement. Students' achievement was assessed using all seven of the Progressive Achievement Test Series (PAT), all of which were normed in New Zealand. The PATs measured achievement in the areas of Reading, Listening, Mathematics, and Study Skills. The PATs were Reading Comprehension (N. A. Reid & Elley, 1991), Reading Vocabulary (N. A. Reid & Elley, 1991), Listening Comprehension (Elley & Reid, 1971), Mathematics (N. A. Reid, 1993), and Study Skills (N. A. Reid et al., 1978). The latter consisted of three subtests, namely, Knowledge and Use of Reference Materials; Reading Maps, Graphs, Tables, and Diagrams; and Reading Study Skills. The PAT tests are group administered, paper-and-pencil scales, typically given by the majority of New Zealand primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in New Zealand (N. A. Reid, 1993; N. A. Reid & Elley, 1991).

8.3.3.2 Reading comprehension performance

The latent construct of reading comprehension performance was represented by two indicators.

Age percentile scores of a standardized measure of reading comprehension, the PAT Reading Comprehension, served as one indicator, since a child's achievement is most often assessed with a standardized test of reading ability (German et al., 1985). Classroom marks served as the second indicator because they provide a very important source of feedback to students.

Teachers provided gradings, expressed as a percentage, for students' performance on different reading tasks in the classroom. These tasks fell into three general categories: in-class seatwork, quizzes and tests, and assignments and projects. An average score for the three categories was calculated with high scores representing good reading performance.

8.3.4 Causal Attribution

Despite the evidence that causal attributions are important determinants of motivation and achievement, relatively little attention has been paid to the reliability and validity of different methods for assessing attributions. As discussed in Chapter 4, section 4.4.1, it was considered important to measure causal attributional dimensions directly so as to avoid the "fundamental attribution researcher error". Thus, in the present study attributional dimensions were measured directly using the Causal Dimension Scale (Russell, 1982). This measure was designed to assess how the student perceives the cause of the reading outcome in terms of the locus of causality, stability, and controllability dimensions described by Weiner (1979, 1985a). The measure consisted of nine semantic differential scales, with three of the scales assessing each causal dimension (see Appendix B, section B.1). Students were asked to think of two to three main reasons why they succeeded or failed in reading and to rate the extent to which these causes, in general, fell on the Causal Dimension Scale (CDS). Scores for the locus of causality, stability, and controllability dimensions were created by summing student's ratings of the attributions on each of the three semantic differential scales representing that dimension. These total scores ranged from 3 to 21 for both the success and failure conditions. A number of studies have provided support for the high reliability and validity of the CDS (e.g., Wilson & Linville, 1985).

The current study provides additional evidence on the CDS's reliability and validity. Factor analyses revealed that the attribution items formed separate factors from that of efficacy, metacognitive knowledge, and strategy use. Reliability estimates computed for the six

subscales in the present study indicated high internal consistency for the whole sample ($N = 407$) with alpha coefficients ranging from .85 to .95 for success subscales and .79 to .93 for failure subscales, thus supporting the formation of the six subscales.

The latent construct of adaptive attributional style for reading was hypothesized to be reflected in six manifest variables of *success* and *failure* on the following three dimensions of attributions: locus of causality, controllability, and stability. For reading success, these scales were scored so that higher values reflected increased adaptive attributions of locus of causality (internality), controllability, and stability. However, for reading failure, these scales were scored so that higher values reflected increased attributions of internality, controllability, and instability to represent adaptive attributions.

8.3.5 Self-efficacy

Two measures were devised by the author to represent the latent construct of self-efficacy. One of the measures was consistent with Bandura's (1977) original definition, "the conviction that one can successfully execute the behaviour required to produce particular outcomes" (p. 193). The other conformed to more recent definitions which emphasize that "self-efficacy is concerned with generative capabilities, and not with component acts" (Bandura, 1986, p. 397). Thus, the first indicator measured students' confidence about outcomes in relation to specific tasks in reading and the second measured students' aggregated reading self-efficacy.

The first measure, called the Task Specific Reading Self-efficacy (TRSE) scale, followed Bandura's original methodology and was more specific in terms of reading tasks. The 6-item scale measured individuals' generalized expectancies concerning their efficacy in performing a variety of reading tasks. Students' self-efficacy for reading different materials (e.g., an encyclopedia, a 150 page novel) was measured on a scale that ranged from 0 to 100 in 10-unit intervals from high uncertainty to complete certainty. Psychometric data for this scale are provided by Pereira-Laird and Deane (1995) (see Appendix C). High reliability and validity was demonstrated for this measure.

The second measure was called the General Reading Self-efficacy (GRSE) scale. In this measure, students were required to respond to 7 items on reading self-efficacy using a 7-point Likert scale. The scale ranged from 1 to 7 with the following verbal descriptors: 1 - *never*,

2 - *almost never*, 3 - *seldom*, 4 - *sometimes*, 5 - *often*, 6 - *almost always*, and 7 - *always*. Items were adapted from the self-efficacy subscale devised by Pintrich and De Groot (1990). The items were modified to make it reading specific. An attempt was made to incorporate many aspects of reading, for example, doing well in reading, ability to understand the material, and so forth. The items had the following stems "I am certain", "I am sure", and "I am confident". Items which indicated comparison to other students were excluded as it was believed that they measured perceived competence rather than self-efficacy. Scores for the seven items were added and then divided by seven to produce a mean score ranging from 1 to 7. Scoring was reversed on the items that were negatively worded, so that for all items a high score indicated high self-efficacy (see Appendix C for psychometric data). The GRSE scale is reported to have good reliability and validity (Pereira-Laird & Deane, 1995).

In the present study, factor analyses of the items from the GRSE and TRSE scales revealed that the items loaded on separate factors. However, in comparison to the single factor structure for NA readers, the TRSE scale items formed two factors for RD readers. Items 1 and 2 loaded on one factor while items 3 through 6 loaded on a second factor. Examination of the two factor structure of the task-specific items suggested that the responses of RD readers could be differentiated for easy and difficult reading tasks. In order to make the scales comparable across groups, items 1 and 2 were deleted from the TRSE scale. A task specific self-efficacy score was computed by calculating the mean for the remaining 4 items in the self-efficacy scale.

In the present study, assessment of reliability, suggested that both measures had satisfactory internal consistency as measured by Cronbach's alpha. Internal consistency for the GRSE measure varied from .64 to .87 for the RD and NA students, respectively; the TRSE measure had Cronbach alphas of .89 and .83 for NA and RD groups, respectively.

8.3.6 Metacognitive knowledge for reading

Two multiple-choice questionnaires were used to measure students' metacognitive knowledge for reading. Using multiple choice tests to assess metacognitive knowledge has a number of advantages. First, it avoids some of the pitfalls of verbal reports noted by Garner (1987). Second, it is more objective than interviews that may involve interpretations of open-ended responses, experimenter bias, or fabricated responses. Although these multiple-choice tests

allow guessing, they do not place shy, inarticulate, or RD students at a disadvantage. Finally, it is easy to administer and score and can be administered to groups in a relatively short period of time.

The Index of Reading Awareness (IRA) was developed by Jacobs and Paris (1987) and assesses children's metacognitive knowledge on four separate subscales: evaluation, planning, regulation, and conditional knowledge of reading. This measure, or an adaptation of this measure, has been used by a number of researchers (Carr et al., 1991; Oka & Paris, 1987; Paris & Oka, 1986; Swanson & Trahan, 1995). The IRA consists of 20 items, each with three alternatives representing an inappropriate response (0 points), a partially adequate answer (1 point), and a strategic response (2 points). This measure was based on empirical research of children's responses to metacognitive questions, so that it accurately reflected children's knowledge about reading strategies, rather than the authors' beliefs about what they know. Test-retest reliability (8 month interval) yielded a Pearson product-moment correlation of $r = .55$, $p \leq .001$ (Jacobs & Paris, 1987).

Investigation into the preliminary reliability and validity of the scale revealed no interpretable factors and a Cronbach's alpha of .61 (McLain et al., 1991). Furthermore, reliabilities of the four subscales (.15 to .32) were too low to support use of the subscales as separate scores for any analysis. Nevertheless, McLain et al. (1991) suggested that the scale was adequate in its present form for measuring metacognitive knowledge if used as a *total score* and only as one measure of the reading process.

In the present study, the wording of some items was slightly changed so that the item reflected a measurement of metacognitive knowledge rather than strategy use (see Appendix B, section B.2). For instance, the IRA item "If you could only read some of the sentences in the story because you were in a hurry, which ones would you read?" was modified to "If you could only read some of the sentences in the story because you were in a hurry, which ones *should* you read?" Analysis of internal consistency of the 20 items led to the deletion of items 1, 5, 9 and 15. This improved alpha considerably. A Cronbach's alpha of .68 was obtained for the whole sample; it varied from .54 for the RD sample to .60 for the NA sample. Scores for the remaining 16 questions on the IRA were combined to produce a total score ranging from 0 to 32 points (see Appendix B, section B.2).

The Metacomprehension Strategy Index (MSI), developed by Schmitt (1988, cited in Schmitt, 1990), is a 25-item, four-option multiple-choice questionnaire with one option representing metacomprehension awareness (1 point) and the other three options representing inappropriate responses (0 points). It was developed by Schmitt to evaluate young adolescent students' awareness of metacomprehension strategies before, during, and after reading a narrative prose. The MSI assesses students' metacognitive awareness within six broad categories: purpose setting; predicting and verifying; previewing; using prior knowledge; summarizing and applying fix-up strategies; and self-questioning. This measure has been used by a number of researchers (e.g., Baumann, Seifert-Kessell, & Jones, 1987; Horrex, 1992).

The MSI has been reported to have an internal consistency value of .87 using the Kuder-Richardson Formula 20 (Lonberger, 1988, cited in Schmitt, 1990). Statistically significant correlations were found between the MSI and the IRA ($r = .48, p \leq .001$). Furthermore, statistically significant correlations were found between the MSI and two comprehension measures commonly used to measure students' metacomprehension ability, an error detection task ($r = .50, p \leq .001$) and a cloze task ($r = .49, p \leq .001$).

Minor modifications were made to the wording of the MSI to improve the face validity of this questionnaire for Form 2 New Zealand students (see Appendix B, section B.1). The alterations acknowledged not only narrative prose reading, but also textbook/book reading. For example in item 2 (A), instead of "Look at the pictures to see what the story is about", the item was changed to "Look at the pictures or *diagrams* to see what the *chapter* is about". Items 2 and 20 were deleted for the present study to improve the reliability of the scale. Internal consistency of the measure for the current study yielded a Cronbach's alpha of .72 for the whole sample. However, internal consistency for the subsamples was lower, varying from .58 for the RD sample to .65 for the NA sample. Scores for the remaining 23 items on the MSI were combined to produce a total score ranging from 0 to 23 points.

Although Schmitt (1990) suggested that both the IRA and MSI instruments may measure similar constructs, the moderate correlation found between these tests suggests otherwise, as they have only 16% of shared variance. For the current study, these tests had 16% to 20% of shared variance across the groups. The MSI and IRA were selected for use in the present study as they are the only available paper-and-pencil tests, with known psychometric properties, currently used by researchers.

8.3.7 Strategy use

Considerable debate has arisen about how to effectively measure students' use of strategies (e.g., Ericsson & Simon, 1980; Meichenbaum et al., 1985; Meyers & Lytle, 1986; Nisbett & Wilson, 1977; Sternberg, 1985). Although this problem has not been satisfactorily resolved, as noted in Chapter 3 (see section 3.3), self-report questionnaires and verbal reporting are two most commonly used methods in measuring strategy use.

Although self-report questionnaires can be used effectively to measure student perceptions of cognitive engagement (Ames & Archer, 1988), researchers have suggested the need to replicate these results with other measures (Zimmerman & Martinez-Pons, 1990), particularly in the light of certain inherent limitations of the questionnaires. In the present study, strategy use was measured by using a questionnaire and a taped concurrent interview (think-aloud procedure). Given that both these measures have methodological limitations, employment of a multimethod assessment approach may produce converging measures of metacognition which may not share the same sources of invalidity (R. E. Reynolds & Wade, 1986; Schneider, 1985).

8.3.7.1 Self-report measures of strategy use

A self-report Reading Strategy Use (RSU) scale with subscales assessing metacognitive strategy use (RSU-meta) and cognitive strategy use (RSU-cog), was devised by the researcher for the purpose of the present study. Items were adapted from a number of published and experimental instruments and modified to make them more specific to reading. The development of and psychometric data for the 10-item cognitive and 12-item metacognitive strategy use subscales are reported in Pereira-Laird (1996) (see Appendix D). High reliability and validity of the subscales were demonstrated by the author. The correlation of the two subscales was .50 for the present study, suggesting that they are distinct constructs.

Students were instructed to respond to the items on a 7-point Likert scale in terms of their strategic behaviour. The 22 item scale had the following verbal descriptors: 1 - *never*, 2 - *almost never*, 3 - *seldom*, 4 - *sometimes*, 5 - *often*, 6 - *almost always*, and 7 - *always*. The 12 items included under the metacognitive strategies subscale (RSU-meta) contained items on planning, monitoring, and evaluation of the student's own cognition. Examples of these items included: "I stop once in a while and ask myself questions to see how well I understand what I am reading", and "I read quickly through the whole passage to get the general idea before

I read it thoroughly". The 10 items on the cognitive strategies subscale (RSU-cog) refer to those strategies students use to integrate new material with prior knowledge, that is, strategies that students use to learn, remember, and understand the material (Zimmerman & Martinez-Pons, 1986, 1988). They include rehearsal, elaboration, and organizational strategies; for example, "I learn new words by picturing in my mind a situation in which they occur", and "I make an outline of what I am reading".

Scoring was reversed on four items (3 for cognitive and 1 for metacognitive) that were negatively worded for strategy use, so that for all items a high score indicated a higher frequency in strategy use. Internal consistency of the scales for the current study ($N = 407$) was high for the whole sample, yielding a Cronbach's alpha of .93 for the RSU-meta subscale and .84 for the RSU-cog subscale. Although the reliability estimate for the cognitive strategy use subscale was lower than that for the metacognitive strategy use scale, they are in keeping with previous research (e.g., Pokay & Blumenfeld, 1990). One possible explanation for the variability in the reliability estimates for cognitive and metacognitive strategy use is that students may be more selective in their choice of cognitive strategies than in their choice of metacognitive strategies. In other words, reliabilities for metacognitive strategies may be higher because their use would be less situation- or task-dependent than cognitive strategies.

Three self-report strategy use scores were computed for each student. Metacognitive strategy use and cognitive strategy use scores were computed by calculating the mean for the items in each subscale. A combined strategy use score was computed by summing the mean scores of the RSU-meta and RSU-cog subscales.

8.3.7.2 Taped concurrent interview: Behavioural measure of strategy use

A second method of assessing students' strategy usage is verbal reporting. Research seems to support the use of concurrent reporting (*think-aloud* protocols) in preference to retrospective reporting as being a more reliable measure of students' strategy usage (Ericsson & Simon, 1980). While there is greater support for the think-aloud concurrent procedure, a *structured* approach of a think-aloud procedure was thought to yield more information about strategy usage considering the verbal reticence of RD students. One such procedure is known as the Think-Aloud Reading Strategy Use Measure (TARSUM; Pereira-Laird & Deane, 1996). The TARSUM was an adapted version of Paris' (1991) Think-Along-Passage (TAP) procedure. The development of TARSUM involved a number of pilot tests which were

conducted to refine and adapt Paris' interview measure to the present study. The changes involved the incorporation of measures of metacognitive strategy use and a monitoring device, revision of the scoring procedures, and the creation of textual material of three graduated levels of difficulty. These changes are reported in detail in Pereira-Laird and Deane (1996) (see Appendix E). The high reliability and validity of this measure was demonstrated by the aforementioned authors. The final version of the TARSUM is now described.

Three instruments, namely a reading passage, a monitoring device called a "bleep" (van Kraayenoord, 1986), and an interview coding form, were required to conduct the taped concurrent interview which was designed to measure students' strategy usage. Textual material of three graduated levels of difficulty and suitable for reading orally was designed for this study. Students were required to read aloud a passage (appropriate for their reading age) which was tape recorded by the interviewer. While reading, students were stopped at designated points by the interviewer who asked questions that probed their comprehension or metacognitive/cognitive strategy use. Interviewers recorded students' answers to the comprehension and cognitive/metacognitive strategy use questions on the interview coding form. Students used a device called a "bleep" to indicate awareness of monitoring their comprehension (by pressing a button).

(i) *Passage*

Three criteria determined the selection of the passage. First, the passage was required to be of reasonable difficulty and length so as to provide students with the opportunity to generate more strategic responses. A related issue was the confounding effect of task difficulty on the generation of strategic responses. Several researchers have noted that oral reading behaviour alters significantly when children read at different levels of text difficulty (Blaxall & Willows, 1984; Ng, 1979, A. St. George et al., 1985; van Kraayenoord, 1986). In particular, research has indicated that flexibility of strategy usage decreases for RD readers as reading material increases in difficulty (Blaxall & Willows, 1984). Thus, it was considered important to equate difficulty level across groups by having students read at their own "moderately difficult" level. Second, the passage selected was required to be typical of the type of comprehension exercises frequently used in most reading classrooms at the Form 2 level (junior high). Third, the passage was required to be about an unfamiliar topic. This was an important consideration as prior knowledge could negate the need for students to employ a

range of strategies when attempting to comprehend the passage and answer the accompanying questions.

The passage, entitled "The Volcano", was obtained from the Neale Analysis of Reading Ability Manual (Neale, 1989) and was considered appropriate for Form 2 children since it has been used by other New Zealand researchers (e.g., van Kraayenoord, 1986). Furthermore, this test was not used by participating schools and hence the students had not been exposed to the passage before.

Three versions of the passage were generated and rewritten at three different readability levels, in order to exercise some control over task difficulty (see Appendix B, section B.3.1). Prior to generating the versions, six textual irregularities were inserted to measure metacognitive strategy use and a *third* paragraph was appended so as to lengthen the passage. The textual irregularities consisted of two textual inconsistencies, two nonsense words, and two scrambled words. Essentially the story line was the same for all the versions. Every attempt was made to include all information in the three versions and to maintain a comparable length. As a result, length ranged from 211 to 243 words with small differences in the number of idea units expressed (see Appendix B, section B.3.1). In creating the versions, some of the words and sentences from the original passage were rephrased and modified to reduce difficulty of the passage. Consequently, the three versions of the passage differed only in difficulty of the vocabulary words, with version 1 being the hardest and version 3 being the easiest.

The versions were created so that students reading their version found it moderately difficult rather than too difficult or too easy. Previous researchers have used 95% or above in accuracy and 75% or above in comprehension for the easy passage (Ng, 1979; Pohl, 1981; van Kraayenoord, 1986), and 90% or below in accuracy and 50% or below in comprehension for the difficult passage (Johns, 1981; A. St. George et al., 1985; van Kraayenoord, 1986). For this study, a set of criteria had to be determined for a moderately difficult passage. It seemed reasonable to set the criteria somewhere in between those set for the easy and difficult passage. Hence, for this study, the criteria were set at above 90% in accuracy and above 50% in comprehension to be considered moderately difficult. The accuracy score consisted of the proportion of correctly read words in a passage expressed as a percentage of the total number of words. Errors which had been successfully corrected were not included in the error count.

As it was impractical and time consuming to determine the reading age and the appropriate version to be used for each student in the main study, it was decided to administer the versions according to some preset criteria. The first version was administered to students who scored at or above the 70th percentile for PAT Reading comprehension, the second version to students who scored at or above the 45th percentile but below the 70th percentile for PAT Reading comprehension, and the third version to students who scored below the 45th percentile for PAT Reading Comprehension. The high reliability of this procedure was demonstrated by Pereira-Laird and Deane (1996) (see Appendix E, Study 2).

(ii) Monitoring device (bleep)

An unobtrusive measure of monitoring was sought to determine students' awareness of errors in the passage while reading, with minimal disruption in the flow of reading. It was considered important that such a measure should only minimally interfere with the cognitive processing undertaken during reading. Furthermore, it should have validity in that it was an acknowledgement of the student's own awareness of monitoring rather than an inference of monitoring that the researcher made on the basis of puzzled looks or the like.

The bleep was considered a suitable instrument which might fulfill the requirements, and has been used by a number of researchers to investigate on-line monitoring (e.g., van Kraayenoord, 1986). It consists of a red button placed on top of a small rectangular container. When the button is pressed an audible sound is heard. The students were asked to press the button or "bleep" when they came to any errors (textual inconsistencies, scrambled words, or nonsense words) in the text. This allowed for subsequent coding of monitoring behaviour with minimal disruption of students' flow of reading or cognitive processing.

(iii) Interview coding form

The interview coding form was used by the trained interviewers to record behavioural data on strategies by following the TARSUM, which involved students reading aloud to interviewers who asked predetermined questions interspersed along the passage. The questions were asked at designated points in the passage (indicated by an asterisk); they not only probed comprehension, but also queried cognitive strategy use and metacognitive strategy

use. The TARSUM consisted of eleven comprehension, twelve cognition (TARSUM-cog), and nine metacognition (TARSUM-meta) questions. These questions were contained on the left hand side of the interview coding form with several strategies listed opposite each cognition and metacognition question so that the spontaneous strategies used by the students could be recorded (see Appendix B, section B.3.2).

The comprehension questions (part a of questions) assessed students' ability to identify the theme of the "Volcano" passage, to make a prediction or an inference within the passage, to decipher the meaning of a unique word, and to provide a concise summary of the passage. The maximum score obtainable for each comprehension question is found on the interview coding form, beside the question.

The cognition questions that followed students' answers to each comprehension question were designed to reveal the **cognitive strategies** that students used. Examples of cognitive strategies that were examined included: rereading, predicting based on context clues, and using prior knowledge. The cognition questions (part b of questions) were usually in the form "How did you know this?", "Why do think that?", "How could you tell?", or "How did you get your answer?" The question "How did you go about looking for the answer?" was used for students who reported not knowing the answer to a comprehension question. This question was designed to reveal the cognitive strategies students used in their attempt to locate an answer.

For each cognition question, interviewers assigned the response one point for each appropriate cognitive strategy verbalised by the student. Subjects were not awarded any points for appropriate strategies that were observed by the interviewer but not verbalised by the subject (e.g., scanning text, looking at title, rereading silently). This avoided the interviewer making subjective inferences.

The metacognition questions assessed two types of activities involved in **metacognitive strategy** use: monitoring and planning. Evaluation strategies, which are also a form of monitoring, were also measured. Monitoring questions were asked at designated points where the previous segment just read contained textual inconsistencies, nonsense or scrambled words. Planning and evaluation questions were asked after the passage was read.

To assess monitoring strategies, the passage was modified to include two sets each of textual inconsistencies, scrambled words, and nonsense words. Such a method has been used quite extensively by researchers to monitor comprehension (Bos & Filip, 1984; Zabrocky & Ratner, 1986). Students used a monitoring device (a bleep) to indicate any errors and disruption of understanding as they read the passage. Such a device allowed comprehension monitoring to be studied without disrupting the student's flow of reading.

For the monitoring questions, students were asked "Why did you bleep?" or "What are you thinking here?" depending on whether they bleeped. (The second question was asked to take into account students who forgot to bleep but who may have been aware of the error in the text). Two of the six monitoring questions were followed by prompts. The monitoring questions assessing the first textual inconsistency and nonsense word were followed by prompts ("Do you agree with the sentence?" or "Is there such a word as urgsy?") for students who did not show awareness. Irrespective of whether students bleeped or not, two points were awarded for every verbalisation of awareness and correction by the student of scrambled words or textual inconsistency without prompting. Students were awarded only one point if they indicated awareness that something was wrong but were not able to correct the error or showed awareness and correction only after being prompted. One or two points were awarded for a verbalisation of awareness of the nonsense words depending on whether students were prompted (see Table B1 in Appendix B, section B.3.3).

The planning activity of metacognitive strategy use was assessed by an evaluation of the students' recall and summary of the story. A student who has planned the answer well is likely to deliver the main points and organize the ideas in recall. It is unlikely that the student is metacognitively aware if he or she knows that relating main points is a useful strategy but uses it indiscriminately by mentioning the irrelevant points and details, or presents the points in a disorganized manner. Hence, students were awarded one point if they had listed most of the main points in an organized manner. An answer was considered planned if the key points related by the student followed in a similar and logical order to that presented in the story.

To assess evaluation activities of metacognitive strategy use, two questions were generated. The first assessed students' subjective evaluation of difficulty of the passage they had read, and the second assessed students' evaluation of the accuracy of their answers to the comprehension questions. The first evaluation question asked students which of the three

paragraphs they found hardest to read. An objective measure of student's difficulty of each paragraph was determined by calculating students' accuracy in reading each paragraph. One point was awarded if the paragraph mentioned by the student was read less accurately than the other paragraphs. If students had similar accuracy across paragraphs, mean comprehension scores (converted to z-scores) for each paragraph were used as an alternative criterion. The second evaluation question asked students to indicate which of the 11 comprehension questions they had answered correctly. Students were awarded one point for each correct evaluation and zero points for incorrect evaluation. Due to the limited time available during the interview, judgements as to whether subjective and objective evaluations matched were made by the primary researcher after the interview, with interrater reliability subsequently assessed by an independent rater (see section 8.4.3.3).

In summary, the TARSUM was used to provide a second measure of strategy use. The questions, found on the interview coding form, were designed to probe comprehension as well as cognitive/metacognitive strategy use. The cognition questions, which usually followed the comprehension questions, were designed to examine a variety of cognitive strategies (e.g., predicting, using context clues). A number of indicators of metacognitive strategy use were also examined: (a) monitoring understanding when faced with textual inconsistencies, (b) monitoring errors and meaning when faced with scrambled words, (c) monitoring meaning when faced with nonsense words, (d) planning the summary of the story, (e) evaluating the difficulty of the paragraphs, and (f) evaluating how successful students were at answering the comprehension questions of the passage.

Three strategy use scores were computed for the interview measure. A cognitive strategy use score was computed by adding the scores of the twelve cognition questions. A metacognitive strategy use score was computed by adding the scores of the nine metacognition questions. Combined strategy use was calculated by summing the cognitive and metacognitive strategy scores.

8.4 Procedure

8.4.1 First Phase

The first phase involved sample selection (details presented under section 8.2). Parental, child, and school consent were obtained for students who were involved in the initial

screening and for students who were selected to participate in the study. Both parents and students were informed that the students would be questioned about their reading practices. Information and consent sheets are provided in Appendix A.

8.4.2 Second Phase

The second phase was conducted within one week of the first phase. This phase involved administering two self-report questionnaires, Reading Questionnaire I (RQI) and Reading Questionnaire II (RQII), which examined adolescents' attributional style, self-efficacy, metacognitive knowledge, and use of cognitive/metacognitive strategies in the domain of reading (see Appendix B, sections B.1 and B.2). Testing took place during two 45 minute sessions with both questionnaires administered within a period of one week; RQI was administered in the first session and RQII was administered in the second. Students in each school were tested in groups of 15 to 20, irrespective of their reading ability. The questionnaires were read out to each group by the researcher for each session to ensure standardized presentation and to minimize students' reading difficulties. Students were instructed to raise their hands if they had difficulty in understanding an item or wanted a particular item repeated. The students were continuously observed in order to ensure that they were comprehending what was being read out and that they were attentive during testing. At the end of each questionnaire session, students were instructed to check that all items had been answered. After the questionnaires were returned the researcher also ensured that all items were answered by each student. When an item was missed, the researcher would approach the student on the same or next day for it to be completed, provided the student was willing.

In RQI, the cognitive/metacognitive strategy use (RSU) and self-efficacy (GRSE) items were presented in section 1. The items were presented in a random order to avoid response set. The MSI Scale was presented in section 2 followed by the CDS in section 3. The self-report strategy use measure was presented before the metacognitive knowledge items so as not to cue students about potentially useful strategies.

Before section 1 was administered, students were given a short training sequence to familiarize them with the use of the 7-point Likert scale. The instructions (for example, "there are no right or wrong answers to these questions") and sample items were designed to

encourage students to use the full range of responses on the questions. The sample item "I really like spinach", where many children endorsed strongly the "negative" end of the Likert scale and where the researcher endorsed the "positive" end, was included to impress upon the students not to give socially desirable responses and illustrated the idea that there were no right or wrong answers to these questions (Assor & Connell, 1992).

In the second section, a sample item was used to illustrate the response format (multiple-choice) of the metacognitive knowledge scale. Students were instructed to put a circle around the letter in front of the best answer to the question. For the third section, students were familiarized with the format of the scale. Students were instructed to think of two to three main reasons why they were successful/unsuccessful in reading and then to rate the extent to which these causes, in general, fell on the attributional dimensions.

In administering RQII, students were again reminded to use the full range of the response scale. Students were again familiarized with the question formats in each section of RQII before proceeding to answer the items. The task-specific measure of self-efficacy (TRSE) was presented in section 1. Efficacy was measured using a 100 point scale with 10-unit intervals. Students initially received practice by judging their certainty of successfully jumping progressively longer distances. In this concrete fashion, students learned the meaning of the scale's direction and the different numerical values. The IRA, a measure of reading metacognitive knowledge, was presented in section 2.

8.4.3 Third phase

The third phase was conducted within two weeks of the second phase. In this phase, one-to-one interviews following the procedure of TARSUM, which typically took 25 minutes to complete, were conducted to provide a second measure of students' use of strategies. The interviews with each student, were conducted by the researcher and four trained interviewers who were all postgraduate psychology students. The interviewers were familiarized with the procedure, notation (for recording oral reading behaviours), and the strategy types (see Appendix B, section B.3.4, Table B2, for definitions of each strategy type). Detailed verbal explanations were provided to the interviewers and any points of confusion were clarified. The interviewers then listened to a recording of one of the interview sessions (which was pilot tested with a student not included in the study) and were asked to transcribe the session onto

the interview coding forms. Points of confusion and disagreement were discussed and clarified.

8.4.3.1 Training session for using the bleep

Students came to the interview room individually. An attempt was made to initially set them at ease through brief conversation about recent school activities. Students were then given a short training session to familiarize them with the use of the monitoring device ("bleep").

Instructions:

Today I want to hear you read a passage. While you are reading, you will be using a "Bleep" [Press the bleep button].

This machine is a "Bleep". I want you to place this on the table close to you and keep one finger on the button. (Show the child the button. Comment that it's like a Space Invaders game). While you are reading aloud, I want you to push the button when something is not right with the sentence. (Give the sentences without the asterisks to the student. Ask the student to read aloud).

Sentence 1: One day a boy went out for a walk. On his way, he saw a dog. The dog began to him* chase. The boy started to run away from the dog. The dog boinkrad* angrily at the boy. The boy was frightened of the cat*. Luckily, his friend came to help him with wooden stick a big*.

[The button should have been bleeped at the points marked with an asterisk.]

If the student did not bleep at the appropriate points, the student was told where he/she should have bleeped and why. The interviewer then demonstrated this by reading aloud and bleeping at the appropriate points. Students were then asked to try a further two sentences.

Sentence 2: The lady looking* was pet for her daughter. She thought of buying a cute little puppy. She paid \$12.00 for the puppy. Her daughter was happy with the kitten* and soinked* after it well.

Sentence 3: The lady liked to pick flowers. Everyday she croit* flowers for her garden. Sometimes she the* gave flowers to a friend. Her enemy* liked the flowers too.

[The button should have been bleeped at the points marked with an asterisk.]

During each trial the interviewer observed whether or not the button was pressed at the appropriate points. If the button had not been pressed, the student was asked to consider the sentence again and indicate where a bleep should have been made. Students were also asked what was wrong at each point. The interviewer ensured that students knew when to press the bleep before proceeding to the next step.

8.4.3.2 Interview

Following the training session, for each student, the interviewer placed the appropriate version of the passage in front of the student. The interviewer had a similar passage to refer to while the student was reading. Students were then instructed to stop at designated points indicated by an asterisk on the passage, where they were asked questions designed to reflect their comprehension and use of strategies.

Instructions:

I want you to read this passage out loud. I will be tape recording our voices as I will not be able to remember everything you say. Please don't be nervous.

The passage will have some asterisks like this (show the student) at the end of some sentences. I would like you to stop at these points as I will be asking you some questions. If you come to a word that you don't know try your best. I cannot help you. Remember this is not a test, so you can look back at the passage when you are answering the questions, if you wish. Don't forget to use the bleep when you come to something that is wrong. Please keep your finger on the button while you are reading. Right let's start. [Switch on tape recorder].

While the student read each segment of the passage, the interviewer indicated on his/her copy of the passage the points where bleeps, errors, successful and unsuccessful self-corrections were made by the student. Each time the student bleeped for a segment of text read, the

interviewer would ascertain the student's reasons and record them on the interview coding form.

While the student answered the prescribed questions, the interviewer scored the student's answer to the comprehension question and recorded the presence of each strategy (by placing a tick in the list given) articulated by the student. If the strategy mentioned was not listed, it was described in the "other" category. Students were instructed to continue reading the next segment after answering the questions at each designated point. If they failed to stop at the designated points, they were reminded to do so.

Encouragement was given for effort made and if a student was fully aware that he/she had done poorly on a question, the interviewer would say, "Many students find this one hard, don't worry". However, no indication was given as to whether their answers to the comprehension questions were right or wrong so as not to invalidate subsequent attempts to evaluate their responses to the comprehension questions. If the student's answers were unclear, the interviewer would request more detail with prompts such as "How do you mean?" or "Tell me more about it".

For each cognition and metacognition question, the interviewers recorded the presence and categorisation of each strategy as it occurred while questioning the student. All interview data were coded and scored by the researcher to obtain a total score each for comprehension, cognitive strategy use, and metacognitive strategy use based on the transcriptions and tape recordings.

8.4.3.3 Reliability

Fifty percent of each interviewer's coding forms ($n = 198$) were randomly subjected to reliability checks by the researcher. Reliabilities for occurrence and categorisation of strategies were calculated using the point-by-point percentage agreement method (Cooper, Heron, & Heward, 1987). Interrater agreement was calculated on a question-by-question basis by dividing the number of agreements by the number of agreements plus disagreements, multiplied by 100.

Mean interrater agreement for the occurrence and categorisation of cognitive strategies was 99.4% (range 98.2% - 100%) and 90.6% (range 87.0% - 93.5%), respectively. Mean interrater agreement for the occurrence ($M = 100\%$) and categorisation ($M = 98.1\%$, range

= 96.9% - 100%) of the monitoring and planning strategies was higher than for cognitive strategies. Interrater agreement for combined cognitive and metacognitive strategy use (monitoring and planning strategies only) was also calculated. Interrater agreement was 99.7% (range = 99.1% - 100%) for the occurrence of a strategy, and 94.4% (range = 93.5% - 95.5%) for categorisation of a strategy. Interrater agreement was also sought for comprehension scores, bleeps, and for oral reading behaviours. Similar to procedures used in van Kraayenoord (1986), interrater agreement on oral reading behaviours consisted of the number of uncorrected errors and the number of successful and unsuccessful self-corrections where there was agreement as to the site of the error or self-correction. Mean interrater agreement was 95.7% (range = 94.9% - 97.0%) for comprehension scores, 100% for bleeps, 97.5% (range = 96.4% - 98.3%) for successful self-corrections, 97.6% (range = 97.1% - 98.4%) for unsuccessful self-corrections, and 98.1% (range = 96.9% - 99.0%) for uncorrected errors. The percentage agreements for the oral reading behaviours compare favourably with reliability in other New Zealand studies (Clay, 1973; van Kraayenoord, 1986). In addition, the finding that more bleeps occurred at expected points (83.4%) provided a validity check that students were using the bleep correctly.

Reliability checks for the aforementioned categories were again made using an independent rater, a postgraduate psychology student who was not one of the original interviewers. Interrater agreement on the evaluation activities of metacognitive strategy use was also collected. The independent rater received the same training as the other interviewers.

Ten percent of the 396 recordings ($n = 40$) were randomly selected for the reliability check. Comparisons of agreement were made on several variables using the reliability formula as used previously. The percentage agreement between raters is as follows: 99.0% and 93.9% for the occurrence and categorisation of cognitive strategies, respectively; 100% and 98.3% for the occurrence and categorisation of metacognitive strategies, respectively. In addition, interrater agreement for combined strategy use was calculated; 99.5% for occurrence of a strategy, and 96.1% for categorisation of a strategy. Interrater agreement for oral reading behaviours included: 95.8% for comprehension scores, 98.2% for successful self-corrections, 97.6% for unsuccessful self-corrections, and 99.2% for uncorrected errors.

Procedural validity was also assessed by the researcher and the independent rater in terms of the appropriateness of the interviewer's verbalizations during the taped section of the interview. Interviewer verbalizations were categorized into correct prompts, incorrect prompts and other verbalizations. Incorrect prompts occurred when (a) student's comprehension answers were prompted or corrected, (b) the interviewer gave some indication that the student

had answered the question correctly, (c) the interviewer prompted the student as to type of strategy that had been used, or (d) the interviewer had prompted the student to bleep at the appropriate points. Other verbalizations referred to student verbalizations, and interviewer verbalizations which were not related to the TARSUM.

Each verbalisation was categorised and recorded using 15-s intervals in a partial-interval recording procedure (Cooper et al., 1987) for five minutes of 10% of the audiotaped sessions ($n = 40$). The five minute segments were randomly selected from each 25 minute interview session. Rater agreement was calculated on an interval-by-interval basis using point-by-point agreement ratio method (Cooper et al., 1987). High validity was indicated as 97.1% of the interviewer verbalizations were correct, 1.9% were incorrect, and 1% were other verbalizations. There was 100% interrater agreement on the occurrence of interviewer verbalizations.

8.5 Data analyses

Data were analyzed in several steps. First, after checking for assumptions of multivariate normality, univariate statistics for each variable were computed. In addition, the degree of relation between variables was calculated using Pearson product-moment correlations. Second, a discriminant analysis was performed to determine whether RD could be differentiated from NA readers on the basis of the following variables: success/failure attributional dimensions, self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use. The analysis also provided validation data for the multidefinitonal procedure used in this study for selecting RD adolescents.

Third, structural equation modelling procedures based on the analysis of covariance structures were used to test the plausibility of four postulated causal systems presented in Figures 4 through 7 (see Chapter 6). More specifically, multisample analyses were conducted with the LISREL 8 software package (Jöreskog & Sörbom, 1993). This constituted the principal method of data analysis. In using structural equation modelling procedures, researchers must clearly specify the assumptions made in formulating both the measurement and structural models. In the present study, each factor in the measurement model was formed using indicators described in section 8.3. In addition to testing each general hypothesized model, a number of competing models were evaluated with each general model as suggested by MacCallum (1986) and J. C. Anderson and Gerbing (1988). Several statistics were used to evaluate model fit.

CHAPTER 9

RESULTS

This chapter is divided into two main parts. The first part has two sections which comprise a descriptive and correlational analysis of the data. This is followed by discriminant analysis of the data in order to determine the validity of the selection procedures in defining reading disabled adolescents. This analysis was also conducted to identify which predictors combined effectively to distinguish between normal achieving and reading disabled adolescents. The second part of the chapter presents multisample LISREL analyses of the four hypothesized models with normal achieving (NA) and reading disabled (RD) students. It represents the main analyses of this dissertation. Initially, a description of the preliminary LISREL analyses is provided, followed by a description of the analyses for the measurement and structural parts of the four hypothesized models. Supplementary analyses of the four hypothesized models were also conducted and are presented at the end of this chapter.

9.1 Descriptive statistics and correlations

All variables were examined for missing values and for normality distributions for assumptions of multivariate analysis. The variables were examined separately for the RD and NA reader groups. A small number of missing cases occurred in the behavioural strategy use and the classroom reading performance measures. No missing values were found on the self-report measures. Missing values were dealt with through listwise deletion of cases on each analysis. Listwise deletion led to the omission of 2 cases for analyses of Model 1 resulting in 203 cases for NA readers and 202 cases for RD readers. For analyses of models 2, 3, and 4, listwise deletion led to the elimination of 10 cases resulting in 199 cases for NA readers and 196 cases for RD readers. For LISREL analyses, these sample sizes were considered to be moderate and are close to Boomsa's (1985, 1987) and Hayduk's (1987) recommendation of 200. Furthermore, the ratio of cases per parameter estimated (6.8:1 for Model 1; 5.3:1 for Model 2; 4.2:1 for Model 3; and 4.8:1 for Model 4) for each model were close to the 5 cases per parameter guideline suggested by Bentler and Chou (1987).

Evaluation of univariate distributions and outliers for each group indicated that all variables met univariate normality of distributions, except for 2 variables (WISC-R and PAT Reading

Comprehension) in the RD reader group which were moderately kurtotic. As a result of selection procedures employed in this study, it was expected that there would be departures from normality for these measures in the RD reader group considering that these readers generally fall at the lower end of the normal IQ range and have low scores in reading. However, maximum likelihood solutions are known to be fairly robust to departures from multivariate normality (Bollen, 1989). As such, these variables did not warrant transformation.

Evaluation of multivariate assumptions of normality was achieved by examining the properties of the residuals using multiple regression analysis. More specifically, residual analyses were conducted when all variables were regressed on each of the achievement measures. Assumptions of normality of distribution, homoscedasticity and independence of residuals were all met and there was no evidence of curvilinearity. Examination of standardized residuals indicated there were no multivariate outliers. These data demonstrated multivariate normality and thus met one of the assumptions underlying the use of maximum likelihood estimation of latent variables.

Means, standard deviations, and reliability estimates (as measured by Cronbach's alpha) for all the indicators are shown separately for each group in Table 5. Examination of Table 5 shows four main features, described below:

1. On average RD readers scored consistently lower than NA readers across all measures. NA readers tended to be more efficacious, adopted an adaptive attributional style with regard to success and failure outcomes, were knowledgeable about reading and engaged in more strategy use than RD readers. Independent between group *t* tests indicated that in all cases the observed differences were statistically significant ($p \leq .0001$).
2. While standard deviation estimates were generally comparable across groups there is evidence that for some measures, particularly task-specific efficacy, PAT reading comprehension, and IQ, the variances were more constrained in the RD reader group. This is predictable given the group selection procedures that were used.
3. Reliability estimates (as measured by Cronbach's alpha) of each reader group suggest that all measures were of moderate to high reliability with alpha values ranging from .54 to .89. As expected, the metacognitive knowledge measures had somewhat lower alpha

coefficients since they measured different aspects of metacognitive knowledge for reading. The reliability estimates of the metacognitive knowledge measures are similar to those obtained in previous studies (e.g., McLain et al., 1991). It is felt that such reliabilities provide further justification for the use of structural modelling with latent variables in that this procedure directly addresses the use of unreliable measures (Hull & Mendolia, 1991).

Table 5. Means and standard deviations of variables in the NA and RD reader groups

LATENT CONSTRUCT with indicators	code name	NA readers			RD readers			overall reliability ^a
		<i>M</i>	<i>SD</i>	reliability ^a	<i>M</i>	<i>SD</i>	reliability ^a	
ATTRIBUTIONS								
Success								
Internality	inters	16.17	2.80	.65	9.20	3.02	.71	.87
Controllability	contros	16.41	3.23	.81	7.33	2.69	.65	.92
Stability	stabs	15.94	3.23	.79	7.86	2.57	.65	.91
Failure								
Internality	interf	14.32	2.84	.78	10.49	3.95	.84	.93
Controllability	controf	16.39	2.56	.68	7.32	2.99	.78	.93
Stability	stabf	16.59	2.59	.72	8.34	3.19	.81	.92
EFFICACY								
General	effica	5.36	0.93	.87	3.07	0.64	.64	.93
Task-specific	efficb	68.19	16.18	.89	30.09	11.77	.83	.95
METACOGNITIVE KNOWLEDGE								
Measure 1	msi	13.13	3.70	.65	8.96	3.45	.59	.72
Measure 2	ira	24.87	3.61	.57	19.95	4.32	.54	.68
STRATEGY USE								
self-report metacognitive	meta	4.75	0.89	.82	2.63	0.68	.76	.93
behavioural metacognitive	metb	18.43	3.89	-	15.32	3.92	-	-
self-report cognitive	coga	4.69	0.78	.71	3.34	0.81	.74	.84
behavioural cognitive	cogb	12.05	1.67	-	9.77	2.13	-	-
self-report combined ^b	strat1	9.44	1.45	.84	5.96	1.28	.82	.94
behavioural combined ^b	strat2	30.47	4.85	-	25.08	5.25	-	-
PERFORMANCE								
PAT Comprehension	pat	63.89	18.53	-	20.60	12.79	-	-
classroom	class	67.10	14.93	-	37.64	15.17	-	-
ABILITY								
WISC-R	iq	102.95	9.17	-	95.58	8.12	-	-

Means and standard deviations were based on listwise deletion; *n*=199 for NA readers and *n*=196 for RD readers.

^a Cronbach's alpha was estimated for most of the self-report measures; however, item level data was not available to calculate reliability for the performance and ability measures. Reliability data for the behavioural measures is provided in the form of interrater reliability (see Chapter 8, section 8.4.3.3).

^b Combined strategy use refer to the global use of both cognitive and metacognitive strategies.

4. The internal reliability estimates for the whole sample, for all measures except for metacognitive knowledge, were very high ranging from .84 to .95. These reliabilities of the manifest variables were higher than those estimated for the subsamples. This discrepancy in the reliability estimates could be explained as follows. Since the reliability of a test reflects both real individual differences and measurement fluctuations, if everyone were alike, the only measurable differences among them would be due to error variance (J. R. Graham & Lilly, 1984). Thus, for example, the same attributions measure when used with a highly selected group should have a lower reliability when students are more similar (i.e., there are few individual differences) than in a group where less restriction of range of real individual differences is found. Comparison of reliability estimates across groups indicated that, with the exception of failure attributions, the reliability estimates for the indicators were generally higher for NA than for RD readers.

The matrix of correlations among the measured variables for each of the models is presented separately for each group in Tables F1 - F3 in Appendix F. These correlations present in standardized form the covariance data used in the LISREL analyses. All correlations between attributions, efficacy, metacognitive knowledge, strategy use, and performance measures were statistically significant ($p \leq .005$). In Table F2 mean intercorrelations were .62, .60, .50, .66 and .68 respectively for the indicators of attributional style, efficacy, metacognitive knowledge, combined strategy use, and performance for NA readers. Mean intercorrelations for RD readers were .52, .47, .46, .65 and .60 respectively for the indicators of attributional style, efficacy, metacognitive knowledge, combined strategy use, and performance. In general, there were reasonably high intercorrelations among the indicators of the same construct which suggests convergent validity, whereas the intercorrelations among indicators of different constructs were generally moderate in size, which suggests discriminant validity. For instance, in each group, the intercorrelation between the indicators of combined strategy use was higher ($r = .66$ for NA readers, $r = .65$ for RD readers) than the intercorrelations of these indicators with indicators of other constructs ($r = .30 - .61$ for NA readers, $r = .13 - .59$ for RD readers).

In general, Tables F1 through F3 indicate that the *majority* of the correlations were moderate in magnitude and statistically significant. Thus, considerable common variance existed for the structural analyses. The correlations amongst the manifest variables for RD readers were weaker than those for NA readers. The correlations of the performance measures are worthy

of note. A methodological problem with research relating classroom performance scores to standardized scores for reading is that there is usually less error in the standardized scores. This is because schools vary in the level and complexity of tasks assigned to students. Furthermore, tasks assigned to RD readers are likely to be easier than those assigned for NA readers. Thus, for example, in the present study, the classroom measure showed lower correlations with the other variables as compared to the standardized PAT scores for both groups. Furthermore, classroom performance of RD readers generally correlated less strongly with other variables when compared to NA readers.

Another finding that was notable was that the correlation between ability and reading comprehension was much lower than expected when compared to previous research of L. J. Horwood (personal communication, August 23, 1994). In Horwood's research, the observed correlations between the full-scale WISC-R scores and PAT reading comprehension was in the region of .60, compared to the correlation of .40 for NA readers and .21 for RD readers observed in the present study. There may be two explanations for this. First, the reliability of the short-form WISC-R used in the present study is lower than that reported for the full scale IQ, and this would have the effect of attenuating the observed correlations between the WISC-R and PAT scores. Second, the low correlations may be a function of the selection procedures that were used in the present study. One of the selection criteria involved choosing students if they had IQ scores of 85 or above. This procedure would have had the effect of constraining the variance of the ability measure within groups, and thus of producing lower than expected correlations with the performance measures (cf. Prior et al., 1995). The lower correlation between IQ and performance for RD readers was expected, as empirical evidence has indicated that the aforementioned relationship is severely attenuated for RD children (S. Graham & Harris, 1989; Stanovich, 1986a; Swanson, 1991b).

9.2 Validation of selection procedures in defining RD students

Before testing the four hypothesized models, it was considered necessary to determine whether the multidefinitional assessment approach used in the present study was valid in distinguishing RD and NA readers. If classification is valid, then predictable differences between the two groups should emerge on variables not used to form the groups. In general, research has indicated that NA and RD readers differ in their motivational, cognitive, and metacognitive characteristics. Consequently, a stepwise discriminant analysis was performed to assess

whether NA readers could be differentiated from RD readers on the basis of the means of 6 variables explored in this study: attributions for success, attributions for failure, self-efficacy, metacognitive knowledge, metacognitive strategy use and cognitive strategy use. Only variables that were not used in the sample selection procedures were included. In order to remove the effect of differing scales of the discriminating variables from the model, standardized scores were used. As multiple measures were available for each construct, they were aggregated to form one variable for each construct. The dimensions of internality (locus), controllability and stability for success attributions were summed and divided by three to provide an aggregate mean score for success attributions. A similar procedure was used for failure attributions. Aggregate mean scores of self-efficacy, metacognitive knowledge, and cognitive/metacognitive strategy use were formed by summing the scores of their respective indicators and dividing by two.

A significant discriminant function was found (Wilks $\lambda = 0.21$, $\chi^2(5, N = 395) = 608.92$, $p \leq .0001$) that explained 79% of the variance. From a multivariate perspective, 5 of the 6 discriminant variables were significant when a stepwise discriminant analysis was conducted. Table 6 shows the variables which discriminated most significantly between NA and RD readers along with their associated standardized coefficients. A combination of motivational beliefs, metacognitive knowledge, and cognitive strategy use were necessary in differentiating NA from RD readers.

Table 6. Results of discriminant analysis

Discriminating variable	Wilks' lambda	standardized canonical discriminant function coefficient
attributions for success	0.29	0.43
self-efficacy	0.24	0.56
attributions for failure	0.22	0.41
metacognitive knowledge	0.21	-0.28
cognitive strategy use	0.21	0.06

All discriminating variables were significant at $p \leq .0001$.

Examination of the standardized canonical discriminant function coefficients indicated that motivational beliefs, metacognitive knowledge, and cognitive strategy use were predictive of the discriminant function. Efficacy and success/failure attributions were more significantly predictive of the discriminant function than metacognitive knowledge and cognitive strategy use. Moreover, these motivational variables were similar in their discriminating strength. The

standardized canonical discriminant function coefficients take into account the intercorrelations of the variables and provide predictive information about good and poor performance in reading.

Results of the classification analysis indicated that 95.20% of the students could be accurately classified on the basis of the single discriminant function (see Table 7). This model displayed considerable accuracy in classifying NA and RD readers.

Table 7. Classification results from discriminant function analysis: Group membership percentages predicted by the function

Actual Group	No. of cases	Predicted Group Membership	
		NA	RD
NA	199	93.5%	6.5%
RD	197	3.0%	97.0%

All assumptions for this analysis were met except for the equality of covariance matrices across groups. The analysis yielded a significant Box's M ($p \leq .0001$). However, considering that the Box's M test is sensitive to departures from normality and that consistent results regarding the discriminating variables were obtained across the multivariate statistics, the violation of the assumption of equality of covariance matrices may not present much concern.

In summary, it can be concluded that metacognitive, cognitive, and motivational factors combine effectively to distinguish between NA and RD readers. These findings provide support for the validity of the multidefinitonal approach that was employed in this study for defining RD adolescents.

9.3 Model testing

Four models, each with two groups, were developed to compare the relationships amongst motivational, cognitive, and metacognitive characteristics of NA and RD readers in predicting reading comprehension performance. The models were tested using the most recent version of LISREL, LISREL 8 (Jöreskog & Sörbom, 1993), and methods of maximum likelihood estimation. The use of the most recent version of LISREL was considered important as major

advances and changes have occurred in later versions making much of the earlier LISREL work outdated (Fassinger, 1987).

In the present study, the first model most closely represents Carr et al.'s (1991) formulation. This model was tested in an attempt to replicate most of their findings. Model 1 tested the plausibility of a causal system comprising the latent variables of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance. The second general model represents an extension of the first model by incorporating the construct, combined strategy use (cognitive and metacognitive combined). Model 2 served as a basis for the development of Models 3 and 4. The third model, an extension of Model 2, involved examining the influences of the latent variables on the separate components of combined strategy use. Model 4 involved examining the separate effects of success and failure attributions on the other latent variables. All of the specified structural relations in the four models were unidirectional and positive (see Chapter 6, Figures 4 -7).

A multisample comparison approach, which involves estimating the model for the two groups simultaneously, and testing the equality or inequality of estimates of particular parameters in the different groups, was used in this study. It should be noted that variance-covariance matrices (as opposed to the correlations that were presented in Tables F1 - F3 to facilitate reader's interpretation) were used in modelling the data. A number of indices were examined to evaluate model fit. These include the chi-square values, and their associated degrees of freedom and probability levels for the models. In addition, the goodness-of-fit index (GFI), the normed fit index (NFI), the parsimonious normed fit index (PNFI) and the standardized root mean square residual (RMR) were also examined. The reader is reminded that good model fits are indicated by (a) small values of chi-square, (b) large statistically nonsignificant probability values (i.e., $\geq .2$) associated with the model chi-square, (c) values approaching 1 for the GFI, NFI, (d) moderate values of PNFI and (e) values below .1 for the standardized RMR (Bollen, 1989; Hair et al., 1992; Hayduk, 1987).

In addition to testing each general hypothesized model, a number of competing models were evaluated with each hypothesized model as suggested by MacCallum (1986) and J. C. Anderson and Gerbing (1988). It was argued earlier (see Chapter 7, section 7.1.6) that a test of competing models was necessary, given that a model that fits the data may not represent the "best" or "correct" model (Jöreskog & Sörbom, 1993). The competing models differed from the hypothesized models in that certain parameters representing relations between

constructs were constrained or freed. Such a set of models can be said to be nested. James, Mulaik, and Brett (1982) have discussed the need for the development of a sequence of tests for nested models prior to actually conducting the analyses to avoid "exploring" the data. Accordingly, each model serves as a basis of comparison for those prior to it in the sequence, and as such, must differ by having additional parameters constrained, or freed. Since the difference between two chi-squares is also distributed as a chi-square, hierarchically nested subsets of models can be compared in order to examine whether the constraining (or freeing) of parameters significantly improves goodness-of-fit. The null hypothesis is that the restricted parameter(s) equals zero, and a chi-square difference test is used to test the hypothesis. Under these conditions, a comparison of any two sequential models serves as a test of the importance of the restricted parameter(s).

9.3.1 Preliminary LISREL analyses

Problems can be encountered with early attempts to fit large models containing multiple indicators. Nonconvergent models can result from sampling fluctuations in covariance matrices or from inconsistencies among multiple indicators (J. C. Anderson & Gerbing, 1984; Boomsa, 1982, 1985; Hayduk, 1987). Hence, it can be helpful to start with a few well-established indicators, particularly for large models such as the ones tested here (Hayduk, 1987).

In accordance with the recommendation of Lomax (1983), preliminary LISREL analyses of Models 1 through 4 were conducted with the NA and RD groups independently. This provides some indication of possible group differences. Initially, for each group, the model was developed in a stepwise fashion by introducing additional variables one by one, to "get a feel" for the inadequacies of the indicators and the model. Hayduk (1987) suggested that by building models incrementally, one can eliminate many sources of problems with fit and indeterminacies. Indeterminacies can result from collinearity and multicollinearity, and impossible estimates such as negative variances. Sometimes indicative of identification problems are matrices for which no inverse can be found (i.e., a singular matrix or one with a zero determinant).

Lomax (1983) suggested that after testing each sample independently, a "final" LISREL model that will be used in the multisample analysis should be selected, where minor modifications in the model may still be made if necessary.

9.3.1.1 Model modifications

The results of these preliminary analyses indicated that a number of modifications were required before the model could be fitted to the data. A variety of statistics prompted modification of the model, as will be discussed below. The modifications were constrained by theoretical considerations in addition to purely statistical logic. Changes included the fixing of some measurement parameters, elimination of one measured variable as an indicator of a latent variable, and the addition of correlated errors.

With only a single measure (indicator) of the latent construct of "ability", it was necessary to make assumptions about the values of the measurement parameters. Ability is assumed to be a fallible measure, and hence a perfect reliability of this measure cannot be accepted. However, the LISREL program cannot compute the residual of a single indicator (as it does when two or more variables are specified to load on an underlying construct). Consequently, the reliability of the ability factor was estimated for the two groups on the basis of split-half indicators of IQ using the Spearman Brown formula (Maytham, Hendry, & Belser, 1973). Split-half indicators of ability, formed by creating two sums of a verbal test plus a performance test of the WISC-R, correlated .47 for NA readers and .31 for RD readers. These values when inserted in Spearman Brown's formula provided a reliability coefficient of .60 for NA readers and .50 for RD readers. The error variance of ability for each group was fixed at $(1 - r_{xx})$ multiplied by $\text{Var}(X)$, where r_{xx} is the reliability estimate for the indicator and $\text{Var}(X)$ is the observed variance of the indicator (Jöreskog & Sörbom, 1989).

An examination of the parameter estimates in a preliminary test of the models for RD readers indicated the presence of a negative variance or "Heywood case" (Bollen, 1989; Hayduk, 1987) for the PAT measure. The notion of a negative variance clearly is meaningless, and hence such an estimate is considered unacceptable (Hayduk, 1987). To eliminate this problem, the error variance of this measure was fixed for the RD reader group. As there was no available information on split-half indicators of the PAT to estimate the reliability of the measure and how much measurement error should be imposed, the psychometric reliability of the measure was used. Hence, to eliminate an improper communality (a "Heywood case") in the RD reader group, the uniqueness parameter or the measurement error for the PAT indicator was set to 1 minus the reliability of the PAT measure, multiplied by the observed variance.

The next change involved the elimination of internal (locus of causality) attributions for failure as a measure of the latent construct of attributional style. This measure was eliminated on the basis of a low squared multiple correlation, a value indicating the proportion of variance in the measured variable accounted for by the latent variable.

The fourth change involved the addition of correlated errors of measurement. A number of a priori reasons dictated the relaxation of error covariances and these were discussed in section 7.4.2 of Chapter 7. To reiterate, Parkes (1987) suggested that psychometric data obtained from questionnaires are usually prone to being systematically distorted by generalized response biases. Examination of the correlations among the attribution indicators suggested that there may be other a priori reasons for allowing correlated errors of measurement amongst the measures of attributions. For instance, it was noted that in the NA reader group the three success measures were highly correlated ($r = .73$ to $.76$) with each other but had only moderate correlations with the two failure measures. In addition, the correlation between the two failure measures ($r = .60$) was slightly higher than the correlations of these measures with the success measures (see Appendix F, Table F1).

These considerations suggest the presence of method-specific effects corresponding to something like success and failure response sets. The same argument applies to the model for RD readers, although in this case, the failure attributional measures have the higher residual correlation. On a priori grounds, the above method-specific effects were controlled by allowing correlated errors of measurement amongst the measures of success attributions, and similarly allowing correlated errors of measurement amongst the measures of failure attributions.

The other area in which correlations suggested there may be a need to allow covariance is the relationship between self-efficacy and reading achievement. In the present study, self-efficacy (particularly task-specific self-efficacy) correlated more highly with the achievement measures than the model predicted. Thus, it seemed reasonable to permit the error variance of the task-specific measure to covary with the errors of the two achievement measures. These error covariances are also theoretically justifiable as confidence in one's ability to read is very closely related to performance on the tasks (Bandura, 1986).

The "modification indices" in the LISREL program provide a powerful tool for detecting misspecification within the general form of the model specified. In particular, they indicate

which of the formal restrictions (such as error covariances) would improve the fit of a model if they were relaxed. Consistent with the above arguments, examination of the LISREL modification indices showed that the largest values occurred among the off-diagonal elements of the error covariance matrix. This suggested that freeing/relaxing some error covariances (i.e., in this case, allowing them to take on a nonzero value) would result in an improvement in model fit. Jöreskog and Sörbom (1989) recommend that only one parameter be freed at a time on the basis of modification index value. This is because a single change in the parameter specifications for the model changes all the modification indices. Provided that it accords with substantive considerations, the parameter with the largest index is usually chosen so as to bring about the maximum improvement in model fit (Hayduk, 1987; Jöreskog & Sörbom, 1989). Although modifications can often be derived on theoretical grounds, the more common practice is to modify models on the basis of the data. Model modifications, if undertaken, must be made cautiously and must be theoretically and substantively meaningful. For example, allowing error terms to covary might decrease the chi-square value and thereby improve statistical fit by introducing a free parameter to the model. This choice, however, must be based on realistic assumptions, otherwise one is simply capitalizing on chance to improve fit.

Results of the preliminary LISREL analyses also indicated that the path from ability to metacognitive knowledge was not significant in either group. This finding suggests that the ability to metacognitive knowledge path can be set to zero in the hypothesized models.

In summary, preliminary analyses indicated that a number of modifications, based on theoretical considerations and statistical logic, were required before the model could be fitted to the data. Changes included the fixing of some measurement parameters, elimination of one measured variable as an indicator of a latent variable, and the addition of correlated errors.

9.3.1.2 Sensitivity analyses

Although correlated errors of measurement were permitted for each of the hypothesized models based on justifiable grounds, there remains the uncertainty that one may simply be capitalising on chance variation in the data. There is always the concern of overfitting the model; that is, fitting the model to trivial sample-specific artifacts in the data. Researchers have stressed the importance of investigating the sensitivity of parameter estimates to the specification of correlated errors of measurement (e.g., Hughes et al., 1986). Sensitivity

analyses (Byrne, 1989a) were thus conducted to determine the practical significance of including additional parameters like correlated errors (i.e., their importance to the overall meaningfulness of the model). One way of determining this information is to test the sensitivity of the major parameters in the model to the addition of different patterns of error covariances. For instance, in Model 1, six correlated errors of measurement were released for NA readers and three correlated errors of measurement were released for RD readers. However, modification indices suggested that two additional error covariances could be released to improve model fit further. The question that could or should be asked here is: How many error covariances should one release and does this substantially affect the estimates of the major parameters in the model?

Experimentation with the fitted models, using multisample analyses, shows that regardless of the pattern of correlated errors fitted for a given model, the same set of structural relationships (i.e., pattern of covariances estimated between the latent factors) are observed to hold. Thus, the imposition of correlated errors has minimal impact on the estimated structural parameters of the model, but does significantly improve model fit. This provides some reassurance that assumptions concerning the error structure of the data have no impact on the essential conclusions drawn from the fitted models.

9.3.1.3 Alternative factor structures for multifaceted constructs

Two multifaceted constructs, namely attributional style and combined strategy use, formed important components of the hypothesized models. Attributional style was considered a multifaceted construct comprising two major subcomponents of achievement outcome: attributions for success and attributions for failure. Combined strategy use also consisted of two subcomponents: metacognitive strategy use and cognitive strategy use. As mentioned earlier, in reference to multifaceted constructs, Carver (1989) suggested that a test of both the composite index and its subcomponents was necessary to establish the basis by which the composite was associated with the outcome measure. Accordingly, these multifaceted constructs were initially tested as one factor models by combining the subcomponents into a single predictive index followed by an examination of the separate effects of the subcomponents on the rest of the model. Preliminary multisample analyses were conducted on the factor structure of the multifaceted constructs to determine their goodness-of-fit.

(i) *Attributional style*

The patterns of correlations obtained between the success and failure attributions may also suggest that the latent construct of attributional style be specified as comprising separate latent variables for positive and negative outcomes (cf. Hull & Mendolia, 1991). There may be conceptual grounds for treating the latent construct of attributional style as two separate latent factors because a number of studies indicate that "attributions serve different functions depending on whether they account for success or for failure" (Oka & Paris, 1987, p. 123). Conversely, it can be argued that attributional style should be modeled as a single latent construct, in accordance with previous empirical research (Berndt & Miller, 1990; Carr et al., 1991). This is consistent with the argument that students' attributions for both success and failure outcomes are necessary and combine or interact to determine reading metacognitive knowledge or performance.

Separate confirmatory factor analyses were conducted on the set of attribution measures to determine the specification of attributional style before assessing the fit of the larger model. A series of multisample analyses were conducted using only the five measures of attributional style with NA and RD readers. Preliminary fitting of a model that specified all five measures as related to a single latent variable with all factor loadings invariant except for failure attribution indicators yielded a poor fit to the data, $\chi^2(12, N = 405) = 44.70, p = .00, GFI = .95$. Modification indices suggested that the fit of the model could be improved by allowing the errors of the failure indicators to covary in each group. The fit of the model was substantially improved by freeing these error covariances, $\chi^2(10, N = 405) = 7.56, p = .67, GFI, NFI = .99, RMR = .024, PNFI = .50$. A third model was tested as involving separate latent variables for positive and negative event attributions and an invariant factor structure. This model also provided a good fit of the observed data, $\chi^2(11, N = 405) = 10.57, p = .48, GFI, NFI = .99, RMR = .032, PNFI = .54$. The goodness-of-fit indices indicated that both the single factor model and the two factor model present good representations of attributional style.

(ii) *Strategy use*

Specification of the latent construct of strategy use as a one factor model has received support from previous research (Meece et al., 1988), as it is consistent with the conceptualization of self-regulated learning theorists (Zimmerman, 1989a). The examination of separate

components of the latent construct of strategy use has also received theoretical (Borkowski et al., 1989, 1990, 1992) and empirical support (Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990). For instance, within Borkowski's theoretical framework the distinct components of metacognitive and cognitive strategy use have differential effects on the other latent constructs of interest in this study (Borkowski et al., 1989, 1990, 1992).

Separate confirmatory factor analyses were conducted on the strategy use measures to determine the specification of the latent construct of strategy use before assessing the fit of the larger model. A series of multisample analyses were conducted using four measures of strategy use with NA and RD readers. The four measures of strategy use included the self-report measures of cognitive and metacognitive strategy use and the behavioural measures of cognitive and metacognitive strategy use. Preliminary fitting of a model that specified all four measures as related to a single latent variable with all factor loadings invariant yielded a poor fit to the data. Modification indices suggested that the fit of the model could be improved by allowing three correlated errors for the NA reader group and two correlated errors for the RD reader group. The fit of the model was substantially improved by freeing these error covariances, $\chi^2(3, N = 395) = 7.82, p = .05, GFI, NFI = .98, RMR = .046, PNFI = .25$. Another model was tested and involved separate latent variables for cognitive and metacognitive strategy use, an invariant factor structure, and one correlated error for each of the reader groups. This model also provided a good fit of the observed data, $\chi^2(2, N = 395) = 0.52, p = .76, GFI, NFI = 1.00, RMR = .011, PNFI = .17$. The goodness-of-fit indices indicated that although both the one and two factor models provided a good fit, the two factor model presented a better representation of the latent construct of strategy use for the data of this study. Notwithstanding this, the hypothesized models were initially tested with a one factor model on the basis of parsimony and its conceptual consistency with theoretical formulations of previous researchers (Meece et al., 1988).

9.3.2 Test of hypothesized models

In assessing model fit, one needs to pay attention to both the *measurement* and the *structural* parts of the model. Whether this should be done simultaneously or sequentially is still a point of debate in the literature. The majority view seems to be that both the measurement and the structural model should be evaluated simultaneously (Fornell, 1987). However, it has been argued that "proper specification of the measurement model is necessary before meaning can

be attached to the analysis of the structural model. That is, good measurement of the latent variables is prerequisite to the analysis of the causal relations among the latent variables" (J. C. Anderson & Gerbing, 1982, p. 453). Considering the merits of this stepwise approach, in the present study, the measurement model was specified and tested before structural estimation.

In the measurement model, each observed variable was assumed to have a nonzero loading only on the latent variable that it is presumed to measure. For Models 1 through 4, once a parsimonious, well fitting measurement model was identified, a model was tested in which structural relations were imposed among the factors of this measurement model. For the structural model of the relations among the latent variables, it was assumed that causal relations would be recursive, that is, all paths are specified in only one direction. Such recursive models were developed and tested because they are more parsimonious, easier to estimate, and easier to interpret than complex, nonrecursive models which allow reciprocal influences among variables. Furthermore, recursive models are justified on the basis of the theoretical formulations of the author.

LISREL analyses of Models 1 through 4 were conducted by analysing the groups simultaneously. However, the Box's M test for homogeneity of variance-covariance matrices across populations yielded a significant chi-square value, $\chi^2(36, N = 395) = 139.32, p = .00$, suggesting that major differences existed between the two matrices. Researchers investigating multisample relationships have sometimes used a significant Box's M result as justification for conducting separate LISREL analysis of groups (e.g., D. A. Cole, 1989). However, the fact that there are differences in the variance-covariance matrix does not inform one as to the sources of differences, especially when group analyses are conducted separately.

In the present study, multisample LISREL analyses were conducted rather than separate group analyses for the following reasons. First, they are a more powerful form of analysis which provides a test of similarities and differences in relationships across groups. In doing separate group analyses, the sample size would be halved resulting in considerable reduction in power. Second, one would be able to conduct more direct and informative tests of equality of parameters across groups which could be important, given the hypotheses of the present study. In the present study it was hypothesized that particular structural paths were invariant across groups. Thus, it is imperative for the data from both groups to be analyzed simultaneously in order to obtain efficient estimates (Jöreskog & Sörbom, 1985). Third, running separate

group analyses does not indicate whether these sources of differences originate from the measurement model or the structural model. Differences in variances across groups does not necessarily mean structural differences. Patterns of relationships amongst the variables may be similar and this may indicate, as in the present study, some similarity in the structural effects across groups.

In multisample analyses, the chi-square goodness-of-fit generated by LISREL enables one to test the hypothesis that the model holds in each group simultaneously, that is the model is invariant across groups. The more appropriate way to test for invariance is by constraining some, or all, parameter estimates to be equal across groups (see Jöreskog & Sörbom, 1989; Marsh & Hocevar, 1985). Omnibus tests of total invariance include tests of the equality of covariance matrices of relations among the measured variables for the different groups and, models that constrain all parameters in a particular model to be equal across the multiple groups. These omnibus tests nearly always result in statistically significant chi-square values and substantial decrements in fit, so more limited tests are pursued in which specific sets of parameters are posited to be invariant over groups (e.g., factor loadings, measurement error, factor correlations, or path coefficients). The LISREL program provides goodness-of-fit statistics for each group's model and a chi-square measure of overall goodness-of-fit for the two groups together. This statistic, therefore, indicates a fit (or lack of fit) of the model to the data for both groups, including all constraints.

In the present study, the extent of structural invariance across groups was examined. According to theory and empirical research, it was expected that a number of the causal paths between the latent variables would be similar in magnitude across groups. In the present study, particular elements of the Γ and the B matrices were constrained equal across groups according to substantive theory (see section 7.4.3.1 (ii) of Chapter 7).

It has been recommended, for comparisons across multiple groups, that the general form of the measurement model must hold for all groups for which comparisons are to be made. Two models have the same form if the model for each group has the same parameter matrices with the same dimensions and the same location of fixed, free, and constrained parameters. For valid and meaningful comparisons to be made across groups, between-group invariance in the number of factors, the pattern of factor loadings, and the magnitude of one or more factor loadings in the measurement model must be demonstrated (Bollen, 1989; Byrne et al., 1989).

The analyzed models were based on two covariance data matrices, one from each group, that used listwise deletion of missing data. These analyses were performed on covariances, rather than correlations, and thus comparisons are not confounded by differences in group variances (cf. Alwin & Jackson, 1981). The use of LISREL maximum likelihood techniques allows the significance of changes in goodness-of-fit resulting from alterations in the specification of model parameters to be assessed by calculating the significance of the change in the chi-square statistic. The approach adopted in the present study was to initially formulate a model with few parameter constraints across groups. Subsequently, a systematic series of competing models were estimated which were designed to test the effects of increasing or decreasing the constraints imposed on the model in accordance with specified hypotheses. The statistical significance of improvement or deterioration in goodness-of-fit resulting from changes in the extent to which the model was constrained was determined by examining the change in the chi-square statistic relative to the change in degrees of freedom. Models that resulted in a significant improvement in goodness-of-fit were accepted; models that resulted in significant deterioration in goodness-of-fit were rejected. When the difference in goodness-of-fit between two models was found to be nonsignificant, the model that provided the more parsimonious solution, that is, the more constrained of the two models, was accepted. To supplement the chi-square difference test, goodness-of-fit measures such as the GFI, NFI, PNFI, and RMR were examined. A number of detailed measures of fit were also used to assess the quality and goodness-of-fit of the final model in each analysis. These included the squared multiple correlations, residual matrix correlations, standardized residuals (Q-Q plot) and modification indices.

In the following sections, a detailed description of the basic model, Model 1, is provided initially. As Models 2 through 4 represent extensions of the basic model, it was considered unnecessary to present the same degree of detail in discussing them. Discussion of these models, however, will focus on the newly introduced aspects of that model. For each test of the structural model and its associated measurement model, a series of distinct hierarchical nested models is presented and tested first. These nested models provided a test of the antitheses of the hypotheses that were formulated for the model (see Chapter 6, section 6.2), and at other times, provided a test of competing views or theories on a particular relationship.

The summary goodness-of-fit statistics are also presented for these models, followed by a discussion of the parameter estimates for the model deemed to be best fitting for these data. The individual parameter estimates provide the detail necessary to understand the causal

influences between variables. In causal ordering research, the direction, size, and location of individual parameters are important in understanding the accuracy of the theoretical model. The strength of effect sizes is identified by the LISREL program in the form of maximum likelihood coefficients (regression coefficients) and stated *t* values. A detailed LISREL specification of each of the best fitting models is provided in Appendix G. In summary, in each description of the hypothesized model, the results of a series of competing models tested will be presented first followed by a detailed description of the model of best fit. However, for ease of reading, the results of the competing models for each hypothesized model is presented in the appendix (see Appendix H).

9.3.2.1 Test of proposed theoretical Model 1

(i) *Measurement model*

Model 1 was tested to replicate some of Carr et al.'s (1991) findings. Before the structural or path model of direct and mediating effects for Model 1 was examined, an initial measurement model was estimated to evaluate the latent constructs of ability, attributional style, self-efficacy, metacognitive knowledge, and reading performance, their intercorrelations, and hypothesized factor loadings.

Preliminary analyses indicated that both the single factor model and the two factor model provided equally good representations of attributional style. However, in testing Model 1, the one factor model was preferable on the basis of parsimony and its conceptual consistency with theoretical formulations of Carr et al. (1991).

The indicators specifying each latent construct are described in Chapter 6, section 6.2.1. For each latent construct, one construct loading was set a priori to unity in order to provide a metric for the latent construct and to identify the model (Hayduk, 1987; Long, 1983a). Unless the scale of the latent variables has been established, there exists an indeterminacy between the variance of the latent variable and the loadings of the observed variables on that latent variable. Each measure was allowed to load onto one and only one of the five latent constructs. The loadings of the indicators onto other factors were constrained to equal zero. Table 8 (see p. 193) shows the indicators with their respective latent constructs. The matrix of correlations of the 12 indicators used in this analysis and their standard deviations is presented in Table F1 of Appendix F. These data, in the form of covariances, formed the input for the LISREL analysis of Model 1.

As described previously, a number of other model parameters were also fixed (see section 9.3.1). As ability was specified by a single indicator, the reliability of this measure was fixed by setting the error variance of this measure in each group to be equal to the variance of the ability measure multiplied by one minus the corresponding reliability estimate. Thus, for NA readers the error variance was fixed to 32.98, obtained by calculating $\{(1 - .60) \times 82.46\}$, and for RD readers the error variance was fixed to 33.30, obtained by calculating $\{(1 - .50) \times 66.62\}$. Similarly, the error variance of the PAT indicator was fixed to 8.18 for the RD reader group, to avoid an improper communality.

The model tested by Carr et al. (1991) assumed that measurement errors were uncorrelated, that is, that there were no common response tendencies that influenced the observed measures in a similar way within groups. However, the validity of the assumption of uncorrelated errors was open to doubt, as indicated earlier. Consequently, covariances between pairs of measured variable residuals (unique variances) were added to the model on the basis of the modification indices and on a priori grounds (see section 9.3.1.1). Only statistically significant covariances were retained.

A series of distinct hierarchical measurement models was tested initially, details of which are provided in section H.1.1 of Appendix H. Table H1 of Appendix H presents a summary of the nested models that were tested. It was theorized earlier that an identical factor loading matrix (i.e., $\Lambda_x^{[g1]} = \Lambda_x^{[g2]}$; $\Lambda_y^{[g1]} = \Lambda_y^{[g2]}$) might be expected since the latent variables were measured using the same indicators in both groups. Analyses of these nested models indicated that the initial specification of an invariant factor loading matrix across groups (as specified in section 7.4.3.1 (i) of Chapter 7) did not provide a good fit. In other words unit increases in true scores did not all lead to the same increments in measured variables for RD readers as for NA readers. However, it was acknowledged that there would be poor support for the total invariance of all parameter estimates in the measurement part of the hypothesized models (i.e., $\Lambda_x^{[g1]} = \Lambda_x^{[g2]}$; $\Lambda_y^{[g1]} = \Lambda_y^{[g2]}$; $\Theta_\epsilon^{[g1]} = \Theta_\epsilon^{[g2]}$; $\Theta_\delta^{[g1]} = \Theta_\delta^{[g2]}$). Some evidence for this was seen from the preliminary analyses which yielded different error covariances for the groups. Further analyses indicated that freeing the construct loadings of the failure attributions resulted in a substantial improvement in the fit of the model. This model (Model 1.2) yielded $\chi^2(87, N = 405) = 76.29$, GFI, NFI = .96, PNFI = .64, RMR = .04, and was accepted on statistical grounds ($p = .79$) as producing a good fit for these data (see Appendix H, Table H1). However, this partial factorial invariance still enables multisample comparisons to be made (Bollen, 1989). The LISREL specification for this optimum measurement model is provided in section G.1.2.1 of Appendix G.

Description of results of optimum measurement model

In light of the results of the nested models, the decision was made to accept Model 1.2 as the final optimum model (see Appendix H, section H.1.1, Table H1). The LISREL estimates of the construct loadings, error variances and covariances, factor variances/covariances and correlations of latent constructs for this optimum measurement model were examined. These are shown in Tables 8 through 11. To enable comparisons across groups to be made, the parameter estimates are standardized in relation to a metric that is common across the two groups. This is known as the common metric completely standardized solution (Jöreskog & Sörbom, 1989). As the estimates are standardized over the two groups, the values of the parameter estimates may take on values greater than 1. The common metric completely standardized solution differs from one that would standardize estimates within each group (within group standardized solution) because in the latter solution, constraints across groups would be lost (see Jöreskog & Sörbom, 1989, pp. 238-242).

The raw values (unstandardized) of the individual parameter estimates shown in Table 8 are not meaningful because they depend on the scales of the variables involved. The significance of each parameter can be tested with the value referred to in LISREL as the t value (i.e., value of the parameter estimate divided by its standard error). This information can be used in a normal z test to evaluate whether a parameter estimate is equal to or greater than 1.96 times its standard error (for a two-tailed test) or 1.65 times its standard error (for a one-tailed test). Thus, estimates with a t value smaller than 1.96 are typically described as nonsignificant at the two-tailed level and are candidates for removal from the model (Bollen, 1989). Significant t values (≥ 1.96) can be regarded as indicative of parameter estimates significantly different from zero. Information regarding tests of significance are based on nonstandardized values.

As shown in Table 8, all indicators were significantly related to their underlying latent construct ($t \geq 10.14$, $p \leq .0001$). The moderate-to-high factor loadings of the indicators suggest that they are effective instruments in defining the latent variables. All loadings were constrained to be equal across the groups except for failure attributions which showed much higher loadings for RD readers. This suggests that the association between the latent construct of attributional style and the observed failure attribution measures was stronger in the RD reader group than in the NA reader group.

The squared multiple correlation (R square) for each observed variable is an indication of the reliability of each observed measure with respect to its underlying latent construct (see Table

8). A virtue of SEM over traditional approaches to internal consistency, such as Cronbach's alpha, is its ability to evaluate the consistency of the relations between a set of indicators and the latent variable they represent (Hoyle & Smith, 1994). For example, in considering the reliabilities of attributions for NA readers, it is seen that the success attributions were the most reliable whilst the failure attributions were the least reliable. The converse was seen for RD readers. In general, the measures had higher construct reliabilities for NA readers than for RD readers with the exception of failure attributions. The PAT measure was the most reliable measure for both NA and RD readers. Note, however, that the reliability of the PAT measure was fixed for RD readers to the psychometric reliability of the measure.

Table 8. Factor loadings and reliabilities of indicators for Measurement Model 1

CONSTRUCT/ indicator	NA Readers			RD readers		
	unstandardized (SE)	standardized ^a	reliability ^b	unstandardized (SE)	standardized	reliability ^b
ATTRIBUTIONS						
Success						
internality	0.86 (0.05)	0.69	.72	0.86 (0.05)	0.69	.27
controllability	1.00 ^c	0.79	.74	1.00 ^c	0.79	.44
stability	1.00 (0.05)	0.80	.76	1.00 (0.05)	0.80	.45
Failure						
controllability	0.61 (0.06)	0.51	.44	1.37 (0.12)	1.15	.67
stability	0.58 (0.06)	0.47	.39	1.59 (0.13)	1.28	.79
EFFICACY						
general	0.06 (0.00)	0.74	.59	0.06 (0.00)	0.74	.46
task-specific	1.00 ^c	0.75	.62	1.00 ^c	0.75	.44
METACOGNITIVE KNOWLEDGE						
MSI	1.00 ^c	0.71	.48	1.00 ^c	0.71	.55
IRA	1.02 (0.10)	0.65	.50	1.02 (0.10)	0.65	.36
PERFORMANCE						
PAT	1.00 ^c	0.92	.79	1.00 ^c	0.92	.95 ^c
classroom	0.72 (0.05)	0.69	.60	0.72 (0.05)	0.69	.35
ABILITY						
WISC-R	1.00 ^c	0.75	.60 ^c	1.00 ^c	0.75	.50 ^c

All Loadings that were estimated are significant at $p \leq 0.0001$ (2 tailed).

^a Refers to common metric completely standardized solution.

^b Estimated by LISREL (Jöreskog & Sörbom, 1993).

^c Fixed value or reference indicator.

A number of measures had reliabilities less than .5. These included the success attribution measures, efficacy measures, the IRA measure of metacognitive knowledge, and the classroom

performance measure for RD readers. For NA readers it included the MSI measure of metacognitive knowledge and the failure attribution measures. While these values are still satisfactory, they are indicative of somewhat less reliability than the other measures. The coefficients of determination for the endogenous and exogenous observed measures were .99 and .55 (range .50 to .60), respectively, for both groups. The coefficient of determination provides a generalized measure of the amount of variance in the observed measures accounted for by the measurement model. In this case the measurement model performs well, since it accounts for a high proportion of the variation in the observed scores, suggesting that the reliability of the measurement model as a whole is exceptionally high.

Table 9. Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 1

	1	2	3	4	5	6	7	8	9	10	11	12
<i>NA readers</i>												
1. contros	0.31											
2. stabs		0.28										
3. stabf			0.49									
4. controf			0.16	0.48								
5. inters					0.27							
6. effica			0.10			0.56						
7. efficb							0.50					
8. msi			0.12					0.55				
9. ira		-0.09							0.42			
10. pat							0.26			0.28		
11. class							0.17				0.41	
12. iq												0.44*
<i>RD readers</i>												
1. contros	0.48											
2. stabs	0.13	0.44										
3. stabf			0.26									
4. controf				0.38								
5. inters				-0.10	0.77							
6. effica						0.34						
7. efficb							0.38					
8. msi								0.42				
9. ira									0.74			
10. pat										0.03*		
11. class							-0.10				0.64	
12. iq												0.45*

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

For purposes of simplicity, all error variances and covariances have been combined into a single matrix.

All residuals are significant at $p \leq .05$ (two-tailed).

* Unstandardized values of these parameters were fixed.

The error variances and covariances of both groups are presented in Table 9. Altogether 6 errors were allowed to covary in the NA reader group and 3 errors to covary in the RD reader group. The t values for the variances and covariances of the measurement errors were generally considerably greater than 1.96 and so are statistically significant at better than the $p \leq .05$ level. Thus, the observed scores in these samples were subject to measurement error, and they were also affected by generalized response tendencies, which resulted in nonrandom distribution of errors.

Table 10 shows the common metric completely standardized solution of the variances and covariances of the latent constructs for NA and RD readers. Examination of Table 10 indicates that the variances and covariances of the latent constructs were in general greater for NA than RD readers. This is expected as the selection criteria used in determining the RD reader group would have led to more constrained variances in some of the constructs in this group. In the within group completely standardized solution (see Table 11), the coefficients linking the latent variables of ability, attributional style, efficacy, metacognitive knowledge, and performance are estimates of the correlations among these dimensions. The coefficients ranged from .42 to .78 for NA readers and .15 to .73 for RD readers, suggesting that in general, they are distinct dimensions, that is, the variables tap separate but related aspects of reading performance. For both groups, correlations among attributions, efficacy, and metacognitive knowledge were high.

Table 10. Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 1

	Attributions	Efficacy	Metacognitive Knowledge	Performance	Ability
<i>NA readers</i>					
Attributions	1.42				
Efficacy	1.05	1.46			
Meta. Knowledge	0.93	0.92	0.99		
Performance	0.69	0.78	0.72	1.27	
Ability	0.54	0.64	0.51	0.71	1.19
<i>RD readers</i>					
Attributions	0.58				
Efficacy	0.40	0.54			
Meta. Knowledge	0.53	0.46	1.01		
Performance	0.42	0.46	0.52	0.73	
Ability	0.10 ^a	0.28	0.22 ^a	0.23	0.80

All estimates for NA readers are significant at $p \leq .0001$ (two-tailed).

All estimates for RD readers are significant at $p \leq .01$ (two-tailed) except those indicated by ^a.

Table 11. Correlations of latent constructs for NA and RD readers for Measurement Model 1

	Attributions	Efficacy	Metacognitive Knowledge	Performance	Ability
Attributions		.71	.70	.65	.15 ^a
Efficacy	.73		.63	.73	.42
Meta. Knowledge	.78	.76		.61	.25 ^a
Performance	.51	.58	.64		.29
Ability	.42	.48	.47	.57	

Triangular correlation matrix for NA readers is presented below the main diagonal and the matrix for RD readers is presented above the main diagonal. All correlations are significant at $p \leq .01$ except for those indicated by ^a.

For this sample of NA and RD readers, both the Q-Q plot and the correlations of the parameter estimates confirm the good fit of the estimated model to the data. For present purposes, the abovementioned model was accepted as the optimum parameter specification for the construct loadings and measurement errors. This pattern of fixed, free, and equated parameters for the measurement model was kept constant (but not fixed) in the second part of the analysis which is described in the next section.

(ii) *Structural model*

Given an adequate measurement model, a structural model was tested to explain the relations among the five latent constructs. The hypothesized paths were largely based on the theoretical propositions of Carr et al. (1991). According to Model 1 (see Figure 4, p. 86), metacognitive knowledge should mediate the effects of attributional style and self-efficacy on reading performance. This model specified that attributional style predicts metacognitive knowledge directly or indirectly through self-efficacy. Although largely duplicating Carr et al.'s (1991) pathways, Model 1 also specified additional pathways relating to self-efficacy. For instance, it posited that ability directly predicted efficacy, which in turn, predicted reading comprehension performance. It was also theorized that the model would be essentially the same for NA and RD readers except for the following paths: ability to attributions, ability to metacognitive knowledge, and ability to performance. These paths were hypothesized to be statistically nonsignificant for RD readers but significant for NA readers.

Preliminary analyses of nested models indicated hypothesized Model 1 warranted respecification (see Appendix H, section H.1.2). Table H2 of Appendix H presents a summary of the series of nested models that were tested. The findings indicated that ability

does not directly predict metacognitive knowledge in both groups and hence, this path was fixed to zero in all subsequent analyses. The influence of ability on metacognitive knowledge was not direct (as hypothesized), but was significantly mediated by attributional style and self-efficacy. The results also indicated that, contrary to predictions, the efficacy to performance path differed across groups. This path was significant for RD readers but not for NA readers, which suggested that the efficacy to performance path should not be constrained equal across groups as hypothesized in Figure 4 (see p. 86). The respecified model (Model 1.3) had a chi-square value of 86.15 with 96 degrees of freedom, indicating a very good fit of the model to the data ($p = .75$). Additionally, the goodness-of-fit indices (GFI = .96, NFI = .96, PNFI = .70), and in particular the RMR (RMR = .051), showed a marked improvement of the revised model over the hypothesized model. See section G.1.2 of Appendix G for the LISREL specification of the optimal structural model.

Description of results of optimum structural model

The results of the revised hypothesized model (representing the final optimum model) are presented in Figure 12 for NA and RD readers. Two types of standardized solutions are presented. In the first solution, parameter estimates are standardized in relation to a metric that is common across the two groups, and which differ from those that would result from standardizing estimates for each group separately in which constraints across groups would be lost (see Jöreskog & Sörbom, 1989, pp. 238-242). For comparison across groups, this solution was preferable as a scale common to the two groups was required. Thus, the discussion of this model is based on the common metric completely standardized estimates. However, these standardized coefficients do not represent regression beta weights as would be the case with the within group standardized coefficients.

To help the reader, the estimates of the within group completely standardized solution are also available and are presented in Table J2 (see Appendix J) where it is discussed in relation to the robustness of the model when demographic variables of ethnicity and gender are included as control variables. The estimates in Table J2 show a path analytic representation of the values of the parameter estimates of NA and RD readers. In this solution, each variable has been transformed so that its variance is 1.0. Thus, regressions coefficients can be interpreted as beta coefficients and covariances as correlations. Although these path weights are meaningful within a particular group, it is not appropriate to compare these path weights across groups since these coefficients are not standardized in relation to a metric that is common across the two groups.

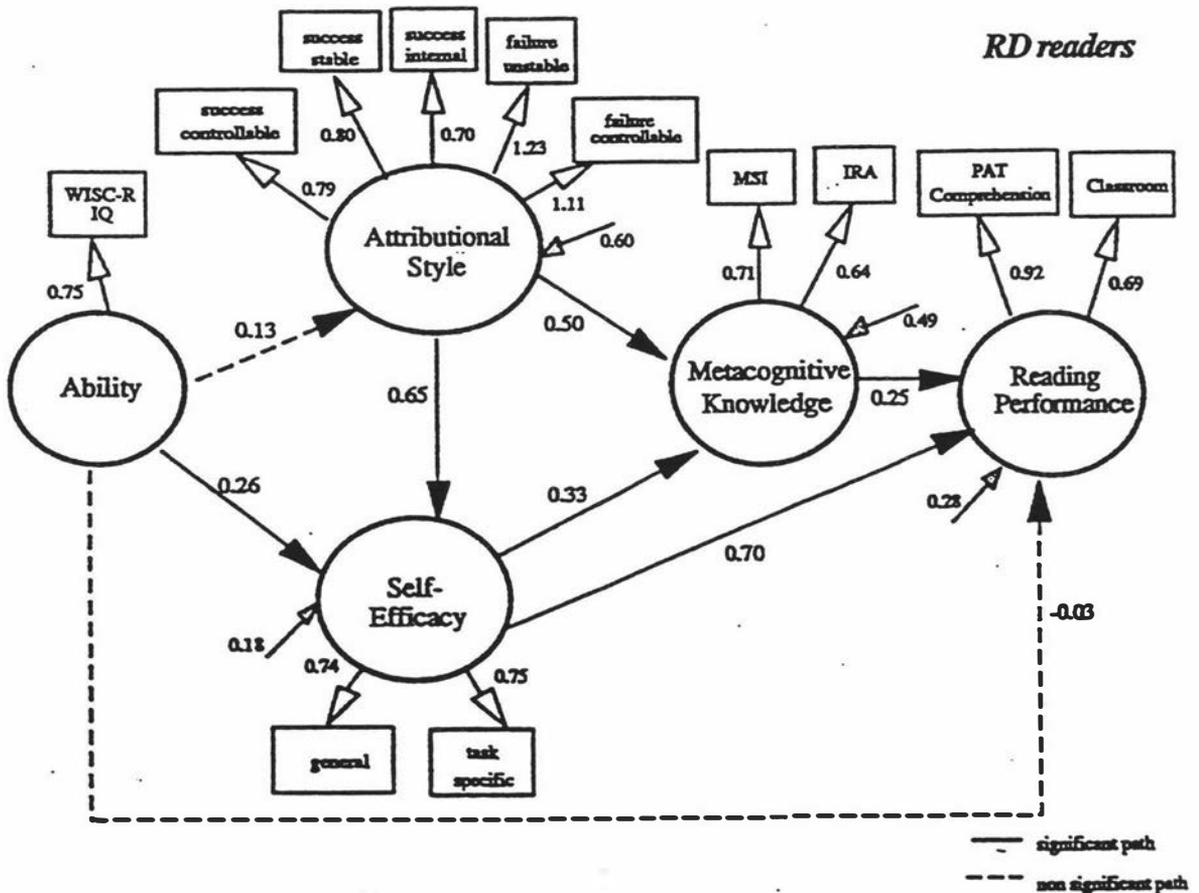
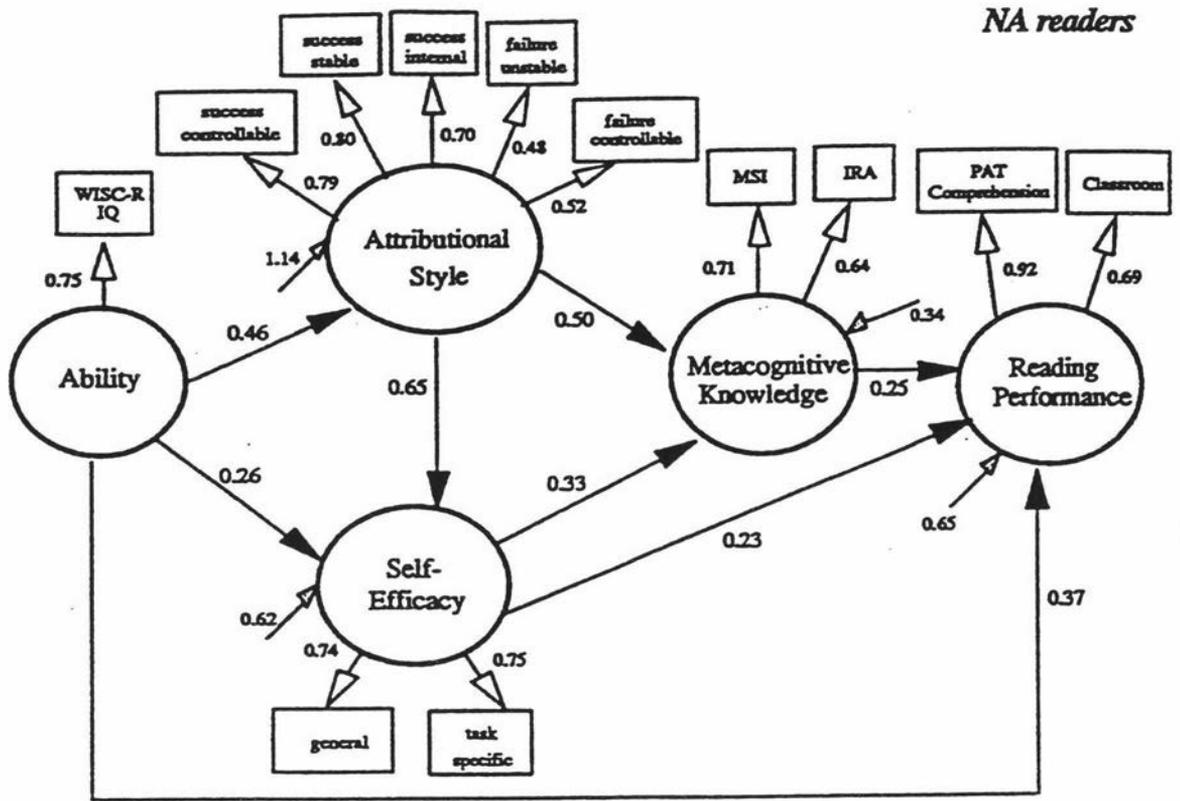


Figure 12. Findings of revised Model 1 for NA and RD readers.

Note: Failure internal attribution was dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero. To avoid clutter, measurement errors are not presented. Common metric completely standardized solution is presented for all parameters.

In keeping with the conventional notation of covariance structure modelling, in Figure 12, measured variables are enclosed in rectangles and latent variables in circles. Filled arrows connecting circles represent causal effects and unfilled arrows relating circles to rectangles represent the latent variable's effects on indicators. Detached arrows indicate residual or structural disturbance variances. Solid lines depict significant effects and dashed lines depict nonsignificant effects. The general convention of assuming sufficient power in the LISREL model, leading to a rejection of the null hypothesis (i.e., testing whether a parameter is significantly different from zero) with $t \geq 1.96$ was followed. All significant effects were in the hypothesized direction (see Appendix J, Table J2, for t values). Measurement errors were deleted from the figure to reduce clutter, but were estimated in the analysis. These error variances and covariances are presented in Table I1 (see Appendix I) for NA and RD readers respectively.

The results of the revised hypothesized structural model for both groups show a fairly good fit and the estimated factor loadings of this model did not deviate markedly from the measurement model. All factor loadings were sizable and significant ($p \leq .0001$). The coefficients of the measurement model in the previous section, although similar, may not be identical to those in this model because the analyses in this model involved solving a full system of equations comprising both measurement and structural effects.

As seen in Figure 12, all causal paths were constrained to be equal across groups except for the paths ability to attributional style, ability to performance, and efficacy to performance. In both groups, attributions significantly predicted self-efficacy (standardized solution = 0.65, $t = 9.89$). Self-efficacy, in turn, predicted metacognitive knowledge (standardized solution = 0.33, $t = 3.15$), which in turn predicted reading performance (standardized solution = 0.25, $t = 2.80$). As indicated earlier, the direct effect of ability on metacognitive knowledge was not significant for either group and thus this path was fixed to zero. Results suggest that the relationship between ability and metacognitive knowledge was mainly a result of indirect causal covariation, largely due to the effects of attributions and self-efficacy. It is clear that ability predicted metacognitive knowledge almost exclusively through the indirect mediating variables rather than directly.

The results indicated that ability was directly associated with attributional style (standardized solution = 0.46, $t = 4.47$) for NA readers but there was no direct significant effect for RD readers (standardized solution = 0.13, $t = 1.45$). Ability was significantly and positively

associated with reading performance for NA readers, but not for RD readers. As expected, the relationship between ability and reading performance for RD readers did not replicate Carr et al.'s (1991) finding. The results also indicated that attributional style was more strongly related to metacognitive knowledge than efficacy was, and this finding was consistent with Borkowski's theory (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990).

As hypothesized, self-efficacy directly and positively predicted reading performance. Although the path efficacy to performance was significant for both NA (standardized solution = 0.23, $t = 1.97$) and RD readers (standardized solution = 0.70, $t = 5.07$), this path differentiated between NA and RD readers. Contrary to predictions, the relationship between efficacy and performance was much stronger for RD than for NA readers, suggesting that efficacy may be more important in predicting performance for RD readers than for NA readers.

Of all path coefficients from the independent variables to performance, only those from metacognitive knowledge (standardized solution = 0.25, $t = 2.80$), and efficacy (standardized solution = 0.70, $t = 5.07$) were significant for RD readers. For NA readers, all direct effects on performance were significant with ability rather than metacognitive knowledge or efficacy being the strongest predictor. In contrast, efficacy was the strongest predictor of reading performance for RD readers. The primary distinctions between the two groups were in the different weighting of the paths between ability and attributions, ability and performance, and efficacy and performance.

Next, the amount of variance (squared multiple correlations) accounted for in each latent construct in the model was examined. For NA readers, the squared multiple correlations were .18 for attributions, .58 for efficacy, .71 for metacognitive knowledge, and .49 for performance. For RD readers, the squared multiple correlations were .02 for attributions, .66 for efficacy, .41 for metacognitive knowledge, and .60 for performance. The model accounted for more variance in RD readers' than NA readers' efficacy and performance. In contrast, it accounted for more variance in NA readers' than in RD readers' attributions and metacognitive knowledge. In general, the model accounted for more variance in efficacy than in attributions for both NA and RD readers. The low R^2 value for attributions is to be expected because it is based on only one predictor in the model. It can be concluded that a substantial proportion of the variance in reading achievement was predictable from the model. This was particularly true of RD readers, and such a finding was expected given the restricted

variance of the performance measures of this group. Alternatively, it may suggest that there may be other important variables missing from the NA reader model for predicting performance.

The coefficients of determination for the dependent (endogenous) and independent (exogenous) manifest indicators were 1.00 and .55 (range .50 to .60) respectively for both groups. This evidence suggests that the structural model had a high degree of determinability, both for the predictor and the dependent manifest indicators. Examination of the other detailed measures of fit (modification indices, standardized residuals and Q-Q plot, correlations of estimates) also indicated that the model was reasonably well specified for both groups. The model provided a better fit for NA than RD readers as indicated by the chi-square contribution. The chi-square contribution for NA readers was 41.19% ($\chi^2 = 35.49$) and 58.81% for RD readers ($\chi^2 = 50.67$).

In summary, the relations and hypotheses proposed for Model 1 were well supported with the exception of two hypotheses: (a) that ability directly predicts metacognitive knowledge for NA readers but not for RD readers (as specified by the pathways in the model), and (b) that the relationship between efficacy and reading performance is similar for both NA and RD readers (hypothesis no. 9). This model yielded a number of interesting findings. They are as follows: (a) in addition to an indirect effect, efficacy also has a direct effect on reading performance; (b) attributional style and efficacy differentially predict metacognitive knowledge, with attributional style being the stronger predictor; (c) ability directly predicted self-efficacy; and (d) RD readers differed from NA readers in the strength of effect of the relationships between ability and attributions, ability and reading performance, and efficacy and reading performance.

9.3.2.2 Test of proposed theoretical model 2

(i) *Measurement model*

Model 2 represented an extension of Model 1 by incorporating the latent construct of combined strategy use. This variable represented a global measure of strategic behaviour which included a combination of metacognitive and cognitive strategy use and is consistent with conceptualizations of self-regulation theorists (e.g., Zimmerman, 1989a). Before the structural effects for Model 2 were examined, a final measurement model was specified as

involving six intercorrelated latent variables: ability (one indicator), attributional style (five indicators), self-efficacy (two indicators), metacognitive knowledge (two indicators), combined strategy use (two indicators), and reading performance (two indicators). The behavioural and self-report measures of combined strategy use, each obtained by summing the cognitive and metacognitive strategy use scores, were specified as indicators of the latent construct of combined strategy use. The correlations of the 14 observed measures for each group are given in Table F2 (see Appendix F).

As discussed previously, preliminary analyses indicated that although both the single factor model and the two factor model of strategy use provided reasonably good fit, the two factor model provided a better representation of strategy use (see section 9.3.1.3 (ii)). Nonetheless, for testing Model 2, the one factor model was preferable initially, on the basis of parsimony and its conceptual consistency with theoretical formulations of previous research (Meece et al., 1988).

Similar procedures to those used for Model 1 were followed in specifying the measurement model. For each latent construct, one construct loading was set a priori to unity to establish a scale of measurement for the latent constructs. A number of other parameters were also fixed. As ability was specified by a single indicator, the reliability was fixed by setting the error variance of the ability measure to 33.61 for NA readers and 32.90 for RD readers. The error variance of the PAT indicator was fixed to 8.18 for the RD group, to avoid a Heywood case.

The same methodological arguments underlying the rationale for allowing correlated errors in the self-report questionnaire measures may also be applied to the behavioural measures. It is possible that behaviour scores may be influenced by errors of measurement arising from test unreliability, and method of measurement factors. Hence, as in Model 1, errors of measurement for both self-report and behavioural measures were allowed to covary with each other. In addition to the error covariances in Model 1, Model 2 also included an additional three error covariances for the NA group as a result of the incorporation of the strategy measures. They included errors between self-report combined strategy use and PAT, behavioural strategy use and stable attributions for success, and behavioural combined strategy use and the IRA measure of metacognitive knowledge. As previously mentioned, the inclusion of correlated errors resulted in significant improvement of model fit but had almost no impact on the estimated factor correlations.

The results of the measurement part of Model 1 suggested that the construct patterns would be similar across groups except for failure attributions. For Model 2, with the exception of failure attributions, the construct loadings of all variables including combined strategy use were constrained to be equal across groups. This model provided a good fit to the observed data, $\chi^2(121, N = 395) = 113.51, p = .67$; $GF1 = .96$; $NFI = .96$; $PNFI = .64$; and $RMR = .043$. As these constraints did not significantly erode the fit of the model to the data, it was concluded that NA and RD readers, in general, have a common factor structure for these variables. An examination of the modification indices suggested that the model could not be substantially improved. None exceeded the 5.00 cutoff recommended by Jöreskog and Sörbom (1989). The LISREL specification of this optimal measurement model is presented in section G.2.2.1 of Appendix G.

The goodness-of-fit indices indicate that one combined strategy use factor can parsimoniously summarize different types of strategy use. The alternative model that specified a similar correlational pattern among the six latent variables across groups was not tested as results for measurement Model 1 indicated that this would result in deterioration of fit.

The LISREL estimates for the above optimum measurement model are presented in Tables 12 through 15. Data from Measurement Model 2 appear to replicate the relationships found in Measurement Model 1. All construct loadings ($t \geq 10.01, p \leq .0001$) were highly statistically significant, as is apparent from the small standard errors, and there is good evidence of convergent validity (see Table 12). As these indicators did not load onto factors they were not designed to measure, they showed strong evidence of discriminant validity as well (Bollen, 1989). LISREL estimates for the indicators of attributional style, efficacy, metacognitive knowledge, reading performance, and ability were found to be similar to those in Model 1.

The internal-consistency reliability estimates of the indicators were also computed by LISREL, taking advantage of the latent-variable approach. These estimates are interpreted as a lower bound of the true reliability (Jöreskog & Sörbom, 1989). As shown in Table 12, reliabilities varied considerably from .39 to .83 for RD readers and from .28 to .95 for NA readers. The measures of combined strategy use had reasonably high reliability (range .59 - .73). The variations in the reliabilities, as well as the factor loadings, suggest differential precision in measurement. For example, metacognitive knowledge was measured less well

than most of the other factors, although its correlation with reading performance was consistent with previous studies.

Table 12. Factor loadings and reliabilities of indicators for Measurement Model 2

CONSTRUCT/ indicator	NA Readers			RD readers		
	unstandardized (SE)	standardized ^a	reliability ^b	unstandardized (SE)	standardized ^a	reliability ^b
ATTRIBUTIONS						
Success						
internality	0.85 (0.05)	0.70	.71	0.85 (0.05)	0.70	.28
controllability	1.00 ^c	0.80	.76	1.00 ^c	0.80	.45
stability	0.98 (0.05)	0.80	.77	0.98 (0.05)	0.80	.45
Failure						
controllability	0.60 (0.06)	0.51	.43	1.37 (0.12)	1.17	.68
stability	0.57 (0.06)	0.47	.39	1.55 (0.13)	1.27	.78
EFFICACY						
general	0.05 (0.00)	0.73	.57	0.05 (0.00)	0.73	.46
task-specific	1.00 ^c	0.76	.64	1.00 ^c	0.76	.46
METACOGNITIVE KNOWLEDGE						
MSI	1.00 ^c	0.72	.49	1.00 ^c	0.72	.56
IRA	1.00 (0.09)	0.65	.50	1.00 (0.09)	0.65	.37
COMBINED STRATEGY USE						
self-report	1.00 ^c	0.81	.60	1.00 ^c	0.81	.73
behavioural	3.69 (0.25)	0.81	.72	3.69 (0.25)	0.81	.59
PERFORMANCE						
PAT	1.00 ^c	0.93	.83	1.00 ^c	0.93	.95 ^c
classroom	0.69 (0.05)	0.68	.58	0.69 (0.05)	0.68	.34
ABILITY						
WISC-R	1.00 ^c	0.75	.60 ^c	1.00 ^c	0.75	.50 ^c

All Loadings that were estimated are significant at $p \leq 0.0001$ (2 tailed).

^a Refers to common metric completely standardized solution.

^b Estimated by LISREL (Jöreskog & Sörbom, 1993).

^c Fixed value or reference indicator.

Examination of Table 13 indicates that self-report combined strategy use had lower measurement error variance for RD readers than for NA readers. Conversely, combined strategy use as measured by the interview method had lower measurement error variance for

Table 13. Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>NA readers</i>														
1. contros	0.28													
2. stabs		0.28												
3. stabf			0.48											
4. controf			0.15	0.48										
5. inters					0.28									
6. effica			0.09			0.59								
7. efficb							0.47							
8. msi			0.11					0.53						
9. ira		-0.09							0.42					
10. strat1										0.45				
11. strat2		-0.07							-0.09		0.26			
12. pat							0.22			0.12		0.22		
13. class							0.15						0.44	
14. iq														0.45 ^a
<i>RD readers</i>														
1. contros	0.45													
2. stabs	0.13	0.45												
3. stabf			0.26											
4. controf				0.38										
5. inters				-0.12	0.74									
6. effica						0.34								
7. efficb							0.37							
8. msi								0.42						
9. ira									0.73					
10. strat1										0.24				
11. strat2											0.45			
12. pat												0.03 ^a		
13. class							-0.10						0.64	
14. iq														0.44 ^a

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; strat1 = self-report measure of combined strategy use; strat2 = behavioural measure of combined strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

All residuals are significant at $p \leq 0.05$ (two-tailed) for NA readers and significant at $p \leq 0.01$ (two-tailed) for RD readers.

For purposes of simplicity, all error variances and covariances have been combined into a single matrix.

^a Unstandardized values of these parameters were fixed.

NA readers than for RD readers. The total coefficient of determination for the observed measures of the endogenous constructs was 1.00 for both NA and RD readers. In this case the measurement model performs well, suggesting that the reliability of the measurement model as a whole is exceptionally high. Table 13 shows that 9 errors were allowed to covary in the NA reader group and 3 errors were allowed to covary in the RD reader group. This indicates that the observed scores were affected by generalized response tendencies.

Examination of the variance-covariance matrix of the latent constructs revealed that the variance and covariances of most variables including combined strategy use were greater for NA than RD readers (see Table 14).

Table 14. Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 2

	Attributions	Efficacy	Metacognitive Knowledge	Combined strategy use	Performance	Ability
<i>NA readers</i>						
Attributions	1.41					
Efficacy	1.04	1.45				
Meta. Knowledge	0.90	0.90	0.99			
Combined strategy	0.68	0.91	0.79	1.02		
Performance	0.68	0.81	0.71	0.83	1.28	
Ability	0.55	0.64	0.51	0.53	0.71	1.21
<i>RD readers</i>						
Attributions	0.58					
Efficacy	0.40	0.54				
Meta. Knowledge	0.53	0.46	1.01			
Combined strategy	0.43	0.52	0.75	0.98		
Performance	0.42	0.46	0.51	0.59	0.71	
Ability	0.09 ^a	0.28	0.22 ^a	0.21	0.22	0.79

All estimates are significant at $p \leq .0001$ (two-tailed) for NA readers.

All estimates are significant at $p \leq .05$ (two-tailed) for RD readers except those indicated by ^a.

The intercorrelations of the latent construct are more readily interpretable and are provided in Table 15. In general, the intercorrelations of the latent constructs were moderate and significant indicating the potential for path relationships between variables. Apart from the high correlation between combined strategy use and efficacy, combined strategy use is moderately correlated with metacognitive knowledge and reading performance for both NA and RD readers. For the most part these correlations, although moderately high, suggest that combined strategy use, efficacy, metacognitive knowledge, and performance are distinct dimensions. For both groups, metacognitive knowledge correlated more highly with combined strategy use than it did the other factors. Although the correlation (standardized covariance) between combined strategy use and performance was large, as would be expected, efficacy correlated more highly with performance than combined strategy use in the RD reader group. These findings suggest that although combined strategy use is an important factor in predicting reading performance, efficacy may be more important in predicting performance, particularly for RD readers.

Table 15. Correlations of latent constructs for NA and RD readers for Measurement Model 2

	Attributions	Efficacy	Metacognitive Knowledge	Combined strategy use	Performance	Ability
Attributions		.71	.69	.57	.65	.13 ^a
Efficacy	.73		.62	.71	.73	.42
Meta. Knowledge	.77	.75		.76	.61	.24 ^a
Combined strategy	.56	.75	.78		.70	.23
Performance	.50	.59	.63	.72		.29
Ability	.42	.49	.47	.48	.57	

Triangular correlation matrix for NA readers is presented below the main diagonal and the matrix for RD readers is presented above the main diagonal; all correlations are significant at $p \leq .05$ except for those indicated by * .

The Q-Q plot, modification indices and the correlations of the parameter estimates confirm the good fit of the estimated model to the data. For present purposes, the above measurement model was accepted as the optimum parameter specification for the construct loadings and measurement errors. The pattern of fixed, free, and equated parameters for the measurement model was kept constant in the second part of the analysis described in the next section.

(ii) Structural model

Having established an adequate measurement model, a structural model was tested to explain the relations among the six latent constructs in Model 2 (see Figure 5, p. 89). The hypothesized model specified that attributional style predicts metacognitive knowledge directly or indirectly through self-efficacy, and that combined strategy use mediates the effect of metacognitive knowledge on performance. It was further postulated that efficacy directly predicts metacognitive knowledge, combined strategy use, and reading performance. This model compared NA and RD readers to examine discrete differences in the structural paths that would account for RD readers' deficient reading performance. Hypothesized Model 2 predicted that the six latent constructs would be similarly related to one another in NA and RD readers with the following exceptions: the paths between ability and attributions, ability and metacognitive knowledge, ability and performance, and combined strategy use and performance.

Preliminary analyses of nested models (see Appendix H, section H.2) and the findings of Model 1 indicated that hypothesized Model 2 warranted respecification. The following a posteriori modifications were made: (a) the ability to metacognitive knowledge path, was

fixed to zero; and (b) the efficacy to performance path was freed across groups. This respecified model (Model 2.2) provided a very good fit to the data and represented the final optimum model. The model yielded $\chi^2(137, N = 395) = 134.71, p = .95$; GFI = .95; NFI = .95; PNFI = .72; and RMR = .071. See section G.2.2.2 of Appendix G for the LISREL specification of this best fitting structural model.

Description of results of optimum structural model

Figure 13 shows the common metric completely standardized solution of parameter estimates yielded by the analysis of the optimum structural model for NA and RD readers. All significant effects were positive and in the hypothesized direction (see Appendix J, Table J3, for t values). Measurement errors were again deleted from the figure for clarity, but the error variance-covariance matrix was estimated in the analyses (see Appendix I, Table I2). As mentioned previously, filled arrows connecting circles represent causal effects and unfilled arrows relating circles to rectangles represent the latent variable's effects on indicators. Detached arrows indicate residual variances. Solid lines depict significant effects and dashed lines depict nonsignificant effects.

This model adequately fitted the data and the estimated factor loadings for this model were not significantly different from the final measurement model for all 14 observed indicators. On the basis of these analyses, the data from Model 2 would appear to provide a replication of the relationships found in Model 1 except for the relationship between efficacy and performance. This particular path was no longer significant in NA readers, although it was in Model 1. Instead, the effect of efficacy on performance was further mediated via combined strategy use. This mediating effect of combined strategy use on the path from efficacy to performance is highly significant. The indirect relationships account for most of the total effect of efficacy on performance, thus rendering the direct effect of efficacy on performance statistically nonsignificant for NA readers. Contrary to predictions, efficacy directly predicted performance for RD readers only.

As seen in Figure 13, all causal paths were constrained to be equal across groups except for the following paths: ability to attributions, ability to performance, efficacy to performance, and combined strategy use to performance. In both groups, attributions significantly predicted self-efficacy (standardized solution = 0.64, $t = 10.09$) which in turn, predicted metacognitive

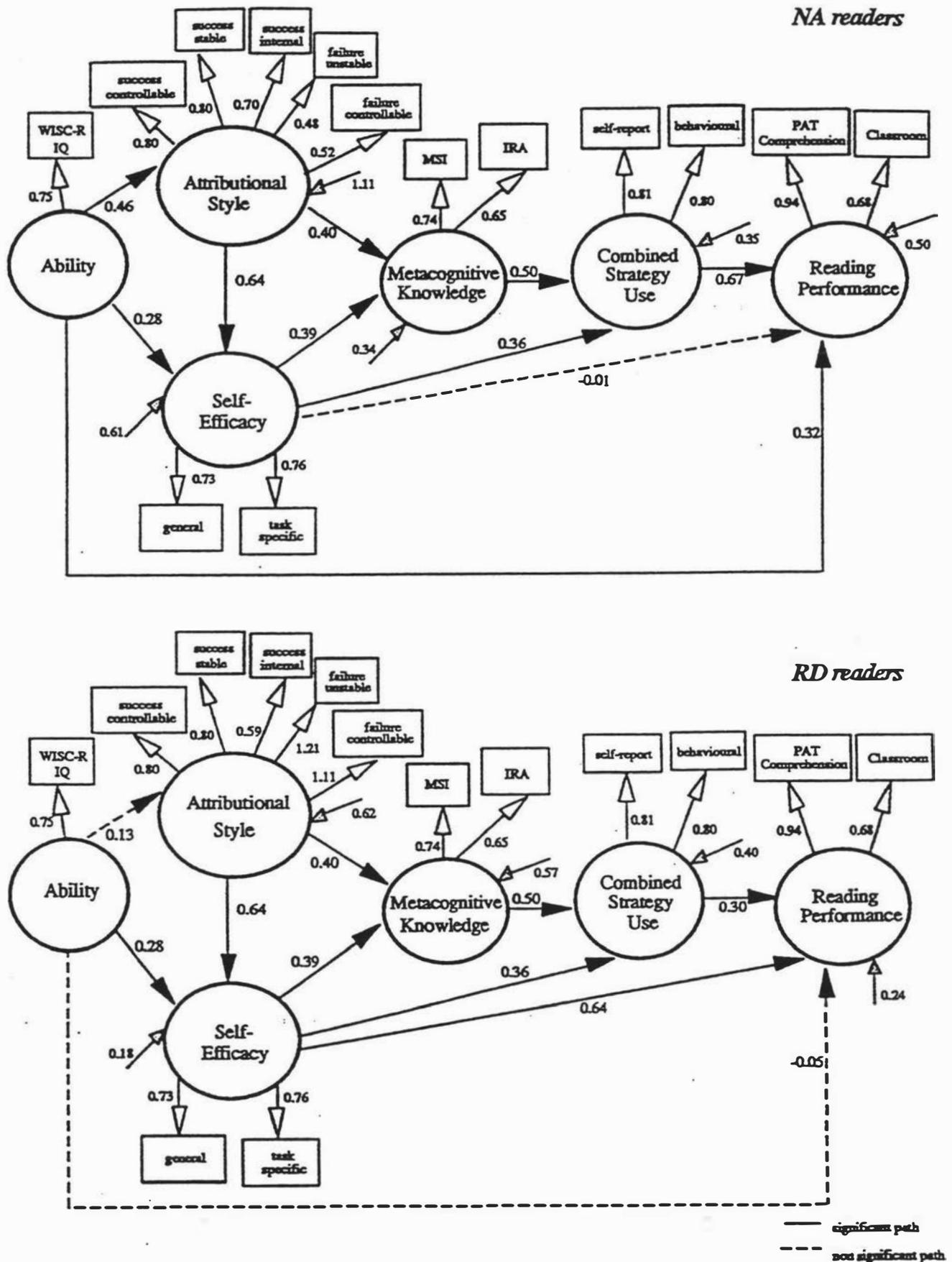


Figure 13. Findings of revised Model 2 for NA and RD readers.

Note: Failure internal attribution was dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero. To avoid clutter, measurement errors are not presented. Common metric completely standardized solution is presented for all parameters.

knowledge (standardized solution = 0.39, $t = 3.87$). The effect of metacognitive knowledge on performance was significantly mediated by combined strategy use for both groups. The direct link between efficacy and combined strategy use was strong and positive (standardized solution = 0.36, $t = 3.61$), which indicates that students with higher self-efficacy were more strategically engaged, irrespective of their skill level. Although the relationship between metacognitive knowledge and reading performance was similar across groups, as indicated by Model 1, the findings of Model 2 suggest that combined strategy use differs in its mediating effect across groups.

Both metacognitive knowledge and self-efficacy predicted combined strategy use, supporting Paris' (1988) suggestion that strategic behaviour is a product of both "skill and will". Metacognitive knowledge was a stronger predictor of combined strategy use than self-efficacy. As hypothesized, the relationship between combined strategy use and performance differed across groups; the coefficient for the direct path between combined strategy use and performance was somewhat larger for NA (standardized solution = 0.67, $t = 4.42$) than for RD readers (standardized solution = 0.30, $t = 3.86$). Efficacy was a stronger predictor of performance (standardized solution = 0.64, $t = 5.50$) than combined strategy use for RD readers, suggesting that motivational rather than cognitive factors may be more important in determining performance for RD readers.

The acceptability of the revised model as an adequate representation of theory was also indicated by the squared multiple correlations, for each of the structural equations. Except for attributions, these were moderately high. For NA readers, the squared multiple correlations were .18 for attributions, .58 for efficacy, .69 for metacognitive knowledge, .70 for combined strategy use, and .62 for performance. For RD readers, the squared multiple correlations were .02 for attributions, .67 for efficacy, .36 for metacognitive knowledge, .52 for combined strategy use, and .63 for performance. On the whole, Model 2 explained 13% more variance in reading performance for NA readers but only 3% more variance for RD readers than did Model 1. This finding again suggests that the incorporation of combined strategy use did not much improve the determinability of the model for RD readers, but did so for NA readers.

Examination of the other detailed measures of fit (modification indices, standardized residuals and Q-Q plot, correlations of estimates) also indicated that the model was reasonably well

specified for both groups. As in the case of Model 1, examination of the chi-square contribution suggested the model provided a better fit for NA readers (chi-square contribution = 42.13%) than for RD readers (chi-square contribution = 57.87%).

In summary, the relations and hypotheses proposed for Model 2 were well supported. Metacognitive knowledge significantly mediated the relationship between motivational beliefs (attributional style, self-efficacy) and combined strategy use. Although the relationship between metacognitive knowledge and performance was found to be similar across groups (as indicated by the findings of Model 1), the mediating role played by combined strategy use in this relationship varied across groups. It can be concluded that there were structural differences between the models for NA and RD readers in the relationships between ability and attributions, ability and reading performance, efficacy and reading performance, and combined strategy use and reading performance.

9.3.2.3 Test of proposed theoretical Model 3

(i) *Measurement model*

Model 3 extended Model 2 by examining the separate effects of the different components of the latent construct of combined strategy use. Before the structural effects for Model 3 were examined, a measurement model was specified as involving seven intercorrelated latent variables: ability, attributional style, self-efficacy, metacognitive knowledge, metacognitive strategy use, cognitive strategy use, and reading performance. Metacognitive strategy use as measured by self-report and by interview were specified as indicators of the metacognitive strategy use factor. Similarly, the interview and self-report measures of cognitive strategy use were specified as indicators of the latent construct of cognitive strategy use. The correlation matrix of the 16 measures for each group is shown in Table F3 (see Appendix F).

Similar procedures to those used for Models 1 and 2 were followed in specifying the measurement model. For each latent construct, one construct loading was set a priori to unity. Each measure was allowed to load on only one factor, with the loadings on other factors fixed to zero. As with Model 2, the error variance of ability was fixed to 33.61 for NA readers and 32.90 for RD readers; the error variance of the PAT indicator was fixed to 8.18 for the RD reader group.

As before, errors were allowed to covary in Model 3. Thirteen errors were allowed to covary in the measurement model for NA readers and 8 were allowed to covary in the measurement model for RD readers (see Table 17, p. 214). Although errors of the behavioural measures were allowed to intercorrelate within each group on a priori grounds, modification indices suggested that relaxing these errors would not result in a further significant improvement in goodness-of-fit. Consequently these error covariances were excluded from the model. As with Models 1 and 2, the construct loadings of all variables except failure attributions were constrained to be equal across groups for Model 3. This measurement model accurately reproduced observed covariances; its goodness-of-fit surpassed the .90 indicative of well-fitting models. This model yielded, $\chi^2(155, N = 395) = 122.43, p = .97$; $GF1 = .96$; $NFI = .96$; $PNFI = .62$; and $RMR = .042$. An examination of the modification indices suggested that the model could not be substantially improved. None exceeded the 5.00 cutoff recommended by Jöreskog and Sörbom (1989). No further tests of invariance were done as results of measurement Model 1 indicated that further improvement of fit would be unlikely. See section G.3.2.1 of Appendix G for the LISREL specification of the optimal measurement model.

The LISREL estimates of the construct loadings, error/factor variances and covariances, and correlations of latent constructs for the optimum measurement model are presented in Tables 16 - 19. These LISREL estimates were similar to those found in Models 1 and 2. Table 16 shows generally sizeable factor loadings for both groups thus establishing convergent validity. Moreover, no confidence interval computed around the factor correlations included 1.0 for both groups, signalling discriminant validity (J. C. Anderson & Gerbing, 1988). All factor loadings were highly significant ($t \geq 10.03, p \leq .0001$) as can be seen by the small standard errors, and confirms the hypothesized factor structure.

The measures of metacognitive and cognitive strategy use in both groups had low to moderate reliability (.45 - .70). Examination of Table 16 indicates that self-report metacognitive and cognitive strategy use had higher reliability estimates for RD readers than for NA readers. Conversely, metacognitive and cognitive strategy use as measured by the interview method had higher reliability for NA readers than for RD readers.

Table 16. Factor loadings and reliabilities of indicators for Measurement Model 3

CONSTRUCT/ indicator	NA Readers			RD readers		
	unstandardized (SE)	standardized ^a	reliability ^b	unstandardized (SE)	standardized ^a	reliability ^b
ATTRIBUTIONS						
Success						
internality	0.84 (0.05)	0.69	.71	0.84 (0.05)	0.69	.28
controllability	1.00 ^c	0.80	.77	1.00 ^c	0.80	.46
stability	0.98 (0.05)	0.79	.76	0.98 (0.05)	0.79	.46
Failure						
controllability	0.60 (0.06)	0.51	.43	1.35 (0.12)	1.16	.67
stability	0.57 (0.06)	0.47	.39	1.52 (0.12)	1.25	.76
EFFICACY						
general	0.06 (0.00)	0.76	.58	0.06 (0.00)	0.76	.48
task-specific	1.00 ^c	0.75	.62	1.00 ^c	0.75	.45
METACOGNITIVE KNOWLEDGE						
MSI	1.00 ^c	0.73	.50	1.00 ^c	0.73	.57
IRA	0.99 (0.09)	0.64	.50	0.99 (0.09)	0.64	.37
METACOGNITIVE STRATEGY USE						
self-report	1.00 ^c	0.78	.54	1.00 ^c	0.78	.70
behavioural	4.61 (0.38)	0.72	.59	4.61 (0.38)	0.72	.45
COGNITIVE STRATEGY USE						
self-report	1.00 ^c	0.80	.59	1.00 ^c	0.80	.66
behavioural	2.29 (0.18)	0.75	.69	2.29 (0.18)	0.75	.49
PERFORMANCE						
PAT	1.00 ^c	0.94	.85	1.00 ^c	0.94	.95 ^c
classroom	0.67 (0.05)	0.67	.56	0.67 (0.05)	0.67	.33
ABILITY						
WISC-R	1.00 ^c	0.75	.60 ^c	1.00 ^c	0.75	.50 ^c

All Loadings that were estimated are significant at $p \leq 0.0001$ (2 tailed).

^a Refers to common metric completely standardized solution.

^b Estimated in LISREL (Jöreskog & Sörbom, 1993).

^c Fixed value or reference indicator.

Table 17 presents the variances and covariances of the measurement errors. As for Models 1 and 2, there is evidence of a patterning of nonzero error covariances, suggesting evidence of shared variance in the manifest indicators which is not explained by the latent variables of the model.

Table 17. Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
<i>NA readers</i>																	
1. contros	0.27																
2. stabs		0.27															
3. stabf			0.49														
4. controf			0.16	0.48													
5. inters					0.27												
6. effica			0.09			0.60											
7. efficb					0.07		0.49										
8. msi			0.11					0.53									
9. ira		-0.10							0.41								
10. meta						-0.12				0.59							
11. coga					0.10					0.13	0.42						
12. metb									-0.11			0.41					
13. cogb		-0.08											0.24				
14. pat							0.25			0.15				0.19			
15. class							0.18								0.46		
16. iq																0.45*	
<i>RD readers</i>																	
1. contros	0.46																
2. stabs	0.13	0.43															
3. stabf			0.29														
4. controf				0.38													
5. inters				-0.14	0.71												
6. effica						0.35											
7. efficb			0.06				0.38										
8. msi								0.41									
9. ira									0.73								
10. meta								-0.09		0.23							
11. coga				-0.08							0.36						
12. metb												0.56					
13. cogb												0.12	0.63				
14. pat														0.03*			
15. class							-0.09					0.10			0.64		
16. iq																0.44*	

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; meta = self-report measure of metacognitive strategy use; coga = self-report measure of cognitive strategy use; metb = behavioural measure of metacognitive strategy use; cogb = behavioural measure of cognitive strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R. All residuals are significant at $p \leq .05$ (two-tailed).

For purposes of simplicity, all error variances and covariances have been combined into a single matrix.

* Unstandardized values of these parameters were fixed.

As with Models 1 and 2, the variances/covariances of the latent constructs in the NA reader group are generally higher than those in the RD reader group (see Table 18). Conceptual distinctiveness for the model constructs is suggested by the data in Table 19. In general, the

latent constructs for each group were moderately intercorrelated. Metacognitive strategy use is strongly correlated with metacognitive knowledge, efficacy, reading performance, and cognitive strategy use for both NA and RD readers (see Table 19). For both groups, compared to metacognitive strategy use, correlations between cognitive strategy use and the other latent constructs were lower. This finding suggests that in general metacognitive strategy use may be more strongly related to the other latent constructs than cognitive strategy use. The markedly different correlations between metacognitive knowledge and the two types of strategy use imply that metacognitive knowledge may be a better predictor of metacognitive than of cognitive strategy use.

Table 18. Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 3

	Attributions	Efficacy	Metacognitive Knowledge	Metacognitive strategy use	Cognitive Performance strategy use	Ability
<i>NA readers</i>						
Attributions	1.41					
Efficacy	1.03	1.45				
Meta. Knowledge	0.90	0.89	0.99			
Metacognitive	0.72	0.95	0.76	1.13		
Cognitive	0.44	0.76	0.64	0.69	0.93	
Performance	0.68	0.79	0.70	0.81	0.76	1.29
Ability	0.55	0.66	0.51	0.52	0.47	1.21
<i>RD readers</i>						
Attributions	0.59					
Efficacy	0.39	0.54				
Meta. Knowledge	0.53	0.45	1.01			
Meta. strategy use	0.47	0.49	0.74	0.87		
Cog. strategy use	0.36	0.45	0.68	0.68	1.07	
Performance	0.42	0.45	0.51	0.53	0.53	0.70
Ability	0.09 ^a	0.27	0.21 ^a	0.16 ^a	0.21 ^a	0.22

All estimates are significant at $p \leq .001$ (two-tailed) for NA readers.

All estimates are significant at $p \leq .01$ (two-tailed) for RD readers.

^a Indicates that estimates are not significant.

Inspection of the standardized residuals and the Q-Q plot showed that the model was well fitting, as 95% of the residuals were within 2 standard deviations of zero. Modification indices and the correlations of the parameter estimates also confirm the good fit of the estimated model to the data. For present purposes, the above measurement model was accepted as best fitting the data. The pattern of fixed, free, and equated parameters for the measurement model was kept constant (but not fixed) in the second part of the analysis.

Table 19. Correlations of latent constructs for NA and readers for Measurement Model 3

	Attributions	Efficacy	Metacognitive Knowledge	Metacognitive strategy use	Cognitive Performance strategy use	Ability
Attributions		.69	.69	.65	.45	.65
Efficacy	.72		.61	.71	.60	.42
Meta. Knowledge	.76	.74		.79	.66	.24 ^a
Meta. strategy use	.57	.74	.72		.71	.20 ^a
Cog. strategy use	.38	.65	.67	.67		.23 ^a
Performance	.51	.58	.62	.67	.69	.29
Ability	.42	.50	.47	.44	.44	.57

Triangular correlation matrix for NA readers is presented below the main diagonal and the matrix for RD readers is presented above the main diagonal. All correlations are significant at $p \leq .01$ (two-tailed) except for those indicated by ^a.

(ii) *Structural model*

Theoretical Model 3 (see Figure 6, p. 91) specified that motivational beliefs (attributions, efficacy) predict metacognitive knowledge, and that metacognitive strategy use mediates the effect of metacognitive knowledge on performance. It was further postulated that efficacy directly predicts strategy use (both cognitive and metacognitive) and reading performance. It was also theorized that the model would be essentially the same for NA and RD readers except for the following paths: ability to attributions, ability to metacognitive knowledge, ability to performance, and cognitive strategy use to performance. These paths were hypothesized to be statistically nonsignificant for RD readers but significant for NA readers.

In addition, for this model, the disturbances of metacognitive and cognitive strategy use factors were allowed to covary (represented in the LISREL specification by freeing the corresponding off-diagonal elements of the Ψ matrix; see Appendix G, section G.3.2.3). This assumption was necessary because the model equations were such that the causally antecedent latent variables of metacognitive knowledge, efficacy, attributional style and ability were insufficient to explain all of the estimated covariation/correlation between metacognitive and cognitive strategy use.

A number of competing models were initially tested, the results of which are summarized in Table H4 (see Appendix H, section H.3). The findings of the two previously tested models (Models 1 and 2), and the competing nested models suggested that hypothesized Model 3 warranted respecification. More specifically, ability did not directly predict metacognitive knowledge and the relationship between efficacy and performance was not similar across groups. This respecified model (Model 3.2) provided a good fit to data. The model yielded, $\chi^2(178, N = 395) = 159.01, p = .84, GF1 = .95, NFI = .95, PNFI = .70,$ and $RMR = .074$. A LISREL specification of this optimum model is provided in section G.3.2 of Appendix G.

Description of results of optimum structural model

Comparison of the nested models indicated that Model 3.2 was the best fitting structural model (see Appendix H, section H.3). The common metric completely standardized solution of the parameter coefficients of this revised structural model is presented in Figure 14 for NA and RD readers. All significant effects were positive and in the hypothesized direction (see Appendix J, Table J4, for t values). Although errors of measurement and covariances of the structural disturbances were deleted from the figure for clarity, the measurement error variance-covariance matrix (see Appendix I, Table I3) and the structural disturbance variance-covariance matrix were estimated in the analysis. The covariance between the disturbances of the metacognitive and cognitive strategy use factors was significant ($\psi = 0.16$, $t = 2.13$ for NA readers; $\psi = 0.18$, $t = 2.55$ for RD readers) indicating the existence of sources of variance in common between metacognitive and cognitive strategy use which were not explained by the predictor variables in the model.

The respecified model adequately fitted the data and the estimated factor loadings were not markedly different from the final measurement model for all 16 observed indicators. As mentioned earlier, the parameter estimates of the measurement model may not be identical to those in this model as a full system of equations containing both measurement and structural effects were solved here. All hypothesized loadings were highly significant ($p \leq .0001$). The findings of Model 3 appear to provide a replication of the structural relationships found in Model 2. Model 3 also clarified further the relationship between combined strategy use and performance. Essentially, the findings suggested that the dissimilarity in the relationship between strategy use and performance across groups was due to one component of strategy use, namely, cognitive strategy use. For NA readers cognitive strategy use was a significant predictor of reading performance whereas for RD readers it was not.

In both groups, ability significantly predicted self-efficacy (standardized solution = 0.27, $t = 4.09$) which in turn, predicted metacognitive knowledge (standardized solution = 0.36, $t = 3.82$). The relationship between motivational beliefs and metacognitive/cognitive strategy use was significantly mediated by metacognitive knowledge. As hypothesized, efficacy had a direct and positive effect on both metacognitive and cognitive strategy use, thus supporting previous empirical research. However, this finding is contrary to Borkowski et al.'s (1989, 1990, 1992) theoretical prediction which suggests that efficacy does not directly predict cognitive strategy use.

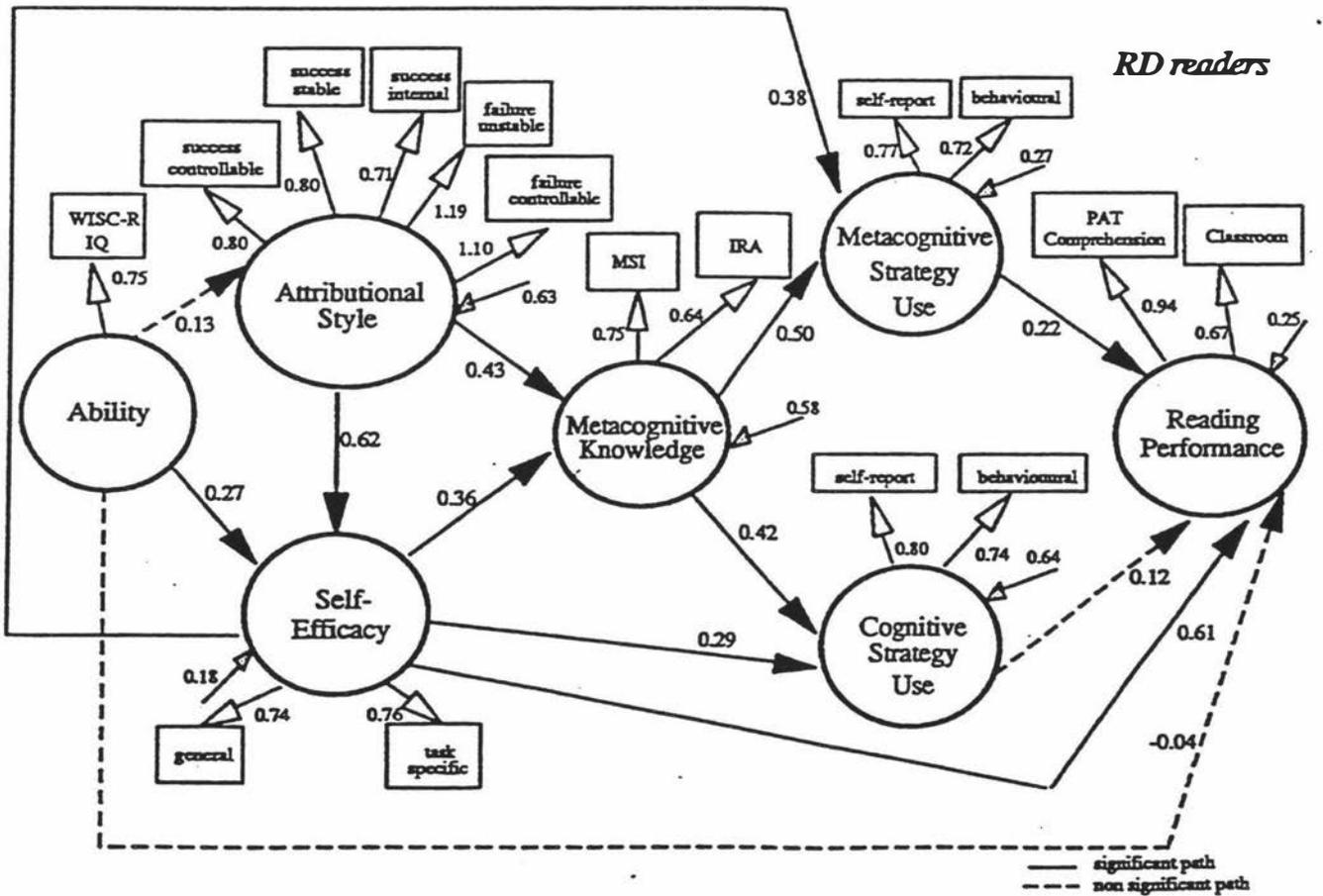
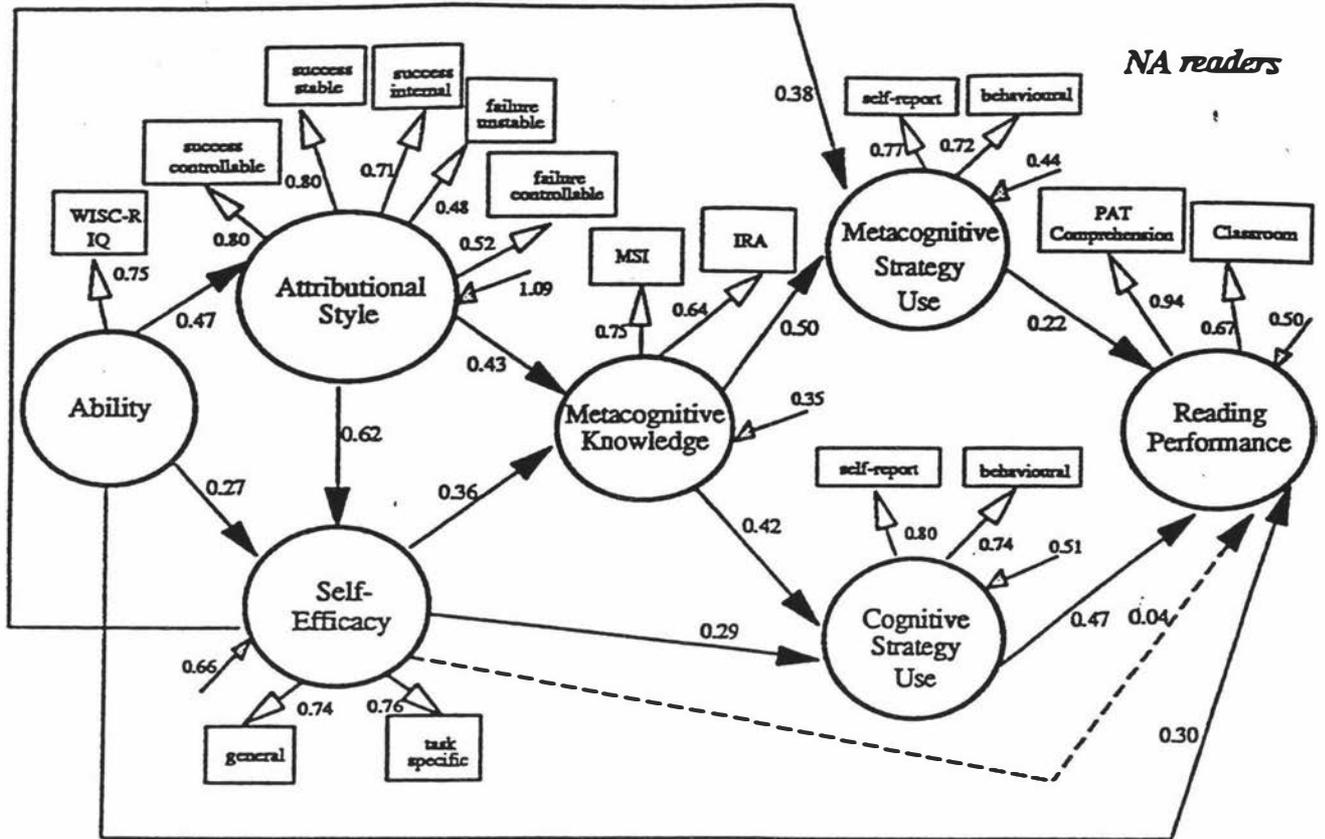


Figure 14. Findings of revised Model 3 for NA and RD readers.

Note: Failure internal attribution was dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero. To avoid clutter, measurement errors and correlated disturbances are not presented. Common metric completely standardized solution is presented for all parameters.

The direct effects of efficacy and metacognitive knowledge on both metacognitive and cognitive strategy use were all positive and significant as hypothesized. Metacognitive knowledge had stronger direct effects on metacognitive strategy use (standardized solution = 0.50, $t = 4.50$) and cognitive strategy use (standardized solution = 0.42, $t = 4.02$) than did any of the other variables in the study. The significant relationship between metacognitive knowledge and cognitive strategy use for RD readers provides indirect support for the inefficient learner view of reading disability.

The findings of Model 3 provide some explication of the observations from Model 2 that the relationship between combined strategy use and reading performance was not similar across groups. Examination of the subcomponents of combined strategy use revealed that only one of the subcomponents of strategy use (cognitive strategy use) contributed to the dissimilarity in the relationship between combined strategy use and performance across groups. The relationship between metacognitive strategy use and performance was similar across groups. More specifically, whereas metacognitive strategy use is significantly associated with reading performance in both groups, cognitive strategy use is significantly related to performance for NA readers only. The nonsignificant relationship between cognitive strategy use and performance in RD readers supports both the "inefficient" and "inactive" learner views of reading disability.

Motivational and metacognitive/cognitive strategy use variables varied in their impact on reading performance across groups. For RD readers, self-efficacy had a stronger direct effect on reading performance (standardized solution = 0.61, $t = 5.14$) than either metacognitive (standardized solution = 0.22, $t = 2.39$) or cognitive strategy use (standardized solution = 0.12, $t = 1.63$). This finding implies that motivational variables may be more important in determining the level of performance for RD readers. Conversely, both metacognitive (standardized solution = 0.22, $t = 2.39$) and cognitive strategy use (standardized solution = 0.47, $t = 4.12$) variables had stronger direct effects on performance than self-efficacy (standardized solution = 0.04, $t = 0.35$) for NA readers.

Both metacognitive and cognitive strategy use had differential effects on performance and these effects varied across groups. As hypothesized, cognitive strategy use was a stronger predictor of reading performance than metacognitive strategy use for NA readers. The converse was true for RD readers. The primary distinctions between the two groups were in the different weighting of the paths between ability and attributions, ability and performance,

efficacy and performance, and cognitive strategy use and performance. For RD readers, with the exception of the path efficacy to performance, the aforementioned paths were not statistically significant.

More than 60% of the variance in reading performance was accounted for in both groups. The coefficient of determination for the dependent (endogenous) and independent (exogenous) manifest indicators were 1.00 and .55 (range .50 to .60) respectively for both groups. This evidence suggests that the structural model not only fitted the data well, but that it also had a high degree of determinability, both for the predictor and the dependent manifest indicators.

In summary, the relations proposed in Theoretical Model 3 were generally well supported. It can be concluded that: (a) metacognitive and cognitive strategy use differentially affect reading performance; (b) efficacy has a direct effect on both metacognitive and cognitive strategy use; (c) metacognitive knowledge significantly mediates the relationship between motivational beliefs and metacognitive/cognitive strategy use; (d) metacognitive strategy use mediates the relationship between metacognitive knowledge and reading performance; (e) the relationship between cognitive strategy use and performance is not similar across groups; (f) the relationship between metacognitive knowledge and cognitive strategy use in RD readers supports the "inefficient" learner view; (g) motivational and cognitive/metacognitive strategy use variables may vary in their importance in determining performance across groups.

9.3.2.4 Test of proposed theoretical Model 4

(i) *Measurement model*

Model 4 examined more precisely the relationships between different dimensions of success and failure attributions on other latent constructs of interest. Before the structural effects for Model 4 were examined, a measurement model was specified as involving eight intercorrelated latent variables: ability, attributions for success outcome, attributions for failure outcome, self-efficacy, metacognitive knowledge, metacognitive strategy use, cognitive strategy use, and reading performance. In line with previous research (Hull & Mendolia, 1991), the construct of attributional style was conceptualized as separate factors for success and failure outcomes. Internal, controllable and stable attributions for success were allowed to load on the success attributional style factor whilst controllable and unstable attributions for failure were allowed to load on the failure attributional style factor.

Similar procedures to those used in previous models were followed in specifying the measurement model. For each latent construct, one construct loading was set a priori to unity. Each manifest variable was allowed to load onto one and only one of the eight latent constructs; its loadings onto other factors were fixed to zero. The error variances of ability for NA and RD readers, and the error variance of the PAT for RD readers, were fixed to values identical to those used in Models 2 and 3.

LISREL analyses of Models 1, 2, and 3 suggested inequality of failure attributions loadings across groups for a single latent factor model of attributions. However, preliminary LISREL analyses indicated equality of failure attributions loadings for the two latent factors model of attributional style. Consequently, a model that specified all factor loadings (including that of failure attributions) to be invariant across groups was tested. This model yielded $\chi^2(146, N = 395) = 121.56, p = .93$; GFI = .96; NFI = .96; PNFI = .58; and RMR = .040. Altogether thirteen errors in the NA reader group and four in the RD reader group were allowed to covary (see Table 21, p. 223). Again, as in the preceding models, the inclusion of correlated errors resulted in significant improvement in model fit but had almost no impact on the estimated latent factor correlations. An examination of the modification indices suggested that the model could not be substantially improved. None exceeded the 5.00 cutoff recommended by Jöreskog and Sörbom (1989).

At first sight, the finding of total factorial invariance across groups seems inconsistent with that from preceding models, which suggested different factor loadings for the failure attribution measures between groups. However, this can be explained by the fact that in previous models group differences were reflected in higher factor loadings for failure measures in the RD reader group, but in Model 4 these differences are now reflected in the variance of the failure attributional style construct for the two groups. The variance of this construct in the RD reader group is substantially higher than for the NA reader group (see Table 22, p. 224). See section G.4.2.1 of Appendix G, for the LISREL specification of the best fitting model.

The LISREL estimates for the optimum measurement model which specified total invariance of factor loadings across groups are presented in Tables 20 through 23. These LISREL estimates are similar to those in previous models except for the attribution indicators. Table 20 shows generally much larger factor loadings for the attributional style indicators compared to those of previous models. Although the parsimony fit index suggests that Measurement Model 3 is more parsimonious than Measurement Model 4, a comparison of the NFI and GFI goodness-of-fit indices suggests that the two models may be comparable, whilst the RMR

index suggests that Model 4's two factor model of attributions is better. Factor loadings of all indicators were highly significant ($t \geq 10.73$, $p \leq .0001$) which confirmed the hypothesized factor structure.

Table 20. Factor loadings and reliabilities of indicators for Measurement Model 4

CONSTRUCT/ indicator	NA readers			RD readers		
	unstandardized (SE)	standardized ^a	reliability ^b	unstandardized (SE)	standardized ^a	reliability ^b
SUCCESS						
ATTRIBUTIONS						
internality	0.82 (0.05)	0.71	.70	0.82 (0.05)	0.71	0.32
controllability	1.00 ^c	0.84	.77	1.00 ^c	0.84	0.58
stability	0.98 (0.05)	0.83	.76	0.98 (0.05)	0.83	0.59
FAILURE						
ATTRIBUTIONS						
controllability	1.09 (0.07)	0.80	.61	1.09 (0.07)	0.80	.78
stability	1.00 ^c	0.84	.57	1.00 ^c	0.84	.69
EFFICACY						
general	1.00 ^b	0.74	.58	1.00	0.74	.47
task-specific	18.08 (1.42)	0.75	.63	18.08 (1.42)	0.75	.45
METACOGNITIVE						
KNOWLEDGE						
MSI	1.00 ^c	0.74	.53	1.00 ^c	0.74	.57
IRA	0.97 (0.09)	0.65	.51	0.97 (0.09)	0.65	.36
METACOGNITIVE						
STRATEGY USE						
self-report	1.00 ^c	0.78	.55	1.00 ^c	0.78	.71
behavioural	4.53 (0.37)	0.71	.59	4.53 (0.37)	0.71	.44
COGNITIVE						
STRATEGY USE						
self-report	1.00 ^c	0.79	.59	1.00 ^c	0.79	.67
behavioural	2.26 (0.18)	0.75	.68	2.26 (0.18)	0.75	.49
PERFORMANCE						
PAT	1.00 ^c	0.94	.86	1.00 ^c	0.94	.95 ^c
classroom	0.67 (0.05)	0.67	.56	0.67 (0.05)	0.67	.33
ABILITY						
WISC-R	1.00 ^c	0.75	.60 ^c	1.00 ^c	0.75	.50 ^c

All Loadings that were estimated are significant at $p \leq 0.0001$ (2 tailed).

^a Refers to common metric completely standardized solution.

^b Estimated by LISREL (Jöreskog & Sörbom, 1993).

^c Fixed value or reference indicator.

The statistically significant variances and covariances of the measurement errors ($p \leq .05$) suggest that the observed scores in the two groups were subject to measurement error and

affected by generalized response tendencies (see Table 21). Inclusion of correlated errors resulted in significant improvement of model fit; however, as in previous models, the underlying pattern of covariances estimated between the latent factors was not noticeably affected regardless of the error covariance structure imposed.

Table 21. Common metric completely standardized solution of error variances and covariances for NA and RD readers for Measurement Model 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
<i>NA readers</i>																		
1. contros	0.27																	
2. stabs		0.28																
3. stabf			0.34															
4. controf				0.33														
5. inters					0.28													
6. effica						0.11	0.57											
7. efficb							0.07	0.49										
8. msi									0.50									
9. ira										0.42								
10. meta											0.57							
11. coga							0.10					0.13	0.40					
12. metb														0.41				
13. cogb															0.25			
14. pat																0.19		
15. class																	0.46	
16. iq																		0.45 ^a
<i>RD readers</i>																		
1. contros	0.34																	
2. stabs		0.33																
3. stabf			0.38															
4. controf				0.25														
5. inters					0.73													
6. effica						0.33												
7. efficb							0.37											
8. msi								0.40										
9. ira									0.74									
10. meta										0.21								
11. coga											0.34							
12. metb												0.57						
13. cogb													0.12	0.64				
14. pat															0.03 ^a			
15. class																0.64		
16. iq																	0.44 ^a	

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; meta = self-report measure of metacognitive strategy use; coga = self-report measure of cognitive strategy use; metb = behavioural measure of metacognitive strategy use; cogb = behavioural measure of cognitive strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R. All residuals are significant at $p \leq .05$ (two-tailed).

For purposes of simplicity, all error variances and covariances have been combined into a single matrix.

^a Unstandardized values of these parameters were fixed.

Table 22 indicates that the variances and covariances of the latent constructs were generally higher for NA readers than for RD readers. Examination of the within group completely standardized solution of the factor variance-covariance matrix (see Table 23) indicated success and failure attributions were highly correlated for both groups (.82 for NA readers and .85 for RD readers). Correlations between success attributions and the latent constructs of efficacy (.73), metacognitive strategy use (.57), and performance (.50) were higher than the correlations of these latent constructs with failure attributions for NA readers. For RD readers, failure attributions had higher correlations with self-efficacy (.70), metacognitive knowledge (.70), and performance (.64) than success attributions. These findings suggest that success and failure attributions may differentially predict self-efficacy across groups, and that failure rather than success attributions may be more important in predicting metacognitive knowledge. Ability correlated highly and comparably with both success (.41) and failure (.42) attributions for NA readers. For RD readers, ability had low correlations with success (.21) and failure (.11) attributions.

Table 22. Common metric completely standardized variances and covariances of latent constructs for NA and RD readers for Measurement Model 4

	Success Attributions	Failure Attributions	Efficacy	Metacognitive Knowledge	Metacognitive strategy use	Cognitive strategy use	Performance	Ability
<i>NA readers</i>								
Success Attrib.	1.32							
Failure Attrib.	0.80	0.71						
Efficacy	1.00	0.62	1.45					
Meta. Knowledge	0.88	0.68	0.91	1.02				
Meta. strategy	0.69	0.49	0.96	0.77	1.13			
Cog. strategy	0.41	0.32	0.75	0.64	0.69	0.93		
Performance	0.66	0.47	0.79	0.72	0.81	0.76	1.30	
Ability	0.52	0.39	0.66	0.52	0.52	0.47	0.71	1.21
<i>RD readers</i>								
Success Attrib.	0.68							
Failure Attrib.	0.79	1.29						
Efficacy	0.40	0.50	0.54					
Meta. Knowledge	0.45	0.79	0.45	0.98				
Meta. strategy	0.49	0.65	0.48	0.72	0.87			
Cog. strategy	0.39	0.47	0.45	0.67	0.69	1.07		
Performance	0.41	0.61	0.45	0.50	0.53	0.52	0.70	
Ability	0.16 ^a	0.12 ^a	0.28	0.21 ^a	0.16 ^a	0.23	0.22	0.79

All estimates are significant at $p \leq .001$ (two-tailed) for NA readers.

All estimates are significant at $p \leq .05$ (two-tailed) for RD readers.

^a Indicates that estimates are not significant.

Table 23. Correlations of latent constructs for NA and RD readers for Measurement Model 4

	Success Attributions	Failure Attributions	Efficacy	Metacognitive Knowledge	Metacognitive strategy use	Cognitive strategy use	Performance	Ability
Success Attrib.		.85	.66	.55	.64	.45	.59	.21 ^a
Failure Attrib.	.82		.70	.70	.61	.39	.64	.11 ^a
Efficacy	.73	.61		.62	.71	.60	.73	.43
Meta. Knowledge	.76	.80	.75		.78	.66	.60	.24 ^a
Meta. strategy	.57	.55	.75	.72		.71	.68	.20 ^a
Cog. strategy	.37	.39	.65	.66	.67		.60	.24
Performance	.50	.49	.58	.62	.67	.69		.29
Ability	.41	.42	.50	.47	.45	.44	.57	

Triangular correlation matrix for NA readers is presented below the main diagonal and the matrix for RD readers is presented above the main diagonal. All correlations are significant except for those indicated by ^a.

Inspection of the standardized residuals and the Q-Q plot suggested that the model fitted well, since 95% of the residuals were within 2 standard deviations of the mean. Modification indices and the correlations of the parameter estimates also confirm the good fit of the estimated model to the data. For present purposes, the above measurement model was accepted as the optimum parameter specification for the construct loadings and measurement errors. This pattern of fixed, free, and equated parameters for the measurement model was kept constant in the second part of the analysis described in the following section.

(ii) *Structural model*

In this model, structural parameters were incorporated into the model and were again estimated simultaneously for both groups. In Model 4, attributional style was respecified to form two latent variables, one for success outcomes and the other for failure outcomes. Previous studies (Berndt & Miller, 1990; Carr et al., 1991) had conceptualized both success and failure attributions as imperfect measures of a more global attributional construct. An implicit assumption underlying attributions in these studies is that the responses to attribution items for both failure and success outcomes are reasonably unidimensional. On the other hand, Sweeney et al. (1986) have provided evidence from the depression literature that information about attributions from different outcomes may be differentially related to other latent constructs of interest. Thus, Model 4 was tested to investigate the relationships between attributions for different outcomes (success vs failure) and ability, self-efficacy, metacognitive knowledge, cognitive/metacognitive strategy use, and reading performance. For Model 4, the disturbances were allowed to covary between the cognitive and metacognitive strategy use

factors, and also between the success and failure attributional style factors (see Appendix G, section G.4.2.3). Correlated structural disturbances were allowed since it was assumed that there was common unexplained variance shared between the subcomponents of attributional style and between the subcomponents of combined strategy use factors that was not accounted for by the other variables in the model.

Theoretical Model 4 (see Figure 7, p. 94) specified that both success and failure attributions predict efficacy and metacognitive knowledge. It was also posited that both these components of attributions predict metacognitive knowledge indirectly through self-efficacy, and that metacognitive strategy use mediates the effect of metacognitive knowledge on performance. The relationships between success/failure attributions and efficacy were hypothesized to be similar across groups, as were the relationships between success/failure attributions and metacognitive knowledge. The paths ability to attributions (both success and failure), ability to metacognitive knowledge/performance, and cognitive strategy use to performance were hypothesized to be statistically nonsignificant for RD readers but significant for NA readers.

A nested sequence of competing models was tested initially, the details of which are presented in section H.4 of Appendix H. Findings of the competing models as well as the results of Models 1 through 3 indicated that hypothesized Model 4 warranted respecification. Based on the results the following respecifications were made: (a) the path ability to metacognitive knowledge was fixed to zero, (b) the path efficacy to performance was freed across groups, (c) the paths success/failure attributions to efficacy was freed across groups, and (d) the path success attributions to metacognitive knowledge was fixed to zero. This respecified model (Model 4.5) provided a substantially good fit to the data and represented the final optimum model. The model yielded $\chi^2(176, N = 395) = 159.67, p = .81, GFI = .95, NFI = .95, PNFI = .69,$ and $RMR = .050$ (see Appendix H, section H.4, Table H5).

Description of results of optimum structural model

Comparison of the nested models indicated that Model 4.5 provided the best fit to the data (see Appendix H, Table H5). Figure 15 presents the results of this revised structural model for NA and RD readers. All significant effects were positive and in the hypothesized direction (see Appendix J, Table J5, for t values, and within group standardized solution). Variances and covariances of measurement errors (see Appendix I, Table I4) as well as the

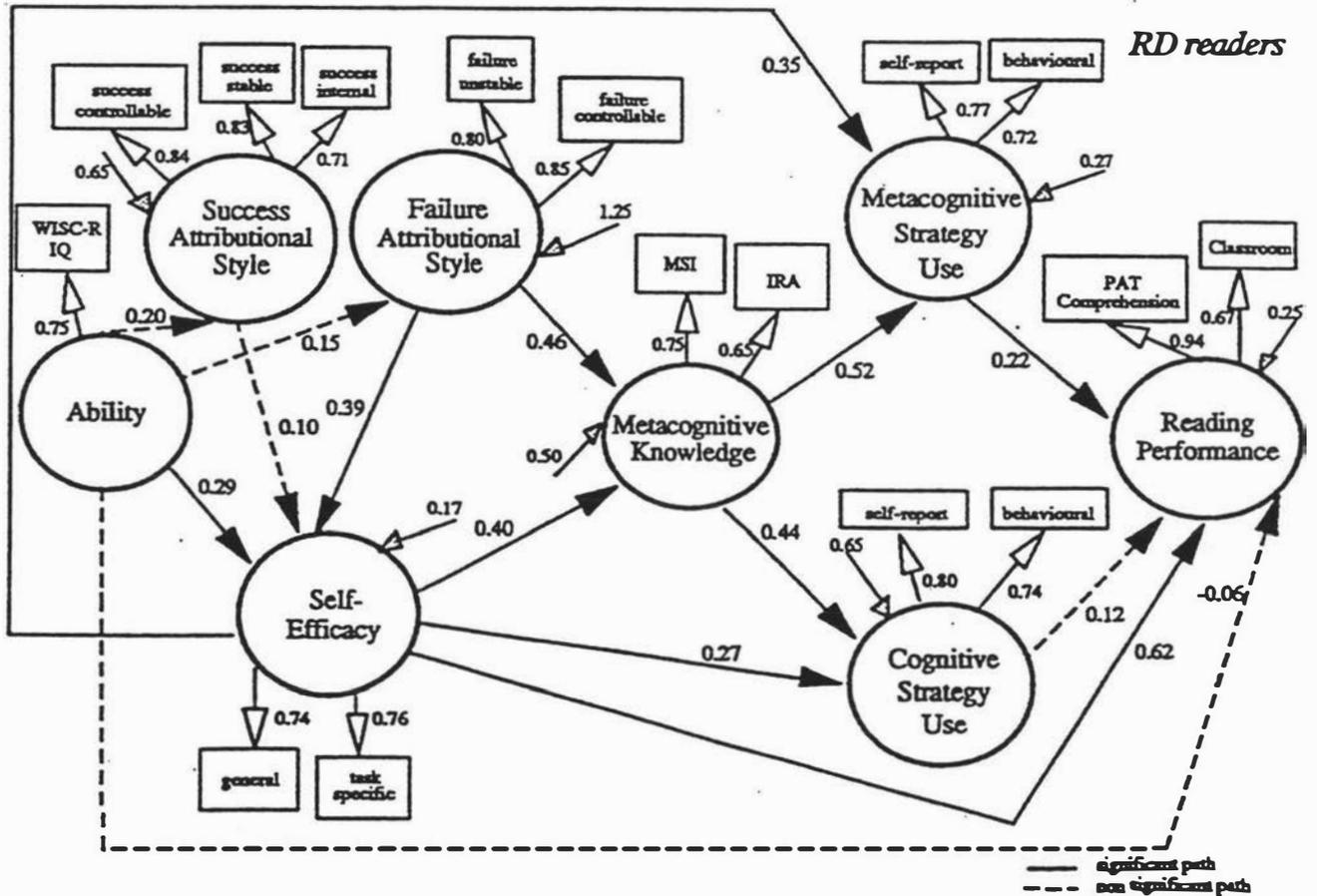
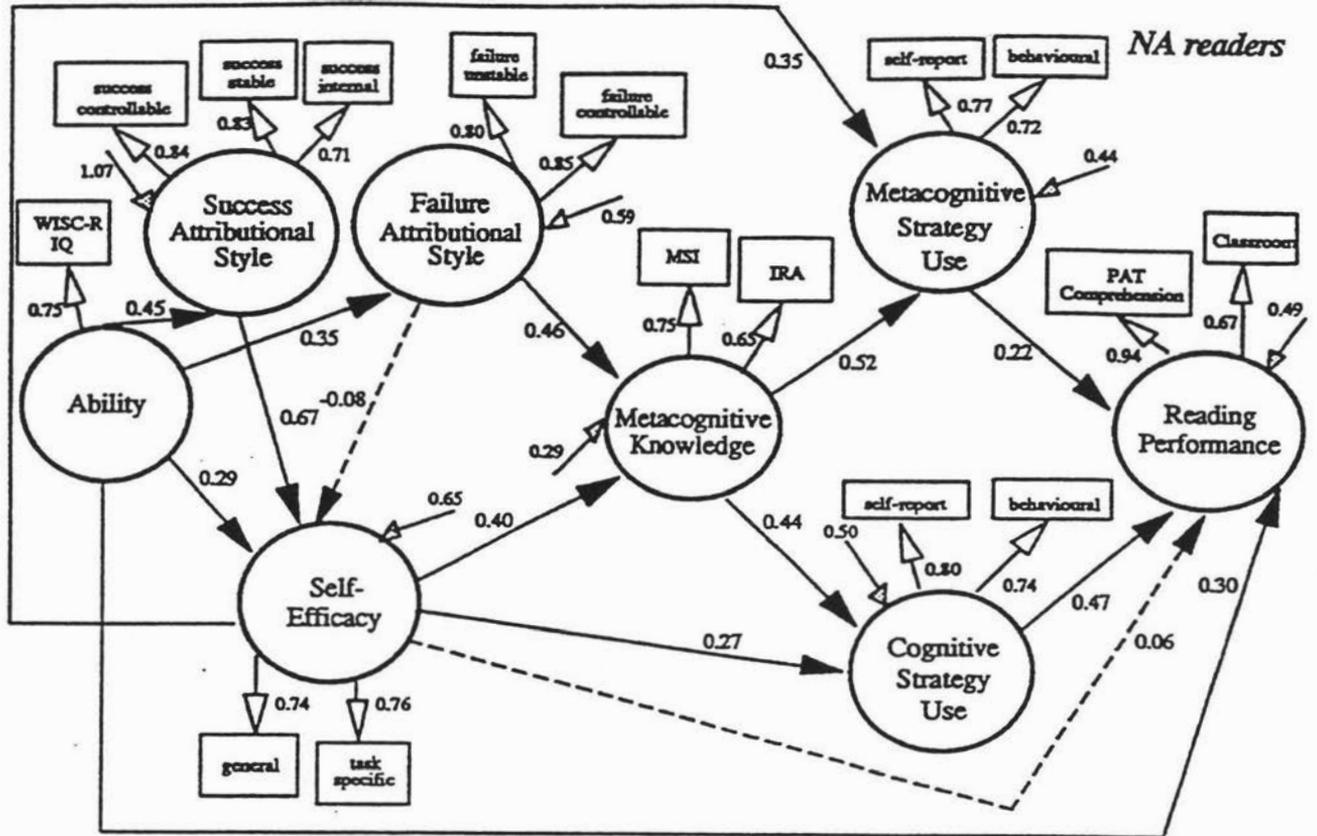


Figure 15. Findings of revised Model 4 for NA and RD readers.

Note: Failure internal attribution was dropped from the model as an indicator because it did not display desirable measurement properties. The path from ability to metacognitive knowledge was fixed to zero. To avoid clutter, measurement errors and correlated disturbances are not presented. Common metric completely standardized solution is presented for all parameters.

covariances between the structural disturbances were deleted from the figure for clarity but were estimated in the analysis. The covariances of the disturbances between the latent constructs of cognitive and metacognitive strategy use ($\psi = 0.15$, $t = 2.13$ for NA readers; $\psi = 0.20$, $t = 2.78$ for RD readers) and between success and failure attributions ($\psi = 0.62$, $t = 6.36$ for NA readers; $\psi = 0.76$, $t = 7.23$ for RD readers) were statistically significant ($p \leq .01$) suggesting that there was some common unexplained variance shared between the attributional style factors and between the strategy use factors, that did not account for all the variance between the factors and that was omitted from the model.

This model provided an extremely good fit to the data, $\chi^2(176) = 159.67$, $p = .81$; GFI = .95; NFI = .95; PNFI = .69; RMR = .050, and the factor loadings were not significantly different from those estimated for the final measurement model for all 16 observed indicators. All hypothesized loadings were highly significant ($p \leq .0001$) and thus confirmed the hypothesized factor structure. The structural relationships in Model 3 were essentially replicated in this model.

The significance of Model 4 rests on its findings that success and failure attributions have differential effects on metacognitive knowledge and self-efficacy and these effects vary across groups. Success attributional style was predictive of efficacy for NA readers; for RD readers, failure attributional style was predictive of efficacy. Another important finding was that in both groups only failure attributions significantly predicted metacognitive knowledge.

At first sight, the above findings particularly that between attributional style and self-efficacy seem contradictory to the findings of Models 1 through 3. The findings of the previous models suggested that the magnitude of the relationship between attributional style and efficacy is similar across groups. One explanation that can be offered for this inconsistency is to suggest that the joint influence of success and failure attributions to efficacy is similar across groups, but the individual effects of these subcomponents vary in magnitude across groups.

As hypothesized, ability predicted both success and failure attributions for NA readers but not for RD readers. Ability was a stronger predictor of success attributions (standardized solution = .45, $t = 4.53$) than failure attributions (standardized solution = .35, $t = 4.36$). For RD readers, as hypothesized, ability did not predict reading performance. Failure attributional

style was a better predictor of metacognitive knowledge (standardized solution = .46, $t = 5.95$) than self-efficacy.

As was seen in Model 3, metacognitive knowledge had stronger direct effects on metacognitive strategy use (standardized solution = .52, $t = 4.81$) and cognitive strategy use (standardized solution = .44, $t = 4.23$) than did efficacy. Cognitive strategy use was a significant predictor of reading performance for NA readers only. For both groups, the effect of metacognitive knowledge on performance was significantly mediated by metacognitive strategy use. The direct and positive links between efficacy and metacognitive/cognitive strategy use indicate that students who feel efficacious are more likely to engage in strategy use. The primary distinctions between the two groups were in the different weighting of the relationships between ability and success/failure attributions, success/failure attributions and efficacy, ability and performance, efficacy and performance, and cognitive strategy use and performance.

For NA readers, the squared multiple correlations were .19 for success attributions, .20 for failure attributions, .56 for efficacy, .69 for metacognitive knowledge, .63 for metacognitive strategy use, .49 for cognitive strategy use, and .62 for performance. For RD readers, the squared multiple correlations were .04 for success attributions, .01 for failure attributions, .68 for efficacy, .54 for metacognitive knowledge, .66 for metacognitive strategy use, .36 for cognitive strategy use, and .64 for performance. As in the previous models, the low R^2 value for attributions is to be expected because it is based on only one predictor (ability). In general, the variance accounted for in each structural equation was similar to that found in Model 3. Inspection of modification indices, standardized residuals and Q-Q plot, and correlations of estimates also indicated that the model was reasonably well specified for both groups.

The above results seem to demonstrate that most of the predictions made for hypothesized Model 4 were well supported. However, the results did not support the following predictions: (a) the relationships between success/failure attributions and efficacy are similar across groups (hypotheses no. 5 and 6), (b) success attributional style predicts metacognitive knowledge, (c) failure attributional style predicts efficacy for NA readers, and (d) success attributional style predicts efficacy for RD readers. It can be concluded that success/failure attributions have differential effects on efficacy and metacognitive knowledge. Model 4 demonstrates the importance of separating out the effects of the subcomponents of attributional style on the outcome measures. For instance, although the joint influence of success and failure attributional style on efficacy may be equal across groups, as was demonstrated in Models 1

to 3, examination of the separate effects of the components reveals that their effects on efficacy may vary in magnitude.

9.3.2.5 Summary of findings

A number of interesting and important results emerged from the analyses of the four hypothesized models. The results revealed that although the measurement models were not identical across groups (as indicated by their residuals), the construct pattern and loadings were similar. The relationships amongst ability, success/failure attributions, efficacy, metacognitive knowledge, cognitive/metacognitive strategy use, and performance are essentially similar for NA and RD readers except for the paths ability to success/failure attributions, ability to performance, efficacy to performance, and cognitive strategy use to performance. In general, the models fitted slightly better for NA than for RD readers.

The causal ordering of the variables were similar across groups. Attributions predicted metacognitive knowledge directly or indirectly through self-efficacy, and combined strategy use significantly mediated the effect of metacognitive knowledge on performance. Of the components of combined strategy use, it was metacognitive strategy use that significantly predicted performance across groups. The mediating role played by cognitive strategy use in the relationship between metacognitive knowledge and performance differed across the groups. For instance, cognitive strategy use significantly predicted performance for NA readers but not for RD readers.

Of the components of attributions, it was failure attributions that significantly predicted metacognitive knowledge. Furthermore, although the joint influence of the attributional style components on efficacy was similar across groups (as indicated in Models 1 through 3), the components differentially influenced efficacy, and the individual effects varied in magnitude across groups. The findings also indicated that the groups varied in terms of which attributional style component significantly predicted efficacy. The effects of self-efficacy on metacognitive variables were similar across groups. Efficacy directly and positively predicted metacognitive knowledge, metacognitive/cognitive strategy use. However, efficacy played a more crucial role in determining performance for RD readers than for NA readers.

In conclusion, although the metacognitive-motivational system of RD students appears to be similar in many aspects to NA readers, the number of subtle but important differences between NA and RD readers suggest that these readers may possess distinct metacognitive-motivational systems. Overall, the models that were specified in this study and the hypotheses that were formulated for each model were generally well supported.

9.3.3 Supplementary analyses

9.3.3.1 Test of impact of ethnicity and gender on hypothesized models

The descriptive results indicated that the proportions of male/female and European/non-European students varied markedly across groups (see Chapter 8, section 8.2.3). Moreover, multivariate analyses of variance revealed gender and ethnicity differences on some variables. Ethnicity differences were found for stability attributions for success and efficacy. Gender differences were found for attributions, efficacy, and metacognitive knowledge.

These findings suggested that it was necessary to reanalyse the four hypothesized models (Figures 4 through 7) again with ethnicity and gender included as control variables to evaluate the robustness of the specified model. Although the addition of these variables may improve the precision of the parameter estimates in the models, these analyses were considered of secondary importance since ethnicity and gender are constructs that are not potentially manipulable or malleable to instruction.

A description of the analyses and the results is provided in detail in Appendix J. Essentially the results suggested that the introduction of ethnicity and gender to the four hypothesized models did not markedly change the structure of the models or the estimates.

9.3.3.2 Replication of results using random sampling technique

The group sample sizes used in the present study were considered to be of modest size considering the complexity of the models tested. As a number of a posteriori respecifications were made to the hypothesized models, the results may have been obtained by capitalizing on chance. Thus, due to the model modifications made in this study, it was considered important to cross-validate the results on independent data. However, an independent sample was unavailable. Furthermore, due to the modest sample size of the present study, cross validation could not be undertaken by splitting the sample into two subsamples and using one subsample as the validation sample.

One approach for validating the results obtained is to use an 80% random sampling technique for each group. Such an approach has been used by Earley and Lituchy (1991). Hence, two additional analyses were conducted, each time using an 80% random sampling technique, to test the robustness of fit. In both replications, the results did not differ markedly from those of the models presented earlier.

CHAPTER 10

DISCUSSION

This chapter presents a discussion of the results of the present study and of the way the findings relate to theory and the research of others. It begins with a brief summary of the purpose of the study. A brief overview of the findings is followed by a more detailed discussion of the results in the following sections. The relationships amongst the motivational, cognitive, and metacognitive variables are presented first. Discussion of these results focuses initially on relationships that are congruent across groups, particularly predictions that were hypothesized and supported by theory and/or empirical research, followed by findings that were contrary to predictions. Findings that distinguished between normal achieving (NA) and reading disabled (RD) readers are discussed next, followed by a discussion of the relationships between metacognitive components and cognitive processes. A discussion of the implications of the findings for Borkowski et al.'s (1989, 1990, 1992) model of metacognition precedes a discussion of the results of the secondary analyses. The chapter concludes with a discussion of implications for educational practice, the study's limitations and strengths, and suggestions for additional research.

10.1 Brief overview of purpose and findings of study

During the last two decades researchers from the fields of reading and educational psychology have investigated students' use of strategies as they learn and read from text. From these studies, they reached consensus that students' metacognitive knowledge and strategy use are important for understanding text, and that students' motivation is important for predicting strategy use. These studies did not, however, present a causal analysis of the way motivational processes and metacognitive knowledge of young adolescents relate to their use of strategies or of how those same variables related to their reading achievement. The study reported here was designed to examine those relationships for young adolescent students in Form 2 Intermediate (junior high) classrooms.

Furthermore, the extension of causal models of reading comprehension to RD children has rarely been attempted. In an attempt to understand and explain the differences in performance between NA and RD readers, the present study compared the interrelationships among

variables such as intellectual ability, attributional style, self-efficacy, metacognitive knowledge, and cognitive/metacognitive strategy use between these groups. Hence, in contrast to most previous research on reading disability, the current study incorporated several of the major variables that have formerly been studied in isolation. In addition, it examined whether a similar set of relationships could explain individual differences in the reading comprehension of both NA and RD students. Two secondary goals of this study were the identification of important predictors that would distinguish between NA and RD adolescents, and the validation of the multidefinitonal approach in defining RD adolescents.

The relationships among the motivational, cognitive, and metacognitive variables were evaluated using four models with the results being analyzed using LISREL structural equation modelling procedures (Jöreskog & Sörbom, 1989, 1993). Figures 12 through 15 present the summarized results of the models of best fit. As was described in Chapter 9, the basic model was developed gradually, culminating in the final model, presented in Figure 15 (see p. 227). This final model fits the data well and, although it contains some modifications, retains to a large extent the basic causal structure that was postulated for this study.

The relationships among the variables in the hypothesized models are well supported and are generally consistent with previous research and theory. The findings of the present study, in general, lend strong support to several theoretical models which are essentially consistent with each other. These include Borkowski's model of metacognition (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990), the Triple Alliance model (Short & Weissberg-Benchell, 1989), and self-regulated learning models (McCombs, 1986; Zimmerman, 1990). Basically, these models suggest that reading and learning are multicausal. In addition, they support the view that an integration of motivational, metacognitive, and cognitive variables is required for a complete understanding of the development of reading.

The metacognitive-motivational model tested in the current study has several points of overlap for NA and RD readers. Replication of the main findings of Carr et al. (1991) as hypothesized in this study was successful. In both groups of readers, attributional beliefs significantly predicted self-efficacy and metacognitive knowledge. However, when the separate effects of attributional style were examined for each outcome, the results revealed that adaptive attributions for failure was the only significant predictor of metacognitive knowledge. Furthermore, the attributional components varied in their impact on self-efficacy and these differential effects also varied across groups.

An important contribution of this study was the incorporation of strategy use to the model. When combined strategy use was included in the model, metacognitive knowledge no longer had a direct impact on reading comprehension (performance); instead combined strategy use played a significant role in mediating this relationship. Self-efficacy as well as metacognitive knowledge predicted combined strategy use, which in turn predicted reading performance. Closer examination of the components of combined strategy use revealed that only metacognitive strategy use directly predicted reading comprehension for both groups. The mediating role played by cognitive strategy use in the relationship between metacognitive knowledge and reading comprehension differed across groups. Self-efficacy directly and positively predicted metacognitive knowledge and metacognitive/cognitive strategy use in both groups.

A number of paths distinguished NA from RD readers. They were the paths between ability and success/failure attributional style, ability and performance, success/failure attributional style and efficacy, cognitive strategy use and performance, and efficacy and performance. These findings are, in general, congruent with the hypotheses for the present study. On the whole, motivational variables appeared more important in determining comprehension for RD readers whilst strategy use variables were more important for NA readers.

10.2 Relationships which were consistent across NA and RD groups

The absence of a widely accepted and multicomponential framework for understanding reading disabilities has hampered attempts to study fundamental characteristics and develop effective remedial reading programs. It is argued that one way to rectify this lack of conceptual clarity consists of examining theoretical models of reading achievement developed for other populations (such as NA populations) and identifying their potential points of overlap with RD children. The results of the present study suggest that for both NA and RD readers, general attributional style (as a single factor latent construct), self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use interrelated in a similar fashion with each other. Moreover, these variables (with the exception of efficacy and cognitive strategy use) predicted reading comprehension performance in a similar fashion across groups.

A number of important theoretical implications follow from the results found for the revised model, shown in Figure 15 (see p. 227). Foremost is the finding that the model accounted for approximately 60% of the sample variability in reading performance for each group. This provides substantial support for the hypothesized structure of the model. The proportion of explained variance in the present study is large, particularly in the realm of social science research (L. J. Horwood, personal communication, March 27, 1995). (Comparison of explained variance with that of the Carr et al. (1991) study could not be made as the authors did not provide the information.) The second theoretical implication stems from the finding that reading performance is multicausal, involving the dynamic interplay of motivational, cognitive, and metacognitive variables. Some of the relationships in the model imply a direct causal relationship between two variables, while others imply more complex, indirect causal relationships. Finally, despite the similarity of a number of relationships between motivational, cognitive, and metacognitive variables, the reading performance of NA and RD readers were explained by *distinct* metacognitive-motivational models.

10.2.1 Relationships between ability, the self-system, and metacognitive knowledge

In the current study, a direct relationship between ability and self-efficacy was specified in accordance with Bandura's (1986) theory. However, Carr et al. (1991) postulated an indirect path between ability and self-esteem (a measure of expectancy). These competing models were tested in the present study and the findings revealed that a model which specified a direct effect between ability and efficacy provided a better fit than a model which only specified an indirect effect through attributions. The direct relationship between ability and self-efficacy is consistent with past research (Earley & Lituchy, 1991).

The self-system (which includes constructs such as self-efficacy, self-esteem, and attributional beliefs) has been described as a complex interdependent system that supports both metacognitive functions and academic performance (McCombs, 1986). According to Borkowski et al. (1989, 1990, 1992), the self-system is strongly directed by attributional beliefs. The findings in the current study suggest that the development of attributional beliefs is closely tied to perceptions of self-efficacy. In addition, consistent with self-efficacy theory (Bandura, 1986), the findings suggest that attributional judgments influence performance through their intervening effects on perceived self-efficacy.

In accordance with Carr et al.'s (1991) model, the importance of self-system constructs in predicting metacognitive knowledge received considerable support from the findings of the present study. The findings suggest that both attributions and efficacy may play key roles in students' acquisition of metacognitive knowledge. Students who have adaptive attributional beliefs and who feel efficacious demonstrate more knowledge about reading strategies. The observed relationship between attributions, efficacy, and metacognitive knowledge is consistent with the findings of previous empirical studies (e.g., Chan, 1995).

The present study extends previous research by examining the relative impact of attributions and efficacy on metacognitive knowledge. As hypothesized, in comparison to efficacy, attributions had a stronger impact on metacognitive knowledge. This finding can be seen as an endorsement of Borkowski's proposition that of all the self-system constructs, attributional style represents the key catalyst in metacognitive development (Borkowski et al., 1990; Groteluschen et al., 1990). However, the results of Carr et al. (1991) suggested that self-esteem was a stronger predictor of metacognitive knowledge than attributions. Two possible explanations are suggested for their findings. First, the accuracy of their predictions may have been undermined by the use of global measures of motivational beliefs, which usually leads to a reduction in the power of prediction (Assor & Connell, 1992; Bandura, 1990). A second explanation involves Carr et al.'s questionable use of subscale scores of the IRA measure as indicators for the construct of metacognitive knowledge (McLain et al., 1991).

The results suggest that attributional judgments of success and failure are likely to affect the acquisition of metacognitive knowledge. Ascribing controllable-unstable factors such as effort and strategies for failure, and internal-controllable-stable factors for success may motivate students to experiment with strategies, thereby enhancing metacognitive knowledge. In the earlier discussion of attributions, attributional style was conceptualized as a single latent construct. Examination of the separate effects of subcomponents of attributional style yielded a more detailed account of the effects of attributions. Although it was hypothesized that both success and failure attributional styles would predict metacognitive knowledge, the results indicated failure attributional style was the only significant and critical component for both groups. This finding is contrary to what was hypothesized and may have important implications for intervention, which will be raised in section 10.7.

By way of explanation it may be helpful to consider the role of failure on learning. The finding that only failure attributions predict metacognitive knowledge suggests the importance

of failure outcomes. Attribution theory posits that it is not so much the event of failure, or even its frequency, that disrupts performance as much as the *meaning* of failure. Also, several theories of motivation and cognition have incorporated factors associated with failure such as challenge, frustration, or anxiety as being necessary for achievement (Atkinson & Birch, 1978; Borkowski et al., 1990; McCombs, 1986). Thus, an adaptive response to failure may be crucial, particularly in view of theories that suggest failure is not only unavoidable but also necessary for academic success and self-regulated learning (Clifford, 1980; Covington, 1989, 1992).

Some researchers have suggested that it is essential for students to learn that inevitable failures can become constructive experiences from which they can learn to solve problems independently (Borkowski et al., 1989; Clifford, 1984; Paris & Byrnes, 1989). Failure can be insightful by providing information about the range of applications of specific strategies and the need to modify strategies while learning (Oka & Paris, 1987). Borkowski et al. (1989) suggested that success is often the result of a series of failures. Information gathered from unsuccessful outcomes can be used to enhance success on subsequent tasks. For instance, failure fosters disequilibrium in that new strategies and new behaviour must be engendered to cope with failure (Paris & Byrnes, 1989). Thus, failure provides an opportunity to obtain feedback about the appropriateness and effectiveness of strategic efforts.

In addition, students who experience failure are more likely to be challenged to experiment with alternative strategies and engage in self-testing so as to find a strategy that works well in the current situation. A student who experiences success is not likely to continue testing alternative strategies as success on the task has already been achieved. It can therefore be argued that failure outcomes provide more opportunity to experiment with alternative strategies than do success outcomes. Such self-testing, in turn, is likely to enhance the acquisition of metacognitive knowledge.

The relative impact of attributional subcomponents on self-efficacy was also examined. It was hypothesized that both components of attributional style would predict efficacy. This prediction was not supported; rather the findings indicated that success attributions significantly predicted efficacy for NA readers and failure attributions significantly predicted efficacy for RD readers. The explanation for this finding will be discussed in section 10.3.1. It appears that the joint influence of success and failure attributions on efficacy may be similar across groups, but the separate effects of the subcomponents varies across groups.

It was predicted that the relationship between ability and metacognitive knowledge would be significant for NA readers but insignificant for RD readers. The insignificant and weak relationship between ability and metacognitive knowledge for NA readers in the present study was contrary to predictions. The absence of this direct effect for both groups does not demonstrate a total lack of relatedness between these two variables, as the relationship is significantly mediated by attributions and efficacy. It is likely that ability may only have an indirect effect on metacognitive knowledge via motivational beliefs which may in turn act as precursors of metacognitive knowledge. The finding of the present study stands in direct contrast to Carr et al. (1991), who found that ability directly predicted metacognitive knowledge for both groups. The variation in findings may be due to differences in how the underachieving group was defined, or the use of different metacognitive knowledge measures.

10.2.2 The effects of self-system constructs and metacognitive knowledge on strategy use and performance

Previous researchers have shown that students' motivational processes have important influences on their use of learning strategies across a range of content areas, including science, English, and mathematics (Meece et al., 1988; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990). The present study indicates that motivational processes also predict metacognitive/cognitive strategy use in the domain of reading.

In the present study, the motivational components of self-efficacy and attributional style were linked in important ways to combined strategy use and reading performance. The present findings support a cognitive theory of learning that includes motivation or motivation related constructs as important mediators of strategy use (Paris et al., 1983, 1991; Pokay & Blumenfeld, 1990). Self-efficacy was positively related to strategy deployment for both NA and RD readers. Students who felt efficacious were more likely to report greater use of cognitive strategies and metacognitive strategies. Self-efficacy may have furnished the motivation to use cognitive and metacognitive strategies that, in turn, triggered specific learning responses. The direct relationship between efficacy and cognitive/metacognitive strategy use is consistent with previous research (Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990; Zimmerman & Martinez-Pons, 1990). However, it differs from Borkowski et al.'s (1989, 1990, 1992) formulation where it was specified that efficacy would only "energize executive processes" (metacognitive strategy use). The findings in the present study suggest that efficacy may also "energize" cognitive strategy use. Nevertheless, this finding

is still consistent with Borkowski et al.'s (1990) overall thesis that motivational variables "power" self-regulated learning.

In addition to efficacy, students must have adaptive attributional beliefs in order for the metacognitive system to work. In the present study, attributional style was strongly related to the deployment of cognitive and metacognitive strategies. Students who ascribed internal-controllable-stable factors for their reading success and controllable-unstable factors for reading failure were more strategically engaged when trying to comprehend text. Although adaptive attributional style was positively and significantly correlated with strategy use and reading performance, it did not have a significant *direct effect* on combined strategy use or on reading performance. More specifically, general attributions indirectly influenced strategy use and performance through their impact on metacognitive knowledge and self-efficacy.

With regard to predictors of metacognitive/cognitive strategy use, the findings revealed that students' metacognitive knowledge and perceived self-efficacy had strong direct effects on strategy use. As Palincsar and Ransom (1988) point out "motivation or will to learn is an important issue in that effective strategy use necessitates both will and skill" (p. 785). Another notable finding that emerged was that both metacognitive knowledge and efficacy had a larger impact on metacognitive strategy use than on cognitive strategy use.

It has been suggested that self-efficacy may have positive effects on dynamic aspects of metacognition (e.g., metacognitive strategy use), rather than on static measures of metacognitive knowledge (Pintrich et al., 1994). The finding that self-efficacy has a significant impact on metacognitive knowledge and metacognitive strategy use suggests that self-efficacy is an important influence on both dynamic and static measures of metacognition.

A notable contribution of the present study was the examination of the mediating influence of metacognitive knowledge and strategy use on the relationship between motivational beliefs and performance. These relationships have only recently been investigated. Using path-analytic procedures, Chan (1995) investigated the relationships among attributional beliefs, self-perceptions of competence, metacognitive knowledge, and reported use of cognitive strategies with 5th, 7th, and 9th grade NA children. Two attributional belief variables were used in her model: a belief in personal control and a learned helplessness variable. In contrast to the younger grades, where attributions were found to predict achievement, the influence of attributions and perceived competence on reading achievement for 9th grade children was

mediated through the use of cognitive strategies. For these children the relationship between metacognitive knowledge and reading achievement was mediated by cognitive strategy use. In addition, for 7th and 9th grade children, metacognitive knowledge mediated the effects of attributions on cognitive strategy use. Perceived competence predicted cognitive strategy use across all grades. Cognitive strategy use directly influenced reading achievement only for 9th grade children. In general, Chan's findings (particularly those of 9th grade children) were consistent with those of the present study, with the exception of the direct effect of attributions on cognitive strategy use.

In the present study, the results confirm the hypothesis that metacognitive knowledge predicts reading comprehension; but when combined strategy use was added to the model, this mediated the relationship between metacognitive knowledge and performance. Thus, when the two components of metacognition were included in the model, the metacognitive strategy use component was directly tied to performance for both NA and RD readers. This finding corroborates the assertion by numerous investigators that self-regulated learning depends on the extent to which learners are able to make judicious choices among learning strategies (Nolen, 1988; Paris & Newman, 1990; Shuell, 1988).

For both NA and RD readers, combined strategy use is significantly predictive of reading comprehension; however, the magnitude of the relationship differs across groups. Examination of the separate effects of the components suggest that only metacognitive strategy use is significantly and similarly related to reading performance. The impact of cognitive strategy use on reading performance varies across groups. These findings demonstrate that it may be useful to examine the effects of metacognitive and cognitive strategy use separately, particularly with regard to providing instruction (cf. Meece et al., 1988).

In accord with prior research and a general model of self-regulated learning, students who reported engaging in higher levels of strategic behaviour had higher metacognitive knowledge and comprehension scores. For learners to be able to use and control their metacognitive knowledge, they need to develop metacognitive skills (Brown, 1980). These findings, in large part, support the current view that comprehension is a highly strategic activity, requiring the coordination of a variety of metacognitive skills.

Two competing theories were tested with regard to the relationship between metacognitive knowledge and cognitive strategy use. In accord with the "inefficient learner view" (Swanson, 1989), it was hypothesized that metacognitive knowledge would predict cognitive strategy use for RD readers. Several other researchers assume the "inactive learner view" where RD students are considered to be passive or strategically inactive learners (Torgesen, 1979, 1980, 1982b; Torgesen et al., 1979; Wong, 1980). This latter hypothesis was tested as a competing model. Under this hypothesis, metacognitive knowledge would not predict cognitive strategy use in RD readers. The difference in chi-square indicated a significant deterioration in fit, resulting in the model being rejected. Thus, the present study provides indirect support for the inefficient learner hypothesis. This finding is consistent with the views of a number of researchers who suggest that RD students' failure to use cognitive strategies effectively during reading is strongly influenced by their limited and less sophisticated metacognitive knowledge (Paris & Oka, 1989; Winograd, 1984). Thus, metacognitive knowledge is an essential prerequisite if students are to become competent strategy users (Baker & Brown, 1984a; Borkowski et al., 1989).

Although the testing of the inactive and inefficient learner views represented a minor part of the present study, it has some important implications for revitalizing the metacognitive orientation to reading disabilities. The field has become so inundated with the focus on phonological skills, that it has not adequately considered the important contribution of higher order skills such as metacognition (H. L. Swanson, personal communication, September 20, 1995). Further discussion of the importance of metacognition for performance is provided in section 10.4.

10.3 Relationships that distinguish NA and RD adolescents

Several major conclusions about processes that distinguish reading achievement and impairment can be drawn from the results of the four models investigated. Although the metacognitive-motivational system of RD students appears to be similar in many aspects to proficient readers, the results indicate that there are a number of subtle but important differences between NA and RD readers. These differences are sufficient to warrant the conclusion that NA and RD readers possess *distinct* metacognitive-motivational systems. The findings indicate that the following relationships may differ across groups: ability and

success/failure attributions, ability and performance, success/failure attributions and efficacy, efficacy and performance, and cognitive strategy use and performance.

10.3.1 Relationships between ability, self-system constructs, and performance

A prominent discriminating attribute of NA and RD students is the nature of the relationship between their intellectual ability and reading performance (Keller & Hallahan, 1987; Short et al., 1984). In the present study, general ability in NA readers is significantly and directly predictive of reading achievement even when the effects of all other variables are controlled. This finding is in accord with previous research and suggests that general ability provides the basic intellectual mechanism for the reading achievement of NA readers.

In contrast, ability did not facilitate reading performance for RD readers. In other words, RD students demonstrated difficulties in reading in a manner inconsistent with their capability to grasp the general knowledge and skills measured on IQ tests (Torgesen & Licht, 1983). This finding is consistent with previous empirical research which has demonstrated that the ability-performance linkage does not hold (or is severely attenuated) for an RD individual (S. Graham & Harris, 1989; Stanovich, 1986a; Swanson, 1991b).

The findings of the present investigation replicate Carr et al.'s (1991) finding that achievers can be distinguished from RD students by the absence of a causal link between their measured ability and attributional beliefs. Carr et al. concluded that, unlike achievers, underachieving children did not credit themselves for their prior performance. Examination of the subcomponents of attributional style for the present study reveals similar results: ability does not predict either success or failure attributions for RD readers but does so for NA readers. Interestingly, ability is a stronger predictor of success attributions than failure attributions.

Although the relationship between general attributional style and self-efficacy is significant and similar across groups, examination of the separate effects of subcomponents of attributions yields a different picture. More specifically, attributions for success predicted efficacy for NA readers but attributions for failure predicted efficacy for RD readers.

One explanation that can be put forward for the differential impact of attributional style subcomponents on efficacy across groups is that NA and RD adolescents vary in taking

responsibility for their failures and successes. RD students take more responsibility for their failure than their successes (Dudley-Marling, Snider, & Tarver, 1982); that is they believe they control failure (since failure is attributed to their lack of ability) but not success. This engenders feelings of powerlessness to influence their environment, which in turn produces debilitating effects on their efficacy. In contrast, NA students take more responsibility for their successes, and this may lead to their preoccupation with the meaning and significance of their success. Their successes lead them to believe they are in control of their environment which in turn helps increase their efficacy.

Another related explanation is that frequency of achievement outcome may determine an individual's level of efficacy. For example, NA readers generally experience success more frequently than failure. These success experiences promote high achievement expectations and engender adaptive attributional beliefs for success outcomes which lead to an increase in positive feelings of self-efficacy. On the other hand, for RD students their inordinately high incidence of failure results in them becoming vulnerable to the debilitating effects of failure. Infrequent success in the face of overwhelming failures would inhibit motivation. Continued failure probably fulfils low achievement expectations, thereby further confirming their maladaptive attributional beliefs. These maladaptive attributions for failure, in turn, are likely to reduce their confidence thus decreasing self-efficacy.

10.3.2 Relationship between strategy use and performance

The positive relationship between strategic learning behaviour and academic performance is acknowledged by several researchers (Alexander & Judy, 1988; Garner, 1990; Palincsar & Brown, 1984; Paris, 1988; Paris et al., 1991; Pressley, Symons, et al., 1989; Weinstein et al., 1988). Paris et al. (1991) refer to strategic reading as "a hallmark of expertise" among readers. They conclude that strategic reading is critical to the educational and cognitive development of children for a number of reasons. First, strategies are amenable to instruction and can be controlled by students. Therefore, strategies empower readers to organize, elaborate, and evaluate text information. Second, the development of cognitive strategies in reading parallels and extends the development of cognitive strategies in other domains. Third, strategic reading mirrors motivation and metacognition, thus enabling readers to become competent and confident learners. Fourth, strategic reading involving monitoring and regulation are critical dimensions of thinking and therefore enables increased learning

throughout the curriculum. Finally, students who believe in the instrumentality of strategies for accomplishing selected tasks and who have a sense of personal causality become self-regulated learners.

One important contribution of the present study is that it extends strategy research by comparing the relationship between cognitive strategy use and performance across NA and RD readers. The findings indicated that the relationship between combined strategy use and performance, although significant for both groups, varied in magnitude. Examination of its subcomponents revealed that the relationship between metacognitive strategy use and performance was significant and similar across groups. However, the relationship between cognitive strategy use and performance differed across groups. Cognitive strategy use predicted reading performance for NA readers but did not for RD readers. The significant relationship between cognitive strategy use and performance for NA readers is consistent with the findings of Pokay and Blumenfeld (1990). However, the importance of the path between cognitive strategy use and performance in explaining reading impairment is not consistent with Borkowski's theory and will be dealt with in detail in section 10.5.2.

The nonsignificant relationship between cognitive strategy use and performance in RD readers is, however, consistent with both the "inefficient learner view" (Swanson, 1989), and the "inactive learner view" (Torgesen, 1980, 1982b; Torgesen et al., 1979; Wong, 1980). These hypotheses implicate deficiencies in the use of strategies for learning; they predict that cognitive strategy use would not predict performance for RD readers.

Further examination of the interview data revealed that 85% of the RD students as compared to 48% of the NA students verbalized inappropriate strategies when answering one or more comprehension questions. This finding is consistent with that of Swanson (1988) who found that LD children were able to "actively" develop strategic thought patterns. Swanson (1989) suggests that the strategic deficits of LD children may be related to their preferred use of heuristics and their limited ability to use strategies flexibly. In other words, LD students often choose an inappropriate strategy from their repertoire of strategies to complete the task at hand.

However, the findings also revealed that RD students were often as likely as NA students to report an appropriate strategy but were less successful in answering one or more comprehension questions. These findings suggest that the strategic deficiencies of some RD

readers may also be due to the *inefficient implementation* of the strategy. Other evidence comes from the results of an interview question, in which students were requested to retell the story they had just read. Reading disabled readers (41%) were more likely than their NA peers (15%) to include unimportant details, although they reported looking for main ideas.

There may be several explanations for RD students' inefficient implementation of strategies. First, RD students may lack the practice to allow them to master the strategic behaviours. This explanation receives support from previous research. For example, Kobasigawa, Ransom, and Holland (1980) found that even among students who were aware of text characteristics and had the ability to use a skimming technique when instructed to do so, the spontaneous employment of skimming developed only gradually. Brown and Smiley (1977) found that when nonspontaneous users were induced to use underlining and note-taking strategies while studying they did so in a more random fashion than spontaneous users.

Second, it may be due to their inability to transform strategies. It has been suggested that one mechanism that promotes expert performance is related to strategy transformation (Chi, Glaser, & Farr, 1988). Children who become experts at certain tasks have often learned simple strategies and, through practice, discover ways to modify them into more efficient and powerful procedures. In contrast, RD children may not be able to transform the strategies easily to match the task at hand as they have not acquired much practice using the strategies. They may learn the skills related to reading and may sometimes perform appropriately on a task by carefully and systematically following prescribed rules or strategies. However, empirical evidence suggests that the difference between RD and NA children is that the latter have modified such strategies to become more efficient (Swanson & Cooney, 1985). It is plausible that the RD child remains a novice in learning new information because he or she fails to transform reading strategies into more efficient forms (cf. Swanson & Rhine, 1985).

The present study was not designed to examine this aspect of strategy use in any detail. However, preliminary findings of the interview data suggest that Bray's (1985) proposition for an examination of strategic flexibility may be necessary for a more complete understanding of the relationship between cognitive strategy use and reading performance.

A third explanation for inefficient implementation of strategy use could be that their deficient monitoring of metacognitive processes hinders efficient use or application of cognitive strategies. They may lack the monitoring skills to effectively transform strategies to fit the designated task at hand. Previous empirical research confirms this. For instance, Pintrich and

De Groot (1990) findings suggest that cognitive strategy use without the concomitant use of metacognitive strategies is not conducive for academic performance. It seems that for effective performance, students require more than reading-oriented cognitive strategies; they need metacognitive strategies that oversee the use of cognitive strategies. The findings of the interview data provide some support for this explanation. For instance, even though some RD readers reported using main ideas in retelling the story, they were less successful in distinguishing main ideas from important details.

Finally, previous research has indicated that RD students' inefficient use of strategies may be the result of their disbelief in the general usefulness of strategies. Such an orientation is likely to inhibit their metacognitive and motivational development because their ineffective strategies will further exacerbate their low self-efficacy and debilitating attributional beliefs (Borkowski et al., 1990).

The present study also found that cognitive strategy use had a larger impact on reading performance than metacognitive strategy use for NA readers. While this has not been previously assessed in relation to reading, in the domain of mathematics, Pokay and Blumenfeld (1990) found that the use of specific cognitive strategies predicted initial achievement whereas metacognitive strategies predicted later achievement for NA children. They suggested that students may be able to make effective use of metacognitive strategies after becoming skilled at the use of specific cognitive strategies. They also argued that as students become more proficient at selecting and using cognitive strategies, their thinking about strategies becomes more important such that students who use metacognitive strategies are more successful.

10.3.3 Relationship between self-efficacy and performance

Although the direct relationship between efficacy and performance was not anticipated to differ across groups, the results indicated that the relationship was significant but only for RD readers. The insignificant relationship for NA readers suggests that self-efficacy is not directly tied to actual performance and plays only a facilitative role in relation to strategic engagement. However, the total effect of efficacy on performance is large and highly significant for NA readers. These findings strengthen Bandura's (1986) claim that self-efficacy beliefs are "key arbiters of human agency", and also lend support to researchers who

contend that student motivation may be better explained by efficacy beliefs than by other cognitive or affective processes (Schunk, 1989b, 1989c, 1991). The importance of the path between efficacy and performance in explaining reading impairment is not specified in Borkowski et al.'s (1989, 1990, 1992) theory. This issue will be discussed in detail under section 10.5.2.

Finding a difference in the paths between self-efficacy and performance for NA and RD readers also provides evidence that motivational characteristics may play a more significant role for RD than for NA readers. This is further corroborated by analyses which indicated that the total effect of efficacy and attributions on performance was much larger for RD readers than for NA readers. Such a finding is in contrast with the research of Paris and his colleagues who found that motivational variables increased in importance with age and skill (Paris & Oka, 1986; Saarnio, Oka, & Paris, 1990).

10.3.4 Summary

The findings in the current study support the importance of attributional style, self-efficacy, metacognitive knowledge, cognitive strategy use, and metacognitive strategy use variables in explaining reading achievement and impairment. The results of the SEM analyses suggest that, in general, the causal ordering of the motivational, cognitive, and metacognitive variables were similar across NA and RD readers. Consistent with previous research and theory, metacognitive knowledge and combined strategy use mediated the relationship between motivational beliefs and reading performance. Although the self-system constructs of attributional style and self-efficacy provided the necessary motivational states in promoting students' reading performance, it was the strategy use components (especially metacognitive strategy use) that provided the means to reach that goal. Consistent with the theories of Bandura (1986) and Weiner (1979, 1985a), and the findings of Schunk (1983, 1984) the data also suggested that the direct causal effect of attributional style was not extended onward to strategies and performance.

The relationships between general attributional style (as a single latent construct), self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use were consistent across groups suggesting that the metacognitive systems of both NA and RD readers are fundamentally similar. However, distinctive patterns of motivation, cognitive, and

metacognitive skills were also found to underlie the reading attainments of NA and RD readers. The metacognitive-motivational systems differed between groups in the way ability and attributions and in the way efficacy and cognitive strategy use related to performance. In addition, the metacognitive systems varied across groups in the way the subcomponents of attributional style related to efficacy.

Although the performance of NA and RD readers was facilitated by the mediating effects of attributions, self-efficacy, metacognitive knowledge, and metacognitive strategy use, the performance of NA readers was also mediated by cognitive strategy use, and the association between ability and attributions. In addition, the performance of RD readers was further mediated by the direct effect of efficacy on performance. The importance of efficacy and cognitive strategy use in explaining differences between NA and RD readers is not supported by Borkowski's theory (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990) and these differences will be discussed in section 10.5.2.

The present study suggests that how students account for their successes and failures determines to a large extent their perceptions of themselves as learners and their attempts to change their metacognitive knowledge and skills. This finding is consistent with Carr et al.'s (1991) suggestion that the effectiveness of the entire metacognitive system depends on whether attributional beliefs develop in collaboration with reading comprehension achievement. The connection between ability and attributions fosters the development of adaptive attributional beliefs, self-efficacy, metacognitive knowledge, and cognitive and metacognitive strategy use. In the present study, the disassociation between ability and attributional beliefs may have altered the effectiveness of the entire metacognitive-motivational model for RD readers. The adoption of maladaptive attributional style patterns may have hindered the development of their self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use, as indicated by their significantly lower mean scores. Therefore, attributions seem to be the driving force behind the self- and metacognitive systems.

Strategy use variables may be more important in predicting performance for NA readers than for RD readers. This was demonstrated by the relatively larger increase in the variance of reading performance for NA readers (13%) than for RD readers (2%) when combined strategy use was incorporated in the model. The findings in the present study parallel the work of Chan (1995). She found that, although motivation variables had a more important role in

explaining reading achievement in the lower school grades, the role of strategic learning variables was as or more important than the motivation variables in higher grades.

The findings of the present study support the importance of positive motivational beliefs in the development of the metacognitive-motivational system. These findings support the growing body of evidence suggesting that a well functioning self-system (e.g., self-efficacy and adaptive attributions) is critical in motivating students to self-regulate their learning and reading (Ames & Archer, 1988; McCombs, 1986, 1989; Nolen, 1988; Pokay & Blumenfeld, 1990).

In summary, this study provides empirical evidence for the way students' attributions, self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use work together in their influence on performance. The results demonstrate that metacognitive processes and motivational beliefs are intimately related, and in combination, explain the emergence of strategies which in turn generate successful performance. Overall, the findings of the present study are theoretically consistent with self-regulated learning theories (McCombs, 1986), the Triple Alliance Model (Short & Weissberg-Benchell, 1989), and Borkowski et al.'s (1989, 1990, 1992) theory.

10.4 Distinction between metacognitive components and their relationship to cognitive processes

Despite the growing body of research documenting the importance of metacognition in cognitive theories, there continue to be problems concerning the definition of metacognition and the differentiation of metacognitive and cognitive strategies. In fact, there is little research that has examined the discriminant validity of metacognitive knowledge, metacognitive strategy use, and cognitive strategy use. Some researchers have argued that there is no distinction between metacognitive and cognitive strategy use (e.g., S. G. Paris, personal communication, April 20, 1994). This study provides evidence that metacognitive knowledge, metacognitive strategy use, and cognitive strategy use represent distinct components. The correlations between metacognitive knowledge and metacognitive strategy use (range = .72 - .79), metacognitive knowledge and cognitive strategy use (range = .67 - .66), and metacognitive strategy use and cognitive strategy use (range = .67 - .71) suggest they are strongly related, but they are not so large as to suggest they are measuring exactly the same construct. That these constructs can be distinguished is important for research

purposes; it also may be helpful when devising intervention programs, as they appear to play unique roles in predicting performance.

Kurtz (1991) maintained that utility of the construct of metacognition could be established if its relationship to cognitive processes is demonstrated. The importance of metacognition to cognitive processes such as reading comprehension has been undermined by the inconsistent empirical findings. The findings of the present study provide additional data on the relationship between metacognition and reading and indicate a strong association between metacognition and reading performance. The present study also extends previous research by examining the relationship between the distinct components of metacognition, namely metacognitive knowledge and metacognitive strategy use, to reading performance. The results suggest that these components have unique roles in predicting performance.

The approach used to examine the relationship between metacognition and comprehension differed from previous research in a variety of ways. First, metacognitive knowledge was specifically defined. As Jacobs and Paris (1987) noted, a precise definition is not always found in studies of metacognition, and this lack of precision may be one reason why its relationship with reading performance has been difficult to determine. Second, to optimize the validity of students' self-reports, the collection of self-report data conformed to the recommendations set forth by Assor and Connell (1992) (see Chapter 8, section 8.3). As a consequence, the validity and reliability of the data collected in the present may be improved over that of other studies, and consequently the associations among the measures are likely to be more accurate.

10.5 Implications of findings for Borkowski's model of metacognition

To a large extent the findings of the present study supports the metacognition model proposed by Borkowski et al. (1989, 1990, 1992). The present study extends Borkowski et al.'s metacognition model by including attributions and self-efficacy as distinct and antecedent components to metacognitive knowledge about reading. In their model, Borkowski and colleagues included motivational belief components as general strategy knowledge correlates. However, they did not delineate how these components interrelated with each other or how they were related together in the prediction of their metacognitive counterparts.

As mentioned in Chapter 5, there are several versions of Borkowski's model (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). These versions are basically consistent with each other with minor refinements (M. Carr, personal communication, February 13, 1995). However, the most recent version presented by Borkowski and Muthukrishna (1992) proved to be an exception, in that it had undergone the most revision (J. G. Borkowski, personal communication, January 31, 1995). First, unlike earlier versions, Borkowski and Muthukrishna (1992) deleted general strategy knowledge and relational knowledge components from the model. This is not surprising considering that there is very little, if any, research on these components. Second, attributional style was included as an independent variable in the model rather than as a general strategy knowledge correlate. Borkowski and Muthukrishna specified relationships between attributions and metacognitive variables, such that attributional beliefs predict executive processes (i.e., metacognitive strategy use). The hypothesized models tested in this study are based on versions preceding the Borkowski and Muthukrishna model.

In this section, the differences in the causal ordering of the variables between those hypothesized in the present study and those posited by Borkowski and Muthukrishna (1992) will be discussed. In addition, relationships that distinguished NA and RD readers in the present study which were contrary to Borkowski et al.'s (1989, 1990, 1992) propositions are also discussed.

10.5.1 Implications for causal ordering of variables

Some of the hypothesized relationships of the present study differ from those offered by Borkowski (Borkowski et al., 1989, 1990, 1992; Borkowski & Muthukrishna, 1992). Before discussing the most recent version of Borkowski's model (Borkowski & Muthukrishna, 1992), some of the proposed relationships of Borkowski's (Borkowski et al., 1989, 1990, 1992) earlier models will be discussed first with reference to the present study. Specifically, the direct relationship between metacognitive and cognitive strategy use is addressed. Although this path was not specified explicitly in the hypothesized models for both groups, it was tested as a competing model. The results of the competing model revealed that the metacognitive strategy use to cognitive strategy use path was significant, and that the model provided a good fit to the data. However, closer examination of the data revealed that a number of paths specified in the model were rendered insignificant contradicting both theory and empirical

research. This model was therefore considered unacceptable (see Appendix H, section H.3, Model 3.10).

Notwithstanding these results, the prediction of cognitive strategy use by metacognitive strategy use is theoretically plausible. As such, there may be methodological factors which clarify these conflicting results. One plausible explanation relates to the operational definition for metacognitive strategy use in the present study. Although similar, the present study's definition is not identical to that of Borkowski (J. G. Borkowski, personal communication, January 31, 1995). The executive processes (or metacognitive strategy use) as defined by Borkowski, include analysing task demands, strategy selection, accommodating the strategy to a new task, monitoring the effectiveness of strategy use, and judging when a problem has been adequately solved (Borkowski & Turner, 1990). This definition involves a focus on task analysis and strategy selection (J. G. Borkowski, personal communication, January 31, 1995), rather than planning and monitoring as is the case in the current study.

In the most recent formulation of the metacognition model (Borkowski & Muthukrishna, 1992), the specification of attributions and metacognitive strategy use to other components in the model varied in a number of ways from earlier versions of the model, and to models specified in the present investigation. In the latest version of the metacognition model, it was posited that (a) attributions predict metacognitive strategy use, and (b) metacognitive strategy use predicts metacognitive knowledge which in turn predicts cognitive strategy use. Earlier versions, however, had specified that metacognitive strategy use predicts cognitive strategy use, and that metacognitive knowledge has reciprocal relationships with metacognitive and cognitive strategy use.

The model specified in the current study differed from that specified by Borkowski and Muthukrishna (1992) in two major ways. In the present study, attributions predicted metacognitive knowledge instead of metacognitive strategy use. In addition, metacognitive knowledge predicted metacognitive strategy use rather than the converse.

That attributions may have a direct impact on metacognitive strategy use as theorized by Borkowski and Muthukrishna (1992) was tested as a competing model in the present study. Although this model provided a good fit to the data, a negatively valenced coefficient resulted (see Appendix H, section H.3, Model 3.9). This result was contradictory to theory and previous empirical findings and was therefore rejected.

The direct influence of metacognitive strategy use on metacognitive knowledge (or specific strategic knowledge) as posited by Borkowski and Muthukrishna (1992) was not specified in the present study. There are several reasons why the aforementioned relationship might not be as theorized. It can be argued that metacognitive knowledge must develop before metacognitive strategy use can develop. In fact, Borkowski and Muthukrishna proposed this themselves (see Borkowski & Muthukrishna, 1992, p. 484).

Borkowski and Kurtz (1987) have suggested that "an important internal dependency exists in the model", whereby each component relies on the quality and development of earlier components (p. 127). For instance, metacognitive strategy use, which involves on-line regulation and selection of cognitive strategies, can only become operative after detailed representations in metacognitive knowledge (and even general strategy knowledge) components have accumulated (see Chapter 5, section 5.2.2, point 4). In other words, before reflection about strategies and self-experimentation with strategies can occur, the student must have some knowledge of these strategies. It therefore seems reasonable to assume that metacognitive knowledge is a precursor to metacognitive strategy use. Consequently, it seems more plausible to suggest that metacognitive knowledge predicts metacognitive strategy use rather than the converse. Alternatively, it may be useful to suggest that there is a reciprocal relationship between metacognitive knowledge and metacognitive strategy use as was indicated in the earlier models (Borkowski et al., 1989, 1990, 1992).

In conclusion, the argument that metacognitive strategy use can only become operative after detailed representations in metacognitive knowledge have accumulated suggests that the relationships between cognitive and metacognitive variables may be better represented in Borkowski's (Borkowski et al., 1989, 1990, 1992) earlier models than in his latest model (Borkowski & Muthukrishna, 1992). One criticism which could be levelled at the most recent development in Borkowski's model is that this refinement may have proceeded too quickly and prior to sufficient empirical testing to inform theory development.

Despite these concerns, Borkowski's (Borkowski & Muthukrishna, 1992) latest refinement of his model (which includes attributional beliefs in the model rather than as a general strategy knowledge correlate) is commendable. Although the model appears to emphasize the integration of self-system constructs, it does not explain how different self-system constructs interact together in predicting their metacognitive counterparts. In addition, the relationships between attributional beliefs and the metacognitive components would benefit from causal

empirical research. The present study is unique in that it provides empirical evidence with causal implications for Borkowski et al.'s (1989, 1990, 1992) model. Moreover, it specifies how self-system constructs of attributions and self-efficacy relate to each other and to their metacognitive and cognitive counterparts. In general, the findings provide support for Borkowski's earlier models.

10.5.2 Implications for relationships that distinguish between groups

Borkowski suggested that three main components distinguished NA and RD readers: attributional beliefs, metacognitive knowledge, and metacognitive strategy use (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). Carr et al. (1991) extended Borkowski's findings by showing that the association between ability and attributional beliefs differed across achieving and underachieving readers. The findings in the current study support the findings of Borkowski et al. (1989, 1990, 1992) and Carr et al., but implicated two other variables, namely cognitive strategy use and self-efficacy, in explaining reading achievement and impairment.

The discriminant analysis results and the SEM finding of a difference in the path between cognitive strategy use and performance across groups support the significant role cognitive strategy use plays in explaining differences between NA and RD readers. This finding is further substantiated by the empirical findings in the strategy use literature that suggest that NA students employ cognitive strategies to enhance their performance (e.g., Chan, 1995; Pokay & Blumenfeld, 1990), whilst RD students do not (e.g., Scruggs & Mastropieri, 1986).

However, there may be methodological factors that may explain the difference in the path between cognitive strategy use and performance across groups. For instance, the difference between groups may be a function of the way cognitive strategy use was assessed and measured in the present study. As noted previously, cognitive strategy use was measured in terms of self-reported strategy use or frequency. This may not be an adequate measure especially in light of findings in the behavioural data which suggested that RD students may employ cognitive strategies similar to NA students, but implement them inefficiently.

In the present study, during the interviews, it was observed that some RD readers were more likely to answer the comprehension questions incorrectly despite their report of a cognitive

strategy that was appropriate for the task. In other words, although RD students may self-report appropriate strategy use, they may not use the strategy effectively. Thus, a frequency measure of strategy use may not explain the relationship between cognitive strategy use and performance accurately, because strategies can be used with differing degrees of effectiveness (R. E. Reynolds, Wade, Trathen, & Lapan, 1989). This discrepancy between reporting cognitive strategy use and using it effectively was highlighted by R. E. Reynolds et al. (1989) who stated that "There seems to be an implicit assumption that if students say they are using a strategy or exhibit behaviours that are seen as reflecting a strategy, then they must be using the strategy effectively" (p. 175).

Structural equation modelling analyses suggested that although the joint influence of success and failure attributions on efficacy was similar across groups, the attributional components varied in their impact on efficacy across groups. The relationship between self-efficacy and performance was also found to differ across groups. Hence, in addition to attributional beliefs, efficacy is also an important variable in explaining reading performance and impairment. The results of the discriminant analysis also indicated that self-efficacy is a key discriminating variable in distinguishing RD and NA students. Consistent with previous research, the findings of the present study suggest that RD students have lower self-efficacy than NA students. The importance of the path between self-efficacy and performance in differentiating NA and RD readers was not specified in Borkowski's theory (Borkowski et al., 1989, 1990, 1992; Borkowski & Turner, 1990; Groteluschen et al., 1990). However, it is important to note that the specific effects of efficacy on other components of the model were not delineated by Borkowski et al. (1989, 1990, 1992) since they did not include efficacy as a separate variable in their model.

The differential importance of motivational and cognitive variables in predicting performance across groups is consistent with a developmental or maturational lag hypothesis (Houck, 1984). Motivational beliefs accounted for more variance in the reading performance of RD than NA readers. In contrast, cognitive and metacognitive strategy use variables accounted for more explained variance for reading performance in NA than RD readers. These findings are consistent with findings of previous research which suggested that metacognitive and cognitive skills become increasingly influential as the child matures and develops academically (Chan, 1995; Paris & Oka, 1986).

In conclusion, the findings of the present study concurred with Borkowski et al.'s (1989, 1990, 1992) prediction that attributions, metacognitive knowledge, and metacognitive strategy use distinguish NA from RD readers. However, the findings of the present study also suggested that two other components may be important in distinguishing between NA and RD readers: cognitive strategy use and self-efficacy. Although Borkowski, in later versions of his model, has begun to focus on the different self-system constructs in predicting their metacognitive counterparts, he does not explicate how these self-system constructs relate together in predicting performance. In addition, it is important to note that the present study goes further by examining whether the *relationships between components* are similar across groups. Further research using SEM procedures is required before any definitive statements can be made regarding the role of these variables in distinguishing NA and RD readers.

10.6 Findings of secondary goals of present study

A stepwise discriminant analysis was conducted to provide validation for the multidefinitonal approach used in identifying RD adolescents and also to identify important predictors that would differentiate between NA and RD readers. The high classification rates (93.5% for NA readers and 97.0% for RD readers) provide validation and support for the utility of the multidefinitonal approach used to define RD adolescents. In employing such an approach, some of the major conceptual and measurement problems associated with the commonly used regression approach in defining reading disabilities were avoided. The multifaceted approach used in the present study shows great promise in defining reading disability and has recently received strong recommendation by researchers (e.g., Shaw et al., 1995). It is anticipated that the strong empirical findings supporting this approach will stimulate renewed discussion as well as the development of alternative criteria for inclusion in the multifaceted definition.

Another goal of the present study was to identify predictors that would discriminate between NA and RD readers. The results of the current study provide additional data on the motivational and metacognitive characteristics of RD and NA readers, given the conflicting findings noted to date. For instance, the literature has revealed inconsistencies in results when comparing the motivational beliefs of RD and NA students. Although the majority of studies indicate that RD students in general have debilitating attributional beliefs and low self-efficacy, some research suggests otherwise (Pintrich et al., 1994; H. Rogers & Saklofske,

1985). The present study concurs with research findings showing that RD readers have maladaptive attributions and lower self-efficacy.

Researchers also disagree on the extent of metacognitive knowledge differences between NA and RD readers. Some researchers have found few metacognitive knowledge differences between NA and RD readers (e.g., van Kraayenoord, 1986). The RD students in the current study exhibited less metacognitive knowledge, which is consistent with most previous research. Also, in comparison to NA readers, RD readers reported less use of cognitive and metacognitive strategies. Students who were more strategically engaged in trying to read had more metacognitive knowledge about reading and performed better than students who reported less strategic engagement.

The discriminant analysis findings of the present study replicate earlier findings regarding factors that distinguish NA from RD students. For example, Oka and Paris (1987) found RD children were less strategic, had less positive attitudes and self-perceptions towards reading, and an extrinsic orientation when compared with NA children. In the present study motivational variables provided the strongest discrimination between NA and RD groups. Metacognitive knowledge and cognitive strategy use also significantly discriminated between the groups, but they were less predictive power than the motivational variables.

In summary, both the discriminant and structural equation modelling analyses suggest that motivational beliefs, metacognitive knowledge, and cognitive/metacognitive strategy use are important predictors of reading comprehension and that they are key discriminating variables in distinguishing NA from RD readers. Hence, the findings suggest that metacognitive and motivational variables combine effectively to distinguish between NA and RD readers. In addition, RD adolescents appear to have functioning metacognitive-motivational systems but the components of the metacognitive-motivational system may be weak or dysfunctional in nature. This retardation appears to take the form of poorer performance on all variables measured in the current study. The results also provide validation for the multidefinitonal approach used in this study for defining RD students.

10.7 Educational implications

The present study provides several implications for educational practice. The motivational,

cognitive, and metacognitive difficulties of RD readers lead to the question of how best to assist students who have reading difficulties. The provision of needs-based special education programmes integrated within general education, as proposed in the Draft review of special education (Department of Education, 1987), are suggested to overcome many of the problems facing RD students. However, the general lack of remedial programmes for students with reading disability in New Zealand schools suggests that RD students' potential will remain largely unrealized (Chapman & Wilkinson, 1988). It is beyond the scope of this thesis to elaborate on complex issues of remedial theory and practice, so the reader is referred to Borkowski and Pressley, whose research may offer some guidelines on interventions that can help maximise motivational beliefs and strategy use (Borkowski et al., 1990; 1992; Borkowski & Muthukrishna, 1992; Carr & Borkowski, 1989; Pressley, Goodchild, et al., 1989; Pressley, Symons, et al., 1989).

The findings of the current study are consistent with Borkowski's research (Borkowski et al., 1992; Borkowski & Muthukrishna, 1992; Groteluschen et al., 1990) which suggest that the relationships between motivational beliefs, metacognitive knowledge, cognitive strategy use, metacognitive strategy use, and achievement are interdependent. Consequently, RD students' lack of reading success seems to be based, at least in part, on skills that are modifiable or amenable to instruction. These would therefore be important components to consider incorporating in training programs as discussed below.

First, helping RD students to become self-reliant and efficient cognitive processors of information is clearly an important educational goal. This goal is based on the findings of the current study and previous research (Borkowski et al., 1989; Swanson, 1990) which suggest that cognitive processes used by RD students do not appear to exhaust or even tap their intellectual capability. Even though cognitive strategies did not directly predict reading performance for RD readers in the present study, the findings for NA readers suggest that cognitive strategies are important for students' reading achievement. Children with reading disabilities are often viewed as inefficient learners who have poor or few strategies for reading. This deficiency leads to an inability to reach their academic potential (Swanson, 1989). The finding that RD students may be inefficiently implementing their cognitive strategies, as revealed in the interview data, suggests that extended practice on cognitive strategies may be required for maintenance and generalization (Symons, Snyder, Cariglia-Bull, & Pressley, 1989). In providing students with extended practice on a wide variety of

carefully monitored ecologically valid tasks, teachers may be able to give corrective feedback, while allowing students ample time to master a strategy to the point of automaticity (Pressley, Symons, et al., 1989). This practice will provide students with first hand experience as to the benefits, limitations, and adaptability of the strategies they have been taught on tasks they meet everyday in the classroom (Valencia & Pearson, 1987).

However, the important role played by metacognitive strategy use in predicting reading performance for both NA and RD readers suggest that teaching cognitive strategies may be insufficient. Students also need metacognitive skills to enable them to use strategies automatically and efficiently, coordinate several strategies needed to achieve a task, and adapt strategies to meet differing task demands. Consistent with empirical research, the present study findings suggests that RD students have problems choosing and using metacognitive strategies to monitor and regulate their use of cognitive strategies. Thus, metacognitive strategy training may also be required. The use of metacognitive strategies is a learned skill which may need to become habitual to be effective. It is possible that, over time, students utilizing metacognitive strategies more frequently will be able to integrate these strategies more efficiently to improve reading comprehension.

The above discussion suggests that teachers might help by providing students with two types of strategy training: cognitive strategy training which focuses on task-specific reading strategies, and metacognitive strategy training which involves instruction on techniques that involve monitoring, planning, and regulation. There are many existing programmes which target these areas and research shows that strategy training programmes improve students' metacognitive knowledge, strategic engagement, and performance (see Weinstein, Goetz, & Alexander, 1988).

With independent practice at applying strategies in a variety of situations students may achieve a level of success they had been unable to achieve previously. Realising success is within their control may encourage students to expend more effort in using strategies on other tasks, and in turn increase the likelihood that they will emphasize strategically based effort over ability when evaluating their own performances. These attributions would in turn lead to an increase in effort and strategic behaviour on future tasks.

Second, the current study also indicates that students' motivation is important for students' strategy use and classroom achievement. As teachers attempt to implement strategy

instruction and foster strategy use in their classrooms, it is important for them to remember that students' motivational processes influence cognitive and metacognitive strategy use. As Borkowski, Weyhing, et al. (1988) pointed out, strategy training "in isolation from motivational histories, is an ineffective method of instruction for students with lengthy records of poor self-esteem and negative attributional beliefs about the importance of personal control" (p. 51). Hence, another general goal in remediation would be the restoration of positive motivational beliefs. Learning and reading tasks could be structured to ensure some degree of genuine success. Providing tasks which are too easy will seldom allow students credible opportunities for demonstrating self-efficacy.

The findings of the present study revealed that efficacy may be potentially important in the development of cognitive and metacognitive strategies. Self-efficacy appears instrumental in the deployment of both cognitive and metacognitive strategies, which in turn affects the quality of reading performance. Priority could be given to improving reading self-efficacy in conjunction with strategy training. Research by Bandura (1986) and Schunk (1990) suggest that efficacy is a malleable construct that can be improved with coaching and modelling. Coaching generally improves efficacy by providing motivational support and increasing persistence (Bandura, 1986), while modelling by peers and teachers also improves the on-line use and regulation of strategies (Zimmerman, 1990).

The importance of teaching attributional beliefs was also highlighted by the results, particularly in light of its crucial role in the development of the entire metacognitive-motivational system. Of the subcomponents of attributional style, the findings in the current study suggest that failure attributions may be more predictive of metacognitive knowledge than success attributions. Previous research has shown that RD children who were provided with attribution retraining showed improvement in their metacognitive knowledge and performance. The findings of the present study refine this notion by suggesting that attribution training may need to focus on failure attributions particularly for failure prone RD children who have maladaptive responses to failure.

The research findings of this study also question the appropriateness of proposals that call for teachers to arrange classroom learning so that students experience either no failure or as little failure as possible (e.g., Glasser, 1969). This view assumes that failure per se causes loss of confidence and self-esteem. A contrary view would suggest that failure can act as a positive force so long as it is properly interpreted by the learner. Instead of focusing on failure as

destructive, educators could illustrate that falling short of one's goals (an inevitable experience), can be interpreted in ways that promote the will to persist. For example, recent research has demonstrated that attributing explanations for failure to faulty strategies is important for future motivation (Weinstein, Hagen, & Meyer, 1991). In her theory of constructive failure, Clifford (1984) suggests that failure is more likely to produce constructive effects if students feel a sense of personal causality. She contends that attributions for failure to poor strategies should be encouraged for two reasons. First, strategy explanations can turn failure outcomes into problem solving situations. Second, instead of focusing on negative implications of failure, students would concentrate on finding more effective strategies. Therefore, it may be important for teachers to train students to discriminate between strategy, effort, and ability attributions in order to enhance future motivation.

Borkowski et al. (1989) have stressed the need for training programs to address RD children's beliefs about the implications of failure. Hence, it may be important for teachers to demonstrate to students that failures in reading can become constructive experiences from which independent problem solving can be learned. This is reinforced by Paris and Byrnes (1989) who stated, "indeed, failure without insight can be defeating in the same way that success without understanding does not promote future learning. Chronic success and mindless failure are to be avoided in the classroom because neither provides the opportunities for students to develop effective coping strategies" (pp. 193-194).

A promising approach may involve teaching students to redefine failure as a common and informative experience that induces action based behaviour. In viewing success and failure as steps in learning rather than terminal outcomes, students would be motivated to identify alternative ways to achieve success, thus "depersonalizing failure" in the process (Borkowski et al., 1989; Oka & Paris, 1987). To inoculate students against interpretations of lack of control, teachers could model failure situations due to ineffective strategy use and demonstrate how strategy based effort can help overcome failure. In addition, students could be shown how information acquired from unsuccessful experiences could be used to improve performance on subsequent tasks.

Finally, it may help to make adolescents aware that ability, attributions, self-efficacy, metacognitive knowledge, cognitive strategy use, and metacognitive strategy use represent six separable components of learning. These components play unique and perhaps compensatory

roles in achievement. Students could be encouraged to understand that differences in general ability need not necessarily limit one's efficacy, nor constrain metacognitive knowledge acquisition or the use of strategies. In addition, students could be encouraged to practice a variety of cognitive and metacognitive strategies until these strategies are automatized and students feel confident using them (Alexander & Judy, 1988; Pressley et al., 1987). Failure to do so may place them at risk for inappropriately attributing reading failures to lack of ability or metacognitive knowledge.

In conclusion, the findings of the study imply that a variety of variables might be used productively in training programmes to improve reading performance. The interdependence of motivation, metacognitive knowledge, and cognitive/metacognitive strategy use variables in the metacognitive-motivational system suggest that enhancement programmes which include strategy and motivational variables could be more effective in improving reading performance than programmes which consider these variables in isolation. Therefore, parents and educators need also consider the affective and motivational well being of students (particularly RD students) when evaluating their progress and creating strategy training programmes.

10.8 Limitations and strengths of present study

10.8.1 Limitations

Although the results of the present study seem to provide relatively strong support for Borkowski et al.'s (1989, 1990, 1992) model, one must nevertheless be cautious in accepting this conclusion. Hence, several caveats should be noted in interpreting the findings of the present study.

The first limitation involves the characteristics of the sample. Although the present research findings confirm many of the findings of previous research, the generalizability of the findings to the broader population of adolescents is tentative as the study involved a highly selected group of RD readers with restricted variance in their IQ and performance scores. Furthermore, the occurrence of a negative variance for the PAT measure for the RD reader group (see Chapter 9, section 9.3.1.1), although circumvented by fixing the measurement error, suggests results must be interpreted cautiously.

Another major limitation involved the intertwining of selection criteria with components of the model. In the present study the group selection process was correlated with two variables of interest, namely, intelligence and reading performance, resulting in questions about the validity of the structural model. Research has demonstrated that sample selection bias can result in biased and inconsistent estimates of structural parameters (see Berk, 1983). The implication of previous research is that even though the same model may hold in different groups, the selection process may lead to a situation in which there are apparent group differences in model structure or model coefficients (L. J. Horwood, personal communication, March 27, 1995). However, some reassurance on this matter can be obtained. First, the selection process consisted of a multidefinitonal approach involving variables that were not included in the model, and the IQ and PAT variables constituted only one aspect of the selection process. Additionally, the models in the present study suggest largely identical structural parameters between groups. Finally, although different selection procedures were used and different samples obtained for the present study and Carr et al. (1991), the present study obtained a model rather similar to that of Carr and her colleagues. This would suggest that the general structural relationships which hold between the variables of interest (and common across studies) are fairly robust to sampling considerations.

The second limitation concerns the structural equation modelling procedures used in the present study. Most of the concerns are relevant to SEM procedures and not specific to the study, but are still worth noting. Although SEM procedures are based on correlations, and correlation does not necessarily imply causality, results gleaned from SEMs are usually discussed in "causal" terms. As Mulaik (1987) has argued, "causation implies correlation" (p. 20), thus when testing hypotheses regarding causal relationships, and when finding support for these relationships, it only makes sense to discuss the results in terms of support for causality (see Bagozzi, 1980; and Biddle & Marlin, 1987, for further discussion of this point). Although SEM cannot be used to establish causality or the validity of one theory over another, it has implications for theory development. As such, SEM allows a researcher to compare and contrast several alternative theoretical models as possible explanations of performance (Carr et al., 1991).

The next issue relates to the a posteriori modifications made to the hypothesized models. The demonstrated goodness of fit between the models and the data is owed in part to the changes made in the hypothesized model during the specification search. While such a process is generally accepted, it raises concerns about the risk of capitalizing on chance. As various

authors have emphasized, any changes made in the structural model during the specification search must therefore be justified on theoretical as well as mathematical grounds (Cliff, 1983; MacCallum, 1986). In consideration of this, a discussion of such substantive theoretical considerations concerning parameters that were added or deleted in the model was included in the present study (see Chapter 9, section 9.3.1.1). However, one issue, relating to correlated errors, deserves mention again because a number of researchers have reservations concerning the inclusion of correlated errors in models. In the present study, although these correlated errors were permitted on justifiable grounds (see Chapter 7, section 7.4.2), there remains the doubt that by including these correlated errors, one may be capitalising on chance variation in the data. Experimentation with the fitted models demonstrated that regardless of the pattern of correlated errors fitted for a given model, a similar set of underlying structural relationships were observed to hold. Thus, the inclusion of correlated errors had minimal impact on the estimated structural parameters of the models but significantly improved fit of the models. Nevertheless, it is important to cross validate the findings of the study. This was not possible in the present study because of the moderate sample size. It is also notable that the inclusion of a large number of correlated errors may have resulted in overfitting of the models as indicated by the high probability levels.

The third issue relates to the use of single indicators. The LISREL analyses may have been hindered by the lack of multiple indicators for the ability construct. In hindsight, it may have been useful to include more indicators, as the reliability estimate for the measure of the WISC-R for the sample in the present study was lower than that obtained for other studies (L. J. Horwood, personal communication, August 23, 1994).

The fourth issue relates to alternative ways of testing the hypothesized models, particularly with regard to the factor structure of success/failure attributions and cognitive/metacognitive strategy use. Although tests of one factor and two factor representations of attributional style and strategy use measures were conducted, there is a third alternative. This consists of a second order factor model with two first order factors of success/failure attributions and cognitive/metacognitive strategy use, and a single second order factor of attributional style and combined strategy use. In this second order factor model, success/failure attributions and cognitive/metacognitive strategy use variables are merely indicators. In the case of the strategy use factors, such a model examines the question "Does the elaboration of strategy use into separate metacognitive and cognitive components add anything to the explanation of the reading performance over and above combined strategy use?" In the case of the attributional

style factors, such a model examines the question "Does the elaboration of attributional style into separate success and failure attribution components add anything to the explanation of the reading performance over and above attributional style?"

Although it is common practice for social science researchers to test the models such as those in the present study using one and two factor representations, these higher order factor representations of the attributions and strategy use constructs do provide alternative means to synthesize the representation of these measures. Attempts were made to fit models using a second order factor model with two first order factors for the attributions and strategy use constructs. Unfortunately, despite experimentation over a wide range of potential starting values for the model parameters, these models proved impossible to implement in the LISREL programme.

The last issue relates to the hypothesized structure of the model. As with all mediational models, it is probably always possible to conceive of alternative formulations that account for the data set equally well. The results that are reported in this study are based on a hypothesized model of students' strategy use based on the author's interpretation of the literature. It is possible that another model with differently specified paths might also fit the data. Since in principle it is impossible to conclusively "prove a mediational theory", the researcher's hope can only be that over time the accumulated evidence becomes "inescapably consistent with that theory" (S. E. Taylor & Fiske, 1981).

The third limitation involves the nature of the design. The cross sectional design used in the current study focuses on the way adolescents' motivational beliefs, metacognitive knowledge, cognitive/metacognitive strategy use, and reading achievement are related at a particular point in time. It does not consider the possible dynamic and developmental nature of these relationships. Neither does it provide insight into the hypothesized reciprocal relationships between students' motivational beliefs, metacognitive knowledge, and strategy use proposed by Borkowski et al. (1989, 1990, 1992).

The next limitation involves the method by which strategy use was measured. Although data were collected through paper-and-pencil instruments and individual interviews, they still represented self-report data. While self-report methods are generally accepted for collecting data on strategy use, it is not clear the extent to which students accurately report their use of strategies. The results that are reported here are based on the assumption that the instruments

are valid measures of students' strategy use; the psychometric data of these measures in the present study appear to support this. Also, several researchers have documented that self-reported strategy use is significantly correlated with observable strategy use (Brown, 1980; Justice & Weaver-McDougall, 1989).

The results of this study are additionally limited by the nature of the items included on the paper-and-pencil instrument and in the interview for strategy use. The purpose of this study was to examine students' use of strategies with text. To this end the questions about students' strategy use could have been better matched across the questionnaire and interview formats. Questions on the paper-and-pencil instrument related to students' general use of strategies with text, while questions in the interview related to students' use of strategies with a specific text task. It may have been appropriate to design questions so that observations of students' monitoring and regulation of cognitive strategies could be made. Furthermore, the measures of metacognitive strategy use did not often measure aspects of strategy selection and task analysis, which was the focus of Borkowski's definition (Borkowski & Turner, 1990).

In retrospect, a frequency measure of strategy use, as was employed in the present study, may not have provided an accurate appraisal of the relationship between strategy use and performance. It was noted on several occasions that although RD students reported a cognitive strategy that was appropriate for the task, they were not able to answer the question correctly. In other words, students did not use the strategy effectively. Hence, a frequency measure of strategy use may not yield an accurate picture of the relationship between strategy use and performance. Swanson's research also suggested that measures assessing strategy efficiency are likely to provide a clearer picture of the relationship between strategy use and performance (Swanson, 1989; Swanson & Cooney, 1985).

10.8.2 Strengths

Despite the constraints noted in section 10.8.1, the present study has several strengths and provides several contributions to theory and research. First, the present study may have made an important advancement in taking a multidimensional approach both in defining RD readers and in investigating reading comprehension performance. Despite recommendations for a multifaceted definition, this study appears to be one of the first to have implemented such an operational definition in an empirical study. The criteria used for this multidefinitional

process were those that were derived from the learning/reading disability literature and represented common elements across definitions. Support for the validity of such an approach was provided by the results of the discriminant analyses.

Although a number of studies have used unicomponential approaches to investigate reading comprehension performance, the present study developed models based on a multicomponential view of reading comprehension. In addition, these models were developed within a causal framework. Studies that have investigated reading performance using a multicomponential approach have often used designs that established correlational rather than causal relationships (e.g., Paris & Oka, 1986).

A second strength involves the general statistical model applied in the present study which represents a useful new direction for multicomponential research on reading comprehension. Few studies have used structural equation models, which are statistically superior to investigations relying on correlations or path analyses. By using the inferential power of structural equation modelling techniques, the current study has afforded a more stringent test of Borkowski et al.'s (1989, 1990, 1992) theory. The use of such techniques permitted the testing of all of the links in a mediational model simultaneously, rather than in the typical piecemeal fashion. Structural equation modelling also afforded the possibility of clarifying controversial or ambiguous aspects of the hypothesized model (as well as elaborating and refining the model), and of contrasting the hypothesized model with competing alternative conceptualizations. As a result, it has provided preliminary answers to several unresolved or competing issues (concerning adequacy or inadequacy of particular paths) associated with the model. For example, theoretical issues concerning the relationships between ability and performance, ability and efficacy, cognitive strategy use and performance, metacognitive strategy use and cognitive strategy use, and attributions and metacognitive strategy use were all addressed.

The present study also incorporated many of the principal components of Borkowski et al.'s (1989, 1990, 1992) model that have previously been studied in isolation. In fact, the present study may be the first to examine most components in Borkowski's model of metacognition, particularly in relation to metacognitive knowledge, cognitive strategy use, and metacognitive strategy use variables. Path analytic or regression studies to date that have investigated relationships between metacognitive and motivational variables, have usually omitted either the metacognitive knowledge or the cognitive/metacognitive strategy use variables. The

present study provides some preliminary empirical support for the causal mediating role of metacognitive knowledge and strategy use in the relationship between motivational beliefs and reading performance.

A fourth strength of the current study is that it extends the literature on causal attributions in achievement contexts by exploring the separate effects of attributional subcomponents on metacognitive knowledge and self-efficacy. The current study also used domain specific measures whereas previous studies have relied almost exclusively on generalized measures of motivational beliefs, strategy use, and academic performance (Zimmerman & Martinez-Pons, 1992).

Unlike previous research which has relied on standardized tests (e.g., Paris & Oka, 1986), this study included both standardized reading achievement test scores and classroom marks as indicators of reading performance. The use of a standardized reading test as the sole measure of comprehension in previous research has raised some concerns (DeFina, Anstendig, & DeLawler, 1991). It is argued that children are generally well aware of their classroom marks while they are often not informed of their standardized test scores. Therefore, if academic achievement is presumed to influence the child's motivational beliefs, it is more likely that this relationship would be more accurately modeled when classroom marks are included as an indicator of achievement.

10.9 Future research

The findings and design of the present research have some implications for future research on causal models of reading comprehension. One goal of future research would be to further substantiate the mechanisms underlying the links between attribution subcomponents and metacognitive knowledge and the link between attribution subcomponents and efficacy. Another related line of research could examine the separate effects of attributional *dimensions* (e.g., controllability, stability) on metacognitive knowledge and efficacy. Furthermore, attributional theorists have to date been largely unsuccessful in overcoming conceptual and methodological problems related to attribution measurement. Consequently, an important preliminary step for future research in this area could be establishing valid and reliable measures of children's attributional beliefs about reading.

A stronger design than employed here could be attempted in future studies when replicating the findings of this study. Byrne (1984) recommended that studies which address the direction-of-causality question follow these prerequisites: (a) a statistical relation must be established, (b) longitudinal designs should be used with a clearly established time precedence, and (c) a causal model must be tested. Although a longitudinal design and structural modelling improves the capacity to make causal inferences, such inferences are not necessarily conclusive (Holland, 1986). However, by testing alternative models, supplementing these with intervention research, and cross-validating results as a part of the model-building process, investigators are in a better position to interpret their findings with more confidence.

The relationships that were hypothesized in this study between the measures of attributional style, self-efficacy, metacognitive knowledge, and metacognitive/cognitive strategy use were unidirectional. It is likely that the relationships among these variables are more dynamic and reciprocal, as indicated in Borkowski et al.'s (1989, 1990, 1992) model. However, testing reciprocal relationships was beyond the scope of the study given the moderate sample size. A longitudinal design with SEM techniques, coupled with intervention research, could be used in future research to investigate this question.

Given the complex nature of performance, it must be acknowledged that many other variables are likely to influence reading attainment. Therefore it could be argued that the present study focused on a limited number of factors. Moreover, it seems clear from the moderate levels of explained variance (60%) that there are other factors implicated in reading performance. Consequently, future research may benefit from the inclusion of additional factors. Research concerning academic achievement has identified a number of factors that may be important to reading achievement. These include personal goals (Ames & Archer, 1988; Meece et al., 1988), task value and interest (Pintrich & De Groot, 1990), instructional context (Ames & Archer, 1988; Zimmerman, 1989b), and classroom settings (e.g., competitive-cooperative or ability grouped-ungrouped). In addition, family and cultural backgrounds are increasingly being considered by recent research to be important predictors of academic performance (Kurtz, 1990).

Another important direction for future research could be to assess the generality of findings of the present study. Although the present analyses suggest particularly strong interrelationships between attributional style, efficacy, metacognitive knowledge, metacognitive/cognitive strategy use, and reading achievement, it is important to note

developmental limitations of this study. Research has indicated that motivational beliefs, metacognitive knowledge, and strategy use of children change with age. It is quite possible the relationships amongst the aforementioned variables may differ with age (Chan, 1995; Paris & Oka, 1986). Follow-up investigations tracing the development of motivational beliefs, metacognitive knowledge, cognitive/metacognitive strategy use, and reading achievement over time may further clarify effects. Future research assessing the generality of the present study's findings across different types of reading tasks may also be useful. For example, research has indicated that attributions vary with different types of reading tasks such as reading for leisure versus reading for a test.

Future studies of this model and others would benefit from the use of multiple and alternative measures of some constructs (e.g., ability, metacognitive knowledge, strategy use) and the investigation of other academic areas. Model tests across different populations would also be beneficial, especially those who are at risk of school failure. These efforts would provide further insight into the process of metacognitive, cognitive, and affective development and the generalizability of the hypothesized model across other groups of children.

Some of the limitations of the present study addressed in section 10.8.1 could provide fruitful avenues for future research. For instance, the findings could be replicated using a similar operational definition of metacognitive strategy use as was used by Borkowski et al. (1989, 1990, 1992). In addition, efficiency and frequency measures of strategy use could be employed in future research so as to provide a more accurate appraisal of the relationships between cognitive/metacognitive strategy use and reading performance. Future research might also be profitably directed toward determining the factors that predict students' ability to utilize strategies in an effective and efficient manner.

10.10 Summary and conclusions

A number of significant findings and implications have arisen from this dissertation. First, the present study demonstrates the saliency of metacognition to reading disability. Metacognition is relevant to reading disabilities because it broadens researcher's perceptions and understanding of reading problems and highlights the importance of metacognitive skills (Wong, 1987, 1991). Furthermore, it shifts the focus from specific cognitive deficits that may not be malleable to strategic deficits which are amenable to instruction.

Second, the results provide support for the specification and formulation of theoretical linkages amongst students' motivational beliefs, their metacognitive knowledge, and cognitive and metacognitive strategic deployment in predicting reading performance and impairment. The study reported herein provides evidence of the interdependence of ability, attributional style, self-efficacy, metacognitive knowledge, cognitive strategy use, and metacognitive strategy use in predicting reading comprehension. These variables appear to operate as part of a dynamic network. The results corroborate previous empirical findings and provide general support for the metacognition model proposed by Borkowski et al. (1989, 1990, 1992).

Third, there is considerable support for a strong relationship between motivational and cognitive/metacognitive factors, suggesting that metacognitive and cognitive processes have roots in the self-system (self-efficacy, attributional beliefs). Although motivational beliefs and cognitive/metacognitive strategy use are linked, strategy use is promoted by positive perceptions of self-efficacy and a sense of personal causality. Therefore, the incorporation of attributional style and self-efficacy as separate factors in metacognitive models seems pertinent for a more complete understanding of reading achievement and impairment.

Fourth, one significant focus of the present study was to not only explain normal cognitive development but also atypical cognitive development in the domain of reading. This presents an important goal in cognitive developmental research, particularly with regard to developing parsimonious models. The findings of this study suggest that the gap between actual and expected achievement in RD students can be explained by an analysis of cognitive, metacognitive, and motivational variables. The model suggests that the performance of RD readers may be explained by (a) cognitive strategy deficiencies, as reflected in their significantly lower cognitive strategy use scores; (b) metacognitive deficiencies, as reflected in their significantly lower scores of metacognitive knowledge, and metacognitive strategy use; and (c) faulty motivational style, as reflected in their significantly lower scores of self-efficacy and attributions.

While the relationships between motivational beliefs, metacognitive knowledge, and cognitive/metacognitive strategy use were similar across groups, distinct metacognitive-motivational models were required to explain reading achievement and impairment. Furthermore, the results suggest that differences in NA and RD adolescents may be due to differences in the ways attributional beliefs and self-efficacy are related to metacognitive development and also in the way cognitive strategy use is related to reading comprehension.

Thus, the development of skilled reading or its converse requires consideration of how cognitive skills, metacognitive skills, and motivational style interact to affect reading.

Fifth, several educational implications were forthcoming from the findings of the present study. Prospective targets of instruction include inducing high self-efficacy, ensuring adequate metacognitive knowledge, encouraging consistent, reflective, and efficient use of cognitive/metacognitive strategies, and fostering adaptive attributions for success and failure outcomes. Two important educational implications merit mentioning again. First, to enhance the academic outcomes of all students, educators must avoid developing cognitive and motivational enhancement programmes in isolation, particularly in the light of findings that suggest motivational beliefs spur the development of the metacognitive system. Second, the findings support the importance of implementing interventions with a focus on failure attributions. The teaching of effective and adaptive coping mechanisms when faced with failure is likely to "depersonalize" failure experiences. This in turn is likely to promote further development of the strategic behaviour of both NA and RD students.

Sixth, the findings of the present study provide potential insights for the conduct of future research. Two important areas deserve mention again. First, it may be necessary to further substantiate the mechanisms underlying the links between attribution subcomponents and metacognitive knowledge and the links between attribution subcomponents and efficacy. In addition, the findings of the present study could be replicated with a stronger design than that employed here. Second, future research might also be profitably directed toward determining the factors that predict students' use of effective strategies, and in investigating the relationship between strategy use and performance, using improved measures of metacognition and strategy use.

In conclusion, this study demonstrates the utility of integrating motivational, metacognitive, and cognitive approaches to better understand reading achievement in school settings. In exploring the generalizability of models across populations, the present study contributed to scientific research by developing a more parsimonious theory of reading comprehension. By developing models that draw on the strengths of different theoretical approaches, researchers can begin to build more comprehensive models of reading achievement. The findings highlight the importance of formulating achievement models that examine relations among attributional style, self-efficacy, metacognitive knowledge, cognitive/metacognitive strategy use and their causal links to achievement. Such a focus highlights the complexity of the learning process as an interactive and dynamic process.

REFERENCES

- Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *Journal of Learning Disabilities, 24*, 178-186, 191.
- Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358-372.
- Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans. *Journal of Abnormal Psychology, 87*, 49-74.
- Afflerbach, P., & Johnston, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading Behavior, 16*, 307-322.
- Aigner, D. J., & Goldberger, A. S. (1977). *Latent variables in socioeconomic models*. New York: North Holland.
- Alexander, P. A., & Judy, J. E. (1988). The interaction of domain specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.
- Allon, M., Gutkin, T. B., & Bruning, R. (1994). The relationship between metacognition and intelligence in normal adolescents: Some tentative but surprising findings. *Psychology in the Schools, 31*, 93-97.
- Alvermann, D. E., & Moore, D. W. (1991). Secondary school reading. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 951-983). New York: Longman.
- Alwin, D. F., & Jackson, D. J. (1981). Applications of simultaneous factor analysis to issues of factorial invariance. In D. Jackson & E. Borgatta (Eds.), *Factor analysis and measurement in sociological research: A multidimensional perspective* (pp. 249-279). Beverly Hills: Sage.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author
- Ames, C., & Ames, R. (1984). Systems of student and teacher motivation: Toward a qualitative definition. *Journal of Educational Psychology, 76*, 535-556.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267.
- Anderson, C., Jennings, D., & Arnoult, L. (1988). Validity and utility of the attributional style construct at a moderate level of specificity. *Journal of Personality and Social Psychology, 55*, 979-990.
- Anderson, J. C., & Gerbing, D. W. (1982). Some methods of respecifying measurement models to obtain unidimensional construct measurement. *Journal of Marketing Research, 19*, 453-460.
- Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness of fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*, 155-173.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.

- Anthony, G. (1994). Learning strategies in the mathematics classroom: What can we learn from stimulated recall interviews? *New Zealand Journal of Educational Studies*, 29, 127-140.
- Ashman, A., & Elkins, J. (1994). *Educating children with special needs* (2nd ed.). Victoria, Australia: Prentice Hall.
- Assor, A., & Connell, J. P. (1992). The validity of students' self-reports as measures of performance affecting self-appraisals. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 25-47). Hillsdale, NJ: Erlbaum.
- Atkinson, J. W., & Birch, D. (1978). *An introduction to motivation* (2nd ed.). NY: van Nostrand.
- Bagozzi, R. P. (1980). *Causal models in marketing*. New York: Wiley.
- Barclay, C. R., & Hagen, J. W. (1982). The development of mediated behavior in children. In J. P. Das, R. F. Mulcahy, & A. E. Wall (Eds.), *Theory and research in learning disabilities*. New York: Plenum Press.
- Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 155-205). Orlando, FL: Academic Press.
- Baker, L., & Brown, A. L. (1984a). Metacognitive skills and reading. In P. D. Pearson, M. Kamil, R. Barr, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- Baker, L., & Brown, A. L. (1984b). Cognitive monitoring in reading. In J. Floom (Ed.), *Understanding reading comprehension*. Newark, DE: International Reading Association.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J. H. Flavell & L. Ross (Eds.), *Social cognitive development: Frontiers and possible futures* (pp. 200-239). Cambridge, England: Cambridge University Press.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1988). Self-efficacy mechanism in physiological activation and health-promoting behavior. In J. Madden, S. Matthyse, & J. Barchas (Eds.), *Adaptation, learning, and affect*. New York: Raven Press.
- Bandura, A. (1990). Conclusion: Reflections on nonability determinants of competence. In R. J. Sternberg & J. Kolligan (Eds.), *Competence considered* (pp. 315-362). New Haven, London: Yale University Press.
- Bandura, A., Adams, N. E., Hardy, A. B., & Howells, G. N. (1980). Tests of the generality of self-efficacy theory. *Cognitive Therapy and Research*, 4, 39-66

- Bar-Tal, D. (1982). The effects of teachers' behavior on pupils' attributions: A review. In C. Antaki & C. Brewin (Eds.), *Attributions and psychological change: Applications of attributional theories to clinical and educational practice*. New York: Academic Press.
- Bar-Tal, D. (1991). Attribution theory of human development. In K. Marjoribanks (Ed.), *The foundations of students' learning* (pp. 229-240). Oxford, UK: Pergamon Press.
- Bauer, R. H. (1979). Recall after a short delay and acquisition in learning disabled and nondisabled children. *Journal of Learning Disabilities, 12*, 596-607.
- Bauer, R. H., & Emhert, J. (1984) Information processing in reading-disabled and nondisabled children. *Journal of Experimental Child Psychology, 37*, 271-281.
- Baumann, J. F., Seifert-Kessell, N., & Jones, L. (1987, December). *Effects of think-aloud instruction on elementary students' ability to monitor their comprehension*. Paper presented at the National Reading Conference, St. Petersburg, FL.
- Bean, T., Singer, H., & Sorter, J. (1986). Direct instruction in metacognitive strategies. *Journal of Reading Behavior, 18*, 153-169.
- Beck, L. R., & St. George, R. (1983). The alleged cultural bias of PAT Reading Comprehension and Reading Vocabulary tests. *New Zealand Journal of Educational Studies, 18*, 32-47.
- Belmont, J. (1989). Cognitive strategies and strategic learning: The social-instructional approach. *American Psychologist, 44*, 142-148.
- Benson, M. J. (1989). Attributional measurement techniques: Classification and comparison of approaches for measuring causal dimensions. *The Journal of Social Psychology, 129*, 307-323.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology, 31*, 419-456.
- Bentler, P. M. (1982). Linear systems with multiple levels and types of latent variables. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Causality, structure, prediction* (Part I, pp. 101-130). Amsterdam: North Holland.
- Bentler, P. M. (1990). Latent variable structural models for separating specific from general effects. In L. Sechrest, E. Perrin, & J. Bunker (Eds.), *Research methodology: Strengthening causal interpretations of nonexperimental data* (pp. 61-80). Newbury Park, CA: Sage.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588-606.
- Bentler, P. M., & Chou, C. (1987). Practical issues in structural modeling. *Sociological Methods in Research, 16*, 78-117.
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review, 48*, 386-398.
- Berndt, T. J., & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology, 82*, 319-326.

- Berninger, V., Hart, T., Abbott, R., & Karovsky, P. (1992). Defining reading and writing disabilities with and without IQ: A flexible, developmental perspective. *Learning Disability Quarterly, 15*, 103-117.
- Biddle, B., & Marlin, M. (1987). Causality, confirmation, credulity, and structural equation modeling. *Child Development, 58*, 4-17.
- Blalock, H. M. (Ed.) (1985). *Causal models in the social sciences* (2nd ed.) NY: Aldine.
- Blaxall, J., & Willows, D. M. (1984). Reading ability and text difficulty as influences on second graders' oral reading errors. *Journal of Educational Psychology, 76*, 330-341.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. *Developmental Medicine and Child Neurology, 15*, 663-687.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Newbury Park, CA: Sage.
- Boomsa, A. (1982). The robustness of LISREL against small sample size in factor analysis models. In K. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part I, pp. 149-173). Amsterdam: North Holland.
- Boomsa, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika, 50*, 229-242.
- Boomsa, A. (1987). The robustness of maximum likelihood estimation in structural equation models. In P. Cuttance & R. Ecob (Eds.), *Structural modeling by example: Applications in educational, sociological, and behavioral research* (pp. 168-188). New York: Cambridge University Press.
- Borkowski, J. G., Carr, M., & Pressley, M. (1987). "Spontaneous" strategy use: Perspectives from metacognitive theory. *Intelligence, 11*, 61-75.
- Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 53-92). Hillsdale, NJ: Erlbaum.
- Borkowski, J. G., Day, J. D., Saenz, D., Dietmeyer, D., Estrada, T. M., & Groteluschen, A. (1992). Expanding the boundaries of cognitive interventions. In B. Y. L. Wong (Ed.), *Contemporary intervention and research in learning disabilities: An international perspective* (pp. 1-21). New York: McGraw-Hill.
- Borkowski, J. G., Estrada, M. T., Milstead, M., & Hale, C. (1989). General problem-solving skills: Relations between metacognition and strategic processing. *Learning Disability Quarterly, 12*, 57-70.
- Borkowski, J. G., Johnston, M. B., & Reid, M. K. (1987). Metacognition, motivation, and controlled performance. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 147-173). Hillsdale, NJ: Erlbaum.

- Borkowski, J. G., & Kurtz, B. E. (1984). Metacognition and special children. In J. B. Gholson & T. I. Rosenthal (Eds.), *Applications of cognitive developmental theory* (pp. 193-213). New York: Academic Press.
- Borkowski, J. G., & Kurtz, B. E. (1987). Metacognition and executive control. In J. G. Borkowski & J. D. Day (Eds.), *Cognition in special children: Comparative approaches to retardation, learning disabilities, and giftedness* (pp. 123-152). Norwood, NJ: Ablex.
- Borkowski, J. G., Milstead, M., & Hale, C. (1988). Components of children's metamemory: Implications for strategy generalization. In F. E. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 73-100). Hillsdale, NJ: Erlbaum.
- Borkowski, J. G., & Muthukrishna, N. (1992). Moving metacognition into the classroom: Working models and effective strategy teaching. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 478 - 501). San Diego, CA: Academic Press.
- Borkowski, J. G., Peck, V., Reid, M. K., & Kurtz, B. (1983). Impulsivity and strategy transfer: Metamemory as mediator. *Child Development*, 54, 459-473.
- Borkowski, J. G., & Turner, L. A. (1990). Transsituational characteristics of metacognition. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance* (pp. 159-176). New York: Springer-Verlag.
- Borkowski, J. G., Weyhing, R. S., & Carr, M. (1988). Effects of attributional retraining on strategy-based reading comprehension in learning-disabled students. *Journal of Educational Psychology*, 80, 46-53.
- Bos, C. S., & Filip, D. (1984). Comprehension monitoring in learning disabled and average students. *Journal of Learning Disabilities*, 17, 229-233.
- Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology*, 130, 353-363.
- Bowers, P. G., Steffy, R., & Tate, E. (1988). Comparison of the effects of IQ control methods on memory and naming speed predictors of reading disability. *Reading Research Quarterly*, XXIII, 304-319.
- Bray, N. W. (1985). *Why are mentally retarded strategically deficient?* Symposium presented at the 18th annual Gatlinburg Conference on Research and Theory in Mental Retardation and Developmental Disabilities, Gatlinburg, TN.
- Brown, A. L. (1974). The role of strategic behavior in retarded memory. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 7, pp. 55-104). New York: Academic Press.
- Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 103-152). New York: Academic Press.
- Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology, Volume 1*. Hillsdale, NJ: Erlbaum.

- Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension. Perspectives from cognitive psychology, linguistics, artificial intelligence and education* (pp. 453-481). Hillsdale, NJ: Erlbaum.
- Brown, A. L. (1981). Metacognition and reading and writing: The development and facilitation of selective attention strategies for learning from texts. In M. L. Kamil (Ed.), *Directions in reading: Research and instruction* (pp. 21-43). Washington, DC: National Reading Conference.
- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.
- Brown, A. L., Armbruster, B. B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Oransanu (Ed.), *Reading comprehension: From research to practice*. Hillsdale, NJ: Erlbaum.
- Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of child psychology: Cognitive development* (Vol. 3, pp. 77-166). New York: Wiley.
- Brown, A., & Palincsar, A. (1982). Inducing strategic knowledge from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities, 2*, 1-18
- Brown, A. L., & Smiley, S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development, 48*, 1-8.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Burns, R. (1982). *Self-concept development and education*. London: Holt, Rinehart, & Winston.
- Butkowsky, I. S., & Willows, D. M. (1980). Cognitive-motivational characteristics of children varying in reading ability: Evidence for learned helplessness in poor readers. *Journal of Educational Psychology, 72*, 408-422.
- Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research, 54*, 427-456.
- Byrne, B. M. (1989a). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Byrne, B. M. (1989b). Multigroup comparisons and the assumption of equivalent construct validity across groups: Methodological and substantive issues. *Multivariate Behavioral Research, 24*, 503-523.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.
- Campione, J. C. (1987). Metacognitive components of instructional research with problem learners. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 117-140). Hillsdale, NJ: Erlbaum.

- Carmine, E. G. (1986). The analysis of covariance structure models. In W. D. Berry, & M. S. Lewis-Beck (Eds.), *New tools for social scientists: Advances and applications in research methods* (pp. 23-55). Beverly Hills: Sage.
- Carr, M., & Borkowski, J. G. (1989). Attributional training and the generalization of reading strategies with underachieving children. *Learning and Individual Differences, 1*, 327-341.
- Carr, M., Borkowski, J. G., & Maxwell, S. E. (1991). Motivational components of underachievement. *Developmental Psychology, 27*, 108-118.
- Carver, C. S. (1989). How should multifaceted personality constructs be tested? Issues illustrated by self-monitoring, attributional style, and hardiness. *Journal of Personality and Social Psychology, 56*, 577-585.
- Chalfant, J. C. (1987). Learning disabilities: Problems in definition of learning disabilities. In C. R. Reynolds & L. Mann (Eds.), *Encyclopedia of special education* (Vol. 2, pp. 925-929). New York: Wiley.
- Chalfant, J. C. (1989). Learning disabilities: Policy issues and promising approaches. *American Psychologist, 44*, 392-398.
- Chan, K. S. (1995). Relationship of motivation, strategic learning, and reading achievement in Grades 5, 7, and 9. *Journal of Experimental Education, 62*, 319-339.
- Chapman, J. W. (1985, June). *Self-perceptions of ability, learned helplessness and academic achievement expectations of children with learning disabilities: A comparative longitudinal study* (Research Contract 55-2-18). Palmerston North, New Zealand: Massey University, Department of Education.
- Chapman, J. W. (1992). Learning disabilities in New Zealand: Where kiwis and kids with LD can't fly. *Journal of Learning Disabilities, 25*, 362-370.
- Chapman, J. W., St. George, R., & van Kraayenoord, C. E. (1983, August). *Short WISCs, short cuts, and Occam's razor: Identification of learning disabled pupils*. Paper presented at the annual conference on New Zealand Psychological Society, Auckland.
- Chapman, J. W., St. George, R., & van Kraayenoord, C. E. (1984). Identification of "learning disabled" pupils in a New Zealand Form 1 sample. *Australian and New Zealand Journal of Developmental Disabilities, 10*, 141-149.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology, 87*, 154-167.
- Chapman, J. W., & van Kraayenoord, C. E. (1987). Learning disabilities. In D. Mitchell & N. Singh (Eds.), *Exceptional children in New Zealand* (pp. 304-317). Palmerston North, New Zealand: Dunmore Press.
- Chapman, J. W., & Wilkinson, L. (1988). *Self-Concept, identity status and emotional status of form 5 learning disabled students: A longitudinal study of intermediate school predictors* (Research Contract No. 47/17/151). Palmerston North, New Zealand: Massey University, Department of Education.

- Chi, M. T. H., Glaser, R., & Farr, M. (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum.
- Clay, M. M. (1972). Learning disorders. In S. I. Havill & D. R. Mitchell (Eds.), *Issues in New Zealand Special Education*. Auckland, New Zealand: Hodder & Stoughton.
- Clay, M. M. (1973). *Reading: The patterning of complex behavior*. Auckland, New Zealand: Heinemann Educational Books.
- Clay, M. M. (1985). *The early detection of reading difficulties*. Auckland, New Zealand: Heinemann Educational Books.
- Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18, 115-126.
- Clifford, M. M. (1980). Effects of failure: Alternative explanations and possible implications. In L. J. Fyans (Ed.), *Achievement motivation, recent trends in theory and research* (pp. 96-113). New York: Plenum.
- Clifford, M. M. (1984). Thoughts on a theory of constructive failure. *Educational Psychologist*, 19, 108-120.
- Cohen, S. G., McDonell, G., & Osborn, B. (1989). Self-perceptions of "at-risk" and high achieving readers: Beyond Reading Recovery achievement data. In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction, Thirty-eight yearbook of the National Reading Conference* (pp. 117-122). Chicago: National Reading Conference.
- Cole, D. A. (1989). Psychopathology of adolescent suicide: Hopelessness, coping beliefs, and depression. *Journal of Abnormal Psychology*, 98, 248-255.
- Cole, J. Y., & Sticht, T. G. (1981). *The textbook in American Psychology*. Washington: Library of Congress.
- Coles, G. S. (1987). *The learning mystique: A critical look at learning disabilities*. New York: Pantheon Books.
- Cooley, E. J., & Ayres, R. R. (1988). Self-concept and success-failure attributions of nonhandicapped students and students with learning disabilities. *Journal of Learning Disabilities*, 21, 174-178.
- Cooney, J. B., & Swanson, H. L. (1987). Memory and learning disabilities: An overview. In H. L. Swanson (Ed.), *Memory and learning disabilities* (pp. 1-40). Greenwich, CT: JAI Press.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.
- Corno, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology*, 11, 333-346.
- Cornoldi, C. (1990). Metacognitive control processes and memory deficits in poor comprehenders. *Learning Disability Quarterly*, 13, 245-255.
- Costner, H. L. (1985). Theory deduction and rules of correspondence. In H. M. Blalock (Ed.), *Causal models in the social sciences* (pp. 229-250). NY: Aldine.

- Covington, M. V. (1989). Self-esteem and failure in school: Analysis and policy implications. In A. M. Mecca, N. J. Semler, & J. Vasconcellos (Eds.), *The social importance of self-esteem* (pp. 72-124). Berkeley: University of California Press.
- Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. New York: Cambridge University Press.
- Covington, M. V., & Omelich, C. L. (1981). As failures mount: Affective and cognitive consequences of ability demotion in the classroom. *Journal of Educational Psychology, 73*, 799-808.
- Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology, 80*, 131-142.
- Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 512-538). New York: Longman.
- Davey, B. (1987). Postpassage questions: Task and reader effects on comprehension and metacomprehension processes. *Journal of Reading Behavior, 19*, 261-286.
- Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review, 94*, 369-389.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. NY: Plenum.
- DeFina, A. A., Anstendig, L., & DeLawler, K. (1991). Alternative integrated reading/writing assessment and curriculum design. *Journal of Reading, 34*, 354-359.
- Department of Education (1987). *Draft review of special education*. Wellington, New Zealand: Department of Education.
- Deshler, D., Schumaker, J., & Lenz, B. (1984). Academic and cognitive intervention for learning disabled adolescents. *Journal of Learning Disabilities, 17*, 108-119.
- Doris, J. L. (1986). Learning disabilities. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 1, pp. 3-53). Hillsdale, NJ: Erlbaum.
- Doris, J. L. (1993). Defining learning disabilities: A history of the search for consensus. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 97-115). Baltimore: Paul H. Brookes.
- Downing, J., & Leong, C. K. (1982). *Psychology of reading*. New York: Macmillan.
- Dudley-Marling, C. C., Snider, V., & Tarver, S. G. (1982). Locus of control and learning disabilities: A review and discussion. *Perceptual and Motor Skills, 54*, 503-514.
- Duncan, O. D. (1975). *Introduction to structural equation models*. New York: Academic Press.
- Dweck, C. (1987, April). *Children's theories of intelligence: Implications for motivation and learning*. Paper presented at the annual meetings of AERA, Washington DC.

- Earley, P. C., & Lituchy, T. R. (1991). Delineating goal and efficacy effects: A test of three models. *Journal of Applied Psychology, 76*, 81-98.
- Eccles, J. (1987, August). *Understanding motivation: Achievement beliefs, gender-roles, and changing educational environments*. Paper presented at the meeting of the American Psychological Association, New York, NY.
- Eccles (Parsons), J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). San Francisco: Freeman.
- Eden, G. F., Stein, J. F., Wood, M. H., & Wood, F. B. (1995). Verbal and visual problems in reading disability. *Journal of Learning Disabilities, 28*, 272-290.
- Elley, W. B., & Reid, N. A. (1971). *Progressive Achievement Tests: Listening Comprehension*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Ellis, E., Deshler, D., Lenz, K., Schumaker, J., & Clark, F. (1991). An instructional model for teaching learning strategies. *Focus on Exceptional Children, 23*, 1-23.
- Englert, C. S., Raphael, T. E., Fear, K. L., & Anderson, L. M. (1988). Students' metacognitive knowledge about how to write informational texts. *Learning Disability Quarterly, 11*, 18-46.
- Englert, C. S., & Thomas, C. C. (1987). Sensitivity to text structure in reading and writing: A comparison between learning disabled and non-learning disabled students. *Learning Disability Quarterly, 10*, 93-105.
- Epps, S., Ysseldyke, J., & Algozzine, B. (1985). An analysis of the conceptual framework underlying definitions of learning disabilities. *Journal of School Psychology, 23*, 133-144.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.
- Evans, L. D. (1992). Multiple IQ-achievement comparisons: Effects on severe discrepancy determination. *Learning Disability Quarterly, 15*, 167-174.
- Farr, R., & Carey, F. (1986). *Reading: What can be measured?* Newark, DE: International Reading Association.
- Fassinger, R. E. (1987). Use of structural equation modeling in counseling psychology research. *Journal of Counseling Psychology, 34*, 425-436.
- Fergusson, D. M. (1995). A brief introduction to structural equation models. In F. C. Verhulst & H. M. Koot (Eds.), *The epidemiology of child and adolescent psychopathology*. New York: Oxford University Press.
- Fergusson, D. M., & Horwood, L. J. (1987). The trait and method components of ratings of conduct disorders,. Part 1: Maternal and teacher evaluations of conduct disorders in young children. *Journal of Child Psychology and Psychiatry, 28*, 249-260.
- Fergusson, D. M., & Horwood, L. J. (1988). Structural equation modelling of measurement processes in longitudinal data. In M. Rutter (Ed.), *Studies of psychosocial risk: the power of longitudinal data* (pp. 325-353). New York: Cambridge University Press.

- Fergusson, D. M., & Horwood, L. J. (1989). Estimation of method and trait variance in ratings of conduct disorder. *Journal of Child Psychology and Psychiatry*, 30, 365-378.
- Fergusson, D. M., Lloyd, M. & Horwood, L. J. (1991). Family ethnicity, social background and scholastic achievement – an eleven year longitudinal study. *New Zealand Journal of Educational Studies*, 26, 49-63.
- Flavell, J. H. (1970). Developmental studies of mediated memory. In H. W. Reese, & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5, pp. 181-221). New York: Academic Press.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Hillsdale, NJ: Erlbaum.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Fletcher, J. M. (1992). The validity of distinguishing children with language and learning disabilities according to discrepancies with IQ: Introduction to the special series. *Journal of Learning Disabilities*, 25, 546-548.
- Fletcher, J. M., Epsy, K. A., Francis, D. J., Davidson, K. C., Rourke, B. P., & Shaywitz, S. E. (1989). Comparisons of cutoff and regression-based definitions of reading disabilities. *Journal of Learning Disabilities*, 22, 334-338.
- Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1992). The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities*, 25, 555-561.
- Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1993). Classification of learning disabilities: Relationships with other childhood disorders. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 27-55). Baltimore: Paul H. Brookes.
- Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., Francis, D. J., Fowler, A. E., & Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology*, 86, 6-23.
- Fogel, A., & Thelen, E. (1987). Development of early expressive and commutative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, 23, 747-761.
- Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. *Journal of Consumer Research*, 9, 443-448.

- Fornell, C. (1987). A second generation of multivariate analysis: Classification of methods and implications for marketing research. In M. J. Houston (Ed.), *Review of Marketing* (pp. 407-450). Chicago: American Marketing Association.
- Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement errors. *Journal of Marketing Research*, 18, 39-50.
- Forrest, D. L., & Waller, T. G. (1981, April). *Reading ability and knowledge of important information*. Paper presented at the biennial meeting of the Society for Research in Child Development, Boston.
- Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology*, 20, 257-274.
- Frankenberger, W., & Fronzaglio, K. (1991). A review of states' criteria and procedures for identifying children with learning disabilities. *Journal of Learning Disabilities*, 24, 495-500.
- Frankenberger, W., & Harper, J. (1987). States' criteria and procedures for identifying learning disabled children: A comparison of 1981/82 and 1985/86 guidelines. *Journal of Learning Disabilities*, 20, 118-121.
- French, J. N., Ellsworth, N. J., & Amoruso, M. Z. (1995). *Reading and learning disabilities: research and practice*. New York: Garland.
- Garner, R. (1982). Verbal report data on reading strategies. *Journal of Reading Behavior*, 14, 159-167.
- Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research*, 60, 517-529.
- Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and Unanswered questions. *Educational Psychologist*, 24, 143-158.
- Garofalo, J., & Lester, F. K. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, 16, 163-176.
- Gerber, M. M. (1983). Learning disabilities and cognitive strategies: A case for training or constraining problem solving? *Journal of Learning Disabilities*, 16, 255-260.
- Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, 11, 572-580.
- Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (1993), *Testing structural equation models* (pp. 40-65). Newbury Park, CA: Sage.
- German, D., Johnson, B., & Schneider, M. (1985). Learning disability versus reading disability: A survey of practioners' diagnostic populations and test instruments. *Learning Disability Quarterly*, 8, 141-157.
- Ghatala, E. S. (1986). Strategy-monitoring training enables young learners to select effective strategies. *Educational Psychologist*, 21, 43-54.

- Ghatala, E. S., Levin, J. R., Pressley, M., & Lodico, M. G. (1985). Training cognitive strategy-monitoring in children. *American Educational Research Journal*, 22, 199-215.
- Glasser, W. (1969). *Schools without failure*. New York: Harper & Row.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6-10.
- Graham, J. R., & Lilly, R. S. (1984). *Psychological testing*. Englewood Cliffs, NJ: Prentice-Hall.
- Graham, S., & Harris, K. R. (1989). The relevance of IQ in the determination of learning disabilities: Abandoning scores as decision makers. *Journal of Learning Disabilities*, 22, 500-503.
- Graham, S., & Long, A. (1986). Race, class, and the attributional process. *Journal of Educational Psychology*, 78, 4-13.
- Groteluschen, A. K., Borkowski, J. G., & Hale, C. (1990). Interdependency of executive and attributional processes. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 81-101). San Diego, CA: Academic Press, Inc.
- Groves, R. M. (1989). *Survey errors and survey costs*. New York: Wiley.
- Guttentag, R. E. (1984). The mental effort requirement of cumulative rehearsal: A developmental study. *Journal of Experimental Child Research*, 37, 92-106.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). *Multivariate data analysis with readings* (3rd ed.). New York: Macmillan.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20, 261-273.
- Hallahan, D. P., Kauffman, J., & Lloyd, J. (1985). *Introduction to learning disabilities* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Hallahan, D. P., & Reeve, R. E. (1980). Selective attention and distractibility. In B. K. Keogh (Ed.), *Advances in special education, Volume 1*. Greenwich, CT: JAI Press.
- Haller, E. P., Child, D. A., & Walberg, H. J. (1988). Can comprehension be taught? A quantitative synthesis of "metacognitive" studies. *Educational Researcher*, 17, 5-8.
- Hamill, D. D. (1990). On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities*, 23, 74-84.
- Hanushek, E. A., & Jackson, J. E. (1977). *Statistical methods for social scientists*. New York: Academic Press.
- Haring, N. G., & Bateman, B. (1977). *Teaching the learning disabled child*. Englewood Cliffs, NJ: Prentice-Hall.
- Harris, K., & Pressley, M. (1991). The nature of cognitive strategy instruction: Interactive strategy construction. *Exceptional Children*, 57, 392-403.

- Hay, I. (1994). *The formation of self-perception and the interactions between academic self-perception and academic achievement of primary school students*. Unpublished doctoral dissertation, The university of Queensland, Queensland, Australia.
- Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances*. Baltimore: The Johns Hopkins University Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Henk, W. A., & Melnick, S. A. (1992). The initial development of a scale to measure "Perception of self as reader". In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views from many perspectives* (pp. 111-117). Chicago: The National Reading Conference.
- Henk, W. A., & Melnick, S. A. (1995). The reader Self-Perception Scale (RSPS): A new tool for measuring how children feel about themselves as readers. *The Reading Teacher*, 48, 470-482.
- Hiebert, E. H., Winograd, P. N., & Danner, F. W. (1984). Children's attributions of failure and success in different aspects of reading. *Journal of Educational Psychology*, 76, 1139-1148.
- Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-970.
- Hooper, S. R., & Willis, W. G. (1989). *Learning disability subtyping: Neuropsychological foundations, conceptual models, and issues in clinical differentiation*. New York: Springer-Verlag.
- Horn, C., Bruning, R., Schraw, G., Curry, E., & Katkanant, C. (1993). Paths to success in the college classroom. *Contemporary Educational Psychology*, 18, 467-478.
- Horrex, J. (1992). *A multicomponential intervention programme for remediating secondary students' reading comprehension difficulties* (Research Affiliateship Scheme Report No. 2). Palmerston North, New Zealand: Massey University, Department of Education.
- Horwood, L. J. (1987). *Aspects of the theory and methodology of covariance structure modelling with applications to social science data*. Unpublished masters thesis, University of Canterbury, Christchurch, New Zealand.
- Houck, C. K. (1984). *Learning disabilities: Understanding concepts, characteristics and issues*. Englewood Cliffs, NJ: Prentice-Hall.
- Hoyle, R. H., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology*, 62, 429-440.
- Huba, G. J., & Bentler, P. M. (1982). On the usefulness of latent variable causal modeling in testing theories of naturally occurring events (including adolescent drug use): A rejoinder to Martin. *Journal of Personality and Social Psychology*, 43, 604-611.
- Huba, G. J., & Harlow, L. L. (1987). Robust structural equation models: Implications for developmental psychology. *Child Development*, 58, 147-166.
- Huba, G. J., Wingard, J. A., & Bentler, P. M. (1981). A comparison of two latent variable causal models for adolescent drug use. *Journal of Personality and Social Psychology*, 40, 180-193.

- Hughes, M. A., Price, R. L., & Marrs, D. W. (1986). Linking theory construction and theory testing: Models with multiple indicators of latent variables. *Academy of Management Review, 11*, 128-144.
- Hull, J. G., & Mendolia, M. (1991). Modeling the relations of attributional style, expectancies, and depression. *Journal of Personality and Social Psychology, 61*, 85-97.
- Hurford, D. P., Darrow, L. J., Edwards, T. L., Howerton, C. J., Mote, C. R., Schauf, J. D., & Coffey, P. (1993). An examination of phonemic processing abilities in children during their first-grade year. *Journal of Learning Disabilities, 26*, 167-177.
- Hurford, D. P., Schauf, J. D., Bunce, L., Blaich, T., & Moore, K. (1994). Early identification of children at risk for reading disabilities. *Journal of Learning Disabilities, 27*, 371-382.
- Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly, 14*, 299-324.
- Hynd, C. R., Alvermann, D. E., & Carter, S. (1984). Verbal report data: The effect of manipulating text familiarity and interval of reporting. *Journal of Reading Education, 10*, 1-7.
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22*, 255-278.
- James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Models, assumptions, and data*. Beverly Hills, CA: Sage.
- Johns, J. L. (1981). *Basic Reading Inventory: Pre-primer to grade eight* (2nd ed.). Dubuque, IA: Kendall-Hunt.
- Johnson, D. J. (1988). Review of research on specific reading, writing, and mathematics disorders. In J. F. Kavanagh & T. J. Truss (Eds.), *Learning disabilities: Proceedings of the national conference* (pp. 79-163). Parkton, MD: York Press.
- Johnston, P. H., & Winograd, P. N. (1985). Passive failure in reading. *Journal of Reading Behavior, 17*, 279-301.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.
- Jöreskog, K. G., & Sörbom, D. (1985). *LISREL VI: Analysis of linear structural relationships by the method of maximum likelihood*. Mooresville, ID: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd ed.). Chicago: SPSS.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software.
- Jöreskog, K. G., & Wold, H. (Eds.) (1982). *Systems under indirect observation: Causality, structure, prediction*. New York: North Holland.
- 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.

- Justice, E. M., & Weaver-McDougall, R. G. (1989). Adults' knowledge about memory: Awareness and use of memory strategies across tasks. *Journal of Educational Psychology, 81*, 214-219.
- Kaluger, G., & Kolson, C. (1978). *Reading and learning disabilities* (2nd ed.). Columbus, OH: Charles E. Merrill.
- Kamhi, A. G., & Catts, H. W. (1989). Reading disabilities: Terminology, definitions, and subtyping issues. In A. G. Kamhi & H. W. Catts (Eds.), *Reading disabilities: A developmental language perspective* (pp. 35-66). Boston, MA: College-Hill Press.
- Kaufman, N. J., Randlett, A. L., & Price, J. (1985). Awareness of the use of comprehension strategies in good and poor college readers. *Reading Psychology, 6*, 1-11.
- Kavale, K. A. (1987). Theoretical quandaries in learning disabilities. In S. Vaughn & C. Bos (Eds.), *Research in learning disabilities: Issues and future directions* (pp. 19-29). Boston: College-Hill Press.
- Kavale, K. A. (1988). The long term consequences of learning disabilities. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), *Handbook of special education research and practice: Vol. 2. Mildly handicapped conditions* (pp. 303-344). Oxford: Pergamon Press.
- Kavale, K. A. (1993). How many learning disabilities are there? A commentary on Stanovich's "Dysrationalia: A new specific learning disability". *Journal of Learning Disabilities, 26*, 520-523, 567.
- Kavale, K. A., & Forness, S. R. (1985). *The science of learning disabilities*. San Diego, CA: College-Hill Press.
- Kavale, K. A., & Forness, S. R. (1987). The far side of heterogeneity: A critical analysis of empirical subtyping research in learning disabilities. *Journal of Learning Disabilities, 20*, 374-383.
- Kavale, K. A., & Forness, S. R. (1992). History, definition, and diagnosis. In N. N. Singh, & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 3-43). New York: Springer-Verlag.
- Kavale, K. A., Forness, S. R., & Lorsbach, T. R. (1991). Definition for definitions of learning disabilities. *Learning Disability Quarterly, 14*, 257-266.
- Keller, C. E., & Hallahan, D. P. (1987). *Learning disabilities: Issues and instructional interventions*. Washington, DC: National Educational Association Professional Library.
- Keogh, B. K. (1987). A shared attribute model of learning disabilities. In S. Vaughn & C. Bos (Eds.), *Research in learning disabilities: Issues and future directions* (pp. 3-18). Boston: College Hill Press.
- Keogh, B. K. (1988a). Learning disability: Diversity in search of order. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), *Handbook of special education: Research and practice* (Vol. 2, pp. 225-252). New York: Pergamon Press.
- Keogh, B. K. (1988b). Improving services for problem learners: Rethinking and restructuring. *Journal of Learning Disabilities, 21*, 19-22.
- Keogh, B. K., Major-Kingsley, S., Omori-Gordan, H., & Reid, H. P. (1982). *A system of marker variables for the field of learning disabilities*. Syracuse, NY: Syracuse University Press.

- Kidder, L. H. (1981). *Selltiz, Wrightsman, & Cook's research methods in social relations* (4th ed.). New York: Holt, Rinehart, & Winston.
- Kirk, S., & Chalfant, J. (1984). *Developmental and academic learning disabilities*. Denver: Love.
- Kistner, J. A., Osborne, M., & LeVerrier, L. (1988). Causal attributions of learning-disabled children: Developmental patterns and relation to academic progress. *Journal of Educational Psychology*, 80, 82-89.
- Kobasigawa, A., Ransom, C. C. & Holland, C. J. (1980). Children's knowledge about skimming. *Alberta Journal of Educational Research*, 26, 169-182.
- Kuhn, D. (1992). Cognitive development. In M. H. Bornstein, & M. E. Lamb (Eds.), *Developmental Psychology: An advanced textbook* (3rd ed.) (pp 211-272). Hillsdale, NJ: Erlbaum.
- Kurtz, B. E. (1990). Cultural influences of children's cognitive and metacognitive development. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance* (pp. 177-199). New York: Springer-Verlag.
- Kurtz, B. E. (1991). Cognitive and metacognitive aspects of text processing. In G. Denhière & J. P. Rossi (Eds.), *Text and text processing* (pp. 77-103). Amsterdam: Elsevier.
- Kurtz, B. E., & Borkowski, J. G. (1984). Children's metacognition: Exploring relations among knowledge, process, and motivational variables. *Journal of Experimental Child Psychology*, 37, 335-354.
- Kurtz, B. E., Reid, M. K., Borkowski, J. G., & Cavanaugh, J. C. (1982). On the reliability and validity of children's metamemory. *Bulletin of the Psychonomic Society*, 19, 137-140.
- Larson, K., & Gerber, M. M. (1987). Effects of social metacognitive training for enhancing overt social behavior in LD and low-achieving delinquents. *Exceptional Children*, 54, 201-211.
- Larson, K. A., & Gerber, M. M. (1992). Metacognition. In N. N. Singh & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 126-169). New York: Springer-Verlag.
- Lawson, M. (1984). Being executive about metacognition. In J. R. Kirby (Ed.), *Cognitive strategies and educational performance*. Orlando: Academic Press.
- Leong, C. K. (1985). Diagnosis for learning in children with special needs. In D. D. Duane & C. K. Leong (Eds.), *Understanding learning disabilities: International and multidisciplinary views* (pp. 49-63). New York: Plenum Press.
- Leong, C. K. (1987). *Children with specific reading disabilities*. Lisse, The Netherlands: Swets & Zeitlinger.
- Leong, C. K. (1989). The locus of so-called IQ test results in reading disabilities. *Journal of Learning Disabilities*, 22, 507-512.
- Lerner, J. W. (1993). *Learning disabilities: Theories, diagnosis, and teaching strategies* (6th ed.). Boston: Houghton Mifflin.

- Lester, F. K. (1988). Reflections about mathematical problem-solving research. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 115-124). Hillsdale: Erlbaum.
- Licht, B. G. (1983). Cognitive-motivational factors that contribute to the achievement of learning disabled children. *Journal of Learning Disabilities, 16*, 483-490.
- Licht, B. G., & Kistner, J. A. (1986). Motivational problems of learning-disabled children: Individual differences and their implications for treatment. In J. K. Torgesen, & B. Y. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 225-255). Orlando, FL: Academic Press.
- Licht, B. G., Kistner, J. A., Ozkaragoz, T., Shapiro, S., & Clausen, L. (1985). Causal attributions of learning disabled children: Individual differences and their implications for persistence. *Journal of Educational Psychology, 77*, 208-216.
- Lipson, M. Y., Irwin, M., & Poth, L. (1986). The relationship between metacognitive self-reports and strategic reading behavior. In J. A. Niles & R. V. Lalik (Eds.), *Solving problems in literacy: Learners, teachers, and researchers. Thirty-fifth yearbook of the National Reading Conference* (pp. 214-221). Rochester, NY: National Reading Conference.
- Lomax, R. G. (1983). A guide to multiple-sample structural equation modeling. *Behavior Research Methods and Instrumentation, 15*, 580-584.
- Long, J. S. (1983a). *Confirmatory factor analysis*. Beverly Hills: Sage.
- Long, J. S. (1983b). *Covariance structure models: An introduction to LISREL*. Beverly Hills: Sage.
- Luchow, J. P., Crowl, T. K., & Kahn, J. P. (1985). Learned helplessness: Perceived effects of ability and effort on academic performance among EH and LD/EH children. *Journal of Learning Disabilities, 18*, 470-474.
- Lyon, G. R. (1989). IQ is irrelevant to the definition of learning disabilities: A position in search of logic and data. *Journal of Learning Disabilities, 22*, 504-506, 512.
- Lyon, G. R. (Ed.) (1994). *Frames of reference for the assessment of learning disabilities: New views on measurement issues*. Baltimore, MD: Paul H. Brookes.
- MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*, 107-120.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalizing on chance. *Psychological Bulletin, 111*, 490-504.
- Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*, 646-656.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin, 97*, 562-582.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multidimensional self-concepts: Relations with sex and academic achievement. *Journal of Educational Psychology, 77*, 581-596.

- Mateos, M., & Alonso, J. (1991). Metacognition and reading comprehension: Strategies for comprehension monitoring training. In M. Carretero, M. Pope, R. Simons, & J. I. Pozo (Eds.), *Learning instruction: European research in an international context* (Vol. 3, pp. 273-292). Oxford: Pergamon Press.
- Mather, N., & Roberts, R. (1994). Learning disabilities: A field in danger of extinction. *Learning Disabilities Research and Practice, 9*, 49-58.
- Maytham, W., Hendry, J., & Belser, J. R. (1973). *Fundamental statistics in psychology and education*. New York: McGraw-Hill.
- McCombs, B. L. (1986). The role of the self-system in self-regulated learning. *Contemporary Educational Psychology, 11*, 314-332.
- McCombs, B. L. (1988). Motivational skills training: Combining metacognitive, cognitive, and affective learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 141-170). San Diego: Academic Press.
- McCombs, B. L. (1989). Self-regulated learning and academic achievement: A phenomenological view. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 51-82). New York: Springer-Verlag.
- McKinney, J. D. (1988). Research on conceptually and empirically derived subtypes of specific learning disabilities. In M. C. Wang, M. C. Reynolds, and H. J. Walberg (Eds.), *Handbook of special education research and practice: Vol. 2. Mildly handicapped conditions* (pp. 253-282). Oxford: Pergamon Press.
- McLain, K. V. M., Gridley, B. E., & McIntosh, D. (1991). Value of a scale used to measure metacognitive reading awareness. *Journal of Educational Psychology, 85*, 81-87.
- Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology, 80*, 514-523.
- Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment and performance in mathematics. *Journal of Educational Psychology, 82*, 60-70.
- Meichenbaum, D., Burland, S., Gruson, L., & Cameron, R. (1985). Metacognitive assessment. In S. Yussen (Ed.), *The growth of reflection in children* (pp. 3-30). Wisconsin: Academic Press.
- Mercer, C. D., King-Sears, P., & Mercer, A. R. (1990). Learning disabilities definitions and criteria used by state education departments. *Learning Disabilities Quarterly, 13*, 141-152.
- Meyer, J. P. (1980). Causal attribution for success and failure: A multivariate investigation of dimensionality, formation, and consequences. *Journal of Personality and Social Psychology, 38*, 704-718.
- Meyers, J., & Lytle, S. (1986). Assessment of the learning process. *Exceptional Children, 53*, 138-144.
- Ministry of Education (1996). Education statistics of New Zealand. Wellington: Author.

- Mitchell, I. V. (Ed.) (1985). *The Ninth Mental Measurements Year Book*. The Buros Institute of Mental Measurement: University of Nebraska Press.
- Moats, L. C., & Lyon, G. R. (1993). Learning disabilities in the United States: Advocacy, science, and the future of the field. *Journal of Learning Disabilities, 26*, 282-294.
- Moore, A. D. (1995). Structural equation modeling in special education research. *Remedial and Special Education, 16*, 178-183.
- Moore, P. J. (1982). Children's metacognitive knowledge about reading: A selected review. *Educational Research, 24*, 120-128.
- Morris, R. (1993). Issues in empirical versus clinical identification of learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 73-93). Baltimore: Paul H. Brookes.
- Mulaik, S. A. (1987). Toward a conception of causality applicable to experimentation and causal modeling. *Child Development, 58*, 18-32.
- Mulaik, S. A., James, L., van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*, 430-445.
- Murphy, S. T. (1992). *On being L.D.: Perspectives and strategies of young adults*. New York: Teachers College, Columbia University.
- Neale, M. D. (1989). *Neale Analysis of Reading Ability: Manual*. Windsor: NFER-Nelson.
- New Zealand Federation of Specific Learning Disabilities (1985). *Statement on education for children with specific learning disabilities*. Christchurch: Author.
- Ng, S. (1979). *Error behavior and self-correction in reading and oral language*. Unpublished doctoral dissertation, University of Auckland, Auckland, New Zealand.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
- Nist, S. L., Simpson, M. L., & Hoglebe, M. C. (1985). The relationship between the use of study strategies and test performance. *Journal of Reading Behavior, 17*, 15-27.
- Nolen, S. B. (1988). Reasons for studying: Motivational orientation and study strategies. *Cognition and Instruction, 5*, 269-287.
- Norman, C., & Zigmund, N. (1980). Characteristics of children labeled and served as learning disabled in school systems affiliated with child service demonstration centers. *Journal of Learning Disabilities, 13*, 16-21.
- Norman, H., Sritheran, E., & Ridding, C. (1984). *Teachers' perceptions of children with special needs: A national survey of primary school teachers concerning children with special needs in regular classes*. Wellington, New Zealand: Department of Education.
- Oka, E. R. (1985). *Motivational and metacognitive characteristics of under- and overachieving children*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

- Oka, E. R., & Paris, S. G. (1987). Patterns of motivation and reading skills in underachieving children. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 115 - 145). Hillsdale, NJ: Erlbaum.
- Olson, R., Kliegl, R., Davidson, B., & Foltz, G. (1985). Individual and developmental differences in reading disability. In T. Waller (Ed.), *Reading research: Advances in theory and practice* (Vol. 4, pp. 1-64). London: Academic Press.
- O'Neill, S. P. (1992). Metacognitive strategies and reading achievement among developmental students in an urban community college. *Reading Horizons, 32*, 316-330.
- O'Sullivan, J. T., & Pressley, M. (1984). Completeness of instruction and strategy transfer. *Journal of Experimental Child Psychology, 38*, 275-288.
- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117-175.
- Palinscar, A. S., & Ransom, K. (1988). From the mystery spot to the thoughtful spot: The instruction of metacognitive strategies. *The Reading Teacher, 41*, 784-789.
- Palmer, D. J., & Goetz, E. T. (1988). Selection and use of study strategies: The role of the studier's beliefs about self and strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 41-62). San Diego: Academic Press.
- Paris, S. G. (1988). Models and metaphors of learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 299-321). San Diego: Academic Press.
- Paris, S. G. (1991). Assessment and remediation of metacognitive aspects of children's reading comprehension. *Topics in Language Disorders, 12*, 32-50.
- Paris, S. G., & Byrnes, J. P. (1989). The constructivist approach to self-regulation and learning in the classroom. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 169-200). New York: Springer-Verlag.
- Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development, 55*, 2083-2093.
- Paris, S. G., & Lindauer, B. K. (1982). The development of cognitive skills during childhood. In B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 333-349). Englewood Cliffs, NJ: Prentice-Hall.
- Paris, S. G., Lipson, M. Y., & Wixson, K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293-316.
- Paris, S. G., & Oka, E. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6*, 25-56.

- Paris, S. G., & Oka, E. R. (1989). Strategies for comprehending text and coping with reading difficulties. *Learning Disability Quarterly*, 12, 32-42.
- Paris, S. G., & Newman, R. S. (1990). Developmental aspects of self-regulated learning. *Educational Psychologist*, 25, 87-102.
- Paris, S. G., Saarnio, D. A., & Cross, D. R. (1986). A metacognitive curriculum to improve children's reading and learning. *Australian Journal of Psychology*, 38, 107-123.
- Paris, S. G., & Wasik, B. A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 609-640). New York: Longman.
- Paris, S. G., Wasik, B., & van der Westhuizen, G. (1988). Meta-metacognition: A review of research on metacognition and reading. In J. Readance & S. Baldwin (Eds.), *Dialogues in literacy research* (pp. 143-166). Chicago: National Reading Conference.
- Paris, S. G., & Winograd, P. (1990a). How metacognition can promote academic learning and instruction. In B. L. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-51). Hillsdale, NJ: Erlbaum.
- Paris, S. G., & Winograd, P. (1990b). Promoting metacognition and motivation of exceptional children. *Remedial and Special Education*, 11, 7-15.
- Parkes, K. R. (1987). Field dependence and the differentiation of neurotic syndromes. In P. Cuttance & R. Ecob (Eds.), *Structural modeling by example: Applications in educational, sociological, and behavioral research* (pp. 24-50). New York: Cambridge University Press.
- Payette, K. A., Clarizio, H. F., Phillips, S. E., & Bennett, D. E. (1995). Effects of simple and regressed discrepancy models and cutoffs on severe discrepancy determination. *Psychology in the Schools*, 32, 93-102.
- Pearl, R. (1982). LD children's attributions for success and failure: A replication with a labeled LD sample. *Learning Disability Quarterly*, 5, 173-176.
- Pearl, R., Bryan, T., & Donahue, M. (1980). Learning disabled children's attributions for success and failure. *Learning Disability Quarterly*, 3, 3-9.
- Pearl, R., Bryan, T., & Herzog, A. (1983). Learning disabled children's strategy analyses under high and low success conditions. *Learning Disability Quarterly*, 6, 67-74.
- Pereira-Laird, J. A. (1996). *Development and validation of a self-report measure of the Reading Strategy Use (RSU) scale*. Manuscript submitted for publication.
- Pereira-Laird, J. A., & Deane, F. P. (1995). *Validation of a reading self-efficacy scale for adolescents*. Manuscript submitted for publication.
- Pereira-Laird, J. A., & Deane, F. P. (1996). *Development of a behavioral measure of strategy use for adolescents: Think-Aloud Reading Strategy Use Measure (TARSUM)*. Manuscript submitted for publication.
- Perfetti, C. A. (1976). Language comprehension and the deverbalization of intelligence. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 283-292). Hillsdale, NJ: Erlbaum.

- Peterson, C., Semmel, A., von Bayer, C., Abramson, L., Metalsky, G., & Seligman, M. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research*, 6, 287-300.
- Petrauskas, R., & Rourke, B. (1979). Identification of subtypes of retarded readers. A neuropsychological multivariate approach. *Journal of Clinical Neuropsychology*, 1, 17-37.
- Pintrich, P. R. (1988). A process-oriented view of student motivation and cognition. In J. S. Stark & L. Mets (Eds.), *Improving teaching and learning through research: Vol. 57. New directions for institutional research* (pp. 55-70). San Francisco: Jossey-Bass.
- Pintrich, P. R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. In C. Ames & M. Maehr (Eds.), *Advances in motivation and achievement: Vol. 6. Motivation enhancing environments* (pp. 117-160). Greenwich, CT: JAI Press.
- Pintrich, P. R., Anderman, E. M., & Klobucar, C. (1994). Intraindividual differences in motivation and cognition in students with and without learning disabilities. *Journal of Learning Disabilities*, 27, 360-370.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 149-183). Hillsdale, NJ: Erlbaum.
- Pohl, H. J. (1981). *A descriptive analysis of the reading performance and strategies of high and low progress third form readers*. Unpublished Diploma of Education investigation, University of Auckland, Auckland, New Zealand.
- Pokay, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology*, 82, 41-50.
- Pressley, M., Borkowski, J. G., & O'Sullivan, J. (1985). Children's metamemory and the teaching of memory strategies. In D. L. Forrest-Pressley, D. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 111-153). Orlando, FL: Academic Press.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 4, pp. 89-129). New York: JAI Press.
- Pressley, M., Gaskins, I. W., Cunicelli, E. A., Burdick, N. J., Schaub-Matt, M., Lee, D. S., & Powell, N. (1991). Strategy instruction at benchmark school: A faculty interview study. *Learning Disability Quarterly*, 14, 19-48.
- Pressley, M., Goodchild, F., Fleet, J., Zajchowski, R., & Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elementary School Journal*, 89, 301-342.
- Pressley, M., & Levin, J. R. (1987). Elaborative learning strategies for the inefficient learner. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 1, pp. 175-212). Hillsdale, NJ: Erlbaum.

- Pressley, M., Woloshyn, V., Lysynchuk, L. M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review*, 2, 1-58.
- Pressley, M., Symons, S., Snyder, B. L., & Cariglia-Bull, T. (1989). Strategy instruction research comes of age. *Learning Disability Quarterly*, 12, 16-30.
- Prior, M., Sanson, A., Smart, D., & Oberklaid, F. (1995). Reading disability in an Australian community sample. *Australian Journal of Psychology*, 47, 32-37.
- Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly*, XXV, 273-295.
- Reid, M. K., & Borkowski, J. G. (1987). Causal attributions of hyperactive children: Implications for teaching strategies and self-control. *Journal of Educational Psychology*, 79, 296-307.
- Reid, N. A. (1993). *Progressive Achievement Tests of Mathematics*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., Croft, A. C., & Jackson, P. F. (1978). *Progressive Achievement Tests: Study skills*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., & Elley, W. B. (1991). *Progressive Achievement Tests of Reading*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reschly, D. J. (1987). Learning characteristics of mildly handicapped students: Implications for classification, placement, and programming. In M. C. Wang, M. C. Reynolds, and H. J. Walberg (Eds.), *Handbook of special education research and practice: Vol. 1. Learner characteristics and adaptive education* (pp. 35-58). Oxford: Pergamon Press.
- Resnick, L., & Klopfer, L. (1989). Toward the thinking curriculum: An overview. In L. Resnick & L. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 1-18). Alexandria, VA: Association for Supervision and Curriculum Development.
- Reynolds, A. R., & Walberg, H. J. (1992). A structural model of science achievement and attitude: An extension to high school. *Journal of Educational Psychology*, 84, 371-382.
- Reynolds, C. R. (1982). The problem of bias in psychological assessment. In C. R. Reynolds & T. B. Gutkin (Eds.), *Handbook of School Psychology*. New York: Wiley.
- Reynolds, C. R. (1985a). Critical measurement issues in learning disabilities. *Journal of Special Education*, 18, 451-475.
- Reynolds, C. R. (1985b). Measuring the aptitude-achievement discrepancy in learning disability diagnosis. *Remedial and Special Education*, 18, 453-476.
- Reynolds, C. R. (1992). Two key concepts in the diagnosis of learning disabilities and the habilitation of learning. *Learning Disability Quarterly*, 15, 2-12.
- Reynolds, R. E., & Shirley, L. L. (1988). The role of attention in studying and learning. In E. T. Goetz, C. E. Weinstein, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 77-100). Washington, DC: Academic Press.

- Reynolds R. E., & Wade, S. E. (1986). Thinking about thinking about thinking: Reflections on metacognition. *Harvard Educational Review*, 56, 307-317.
- Reynolds, R. E., Wade, S. E., Trathen, W., & Lapan, R. (1989). The selective attention strategy and prose learning. In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications* (pp. 159-190). New York: Springer-Verlag.
- Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *Journal of Educational Psychology*, 85, 41-48.
- Rispens, J., van Yeren, T. A., & van Duijn, G. A. (1991). The irrelevance of IQ to the definition of learning disabilities: Some empirical evidence. *Journal of Learning Disabilities*, 24, 434-438.
- Rogers, C. G. (1987) Attribution theory and motivation in school. In N. Hastings, & J. Schwieso (Eds.), *New directions in educational psychology: Vol. 2. Behavior and motivation in the classroom* (pp. 195-212). London: Falmer Press.
- Rogers, H., & Saklofske, D. H. (1985). Self-concepts, locus of control, and performance expectations of learning disabled children. *Journal of Learning Disabilities*, 18, 273-278.
- Ronis, D. L., Hansen, R. D., & O'Leary, V. E. (1983). Understanding the meaning of achievement attributions: A test of derived locus and stability scores. *Journal of Personality and Social Psychology*, 44, 702-711.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80, 1-28.
- Rottman, T. R., & Cross, D. R. (1990). Using informed strategies for learning to enhance the reading and thinking skills of children with learning disabilities. *Journal of Learning Disabilities*, 23, 270-278.
- Rourke, B. P. (1983). Outstanding issues in research on learning disabilities. In M. Rutter (Ed.), *Developmental Neuropsychiatry* (pp. 564-574). New York: Guilford Press.
- Russell, D. (1982). The causal dimension scale: A measure of how individuals perceive causes. *Journal of Personality and Social Psychology*, 44, 702-711.
- Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry*, 16, 181-197.
- Saarnio, D. A., Oka, E. R., & Paris, S. G. (1990). Developmental predictors of children's reading comprehension. In T. H. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 57-79). San Diego, CA: Academic Press.
- Sattler, J. M. (1992). *Assessment of children* (3rd ed.). San Diego: Author.
- Schmitt, M. C. (1990). A questionnaire to measure children's awareness of strategic reading processes. *The Reading Teacher*, 43, 454-461.
- Schneider, W. (1985). Developmental trends in the metamemory-memory behavior relationship: An integrative review. In D. L. Forrest-Pressley, G. E. MacKinnon & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 57-109). Orlando, FL: Academic Press.

- Schneider, W., & Pressley, M. (1989). *Memory development between 2 and 20*. New York: Springer-Verlag
- Schneider, W., & Weinert, F. E. (1990). The role of knowledge, strategies, and aptitudes in cognitive performance: Concluding comments. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance* (pp. 286-302). New York: Springer-Verlag.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemporary Educational Psychology, 19*, 143-154.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.
- Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology, 75*, 848-856.
- Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist, 19*, 48-58.
- Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools, 22*, 208-223.
- Schunk, D. H. (1986). Verbalization and children's self-regulated learning. *Contemporary Educational Psychology, 11*, 347-369.
- Schunk, D. H. (1989a). Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 83-110). New York: Springer-Verlag.
- Schunk, D. H. (1989b). Self-efficacy and achievement behaviors. *Educational Psychology Review, 1*, 173-208.
- Schunk, D. H. (1989c). Self-efficacy and cognitive skill learning. In C. Ames & R. Ames (Eds.), *Research on motivation in education: Vol. 3. Goals and cognitions* (pp. 13-44). San Diego, CA: Academic Press.
- Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*, 71-86.
- Schunk, D. H. (1991). Self-efficacy and achievement motivation. *Educational Psychologist, 26*, 207-231.
- Scruggs, T. E., & Mastropieri, M. A. (1986). Improving the test taking skills of behaviorally disordered and learning disabled children. *Exceptional Children, 53*, 63-68.
- Seegers, G., & Boekaerts, M. (1993). Task motivation and mathematics achievement in actual task situations. *Learning and Instruction, 3*, 133-150.
- Senf, G. M. (1986). LD research in sociological and scientific perspective. In J. K. Torgesen & B. Y. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 27-53). Orlando, FL: Academic Press.

- Senf, G. M. (1987). Learning disabilities as sociologic sponge: Wiping up life's spills. In S. Vaughn & C. Bos (Eds.), *Research in learning disabilities: Issues and future directions* (pp. 87-96). Boston: College-Hill Press.
- Share, D. L., McGee, R., McKenzie, D., Williams, S., & Silva, P. A. (1987). Further evidence relating to the distinction between specific reading retardation and general reading backwardness. *British Journal of Developmental Psychology*, *5*, 35-44.
- Share, D. L., 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., Moffitt, T. E., & Silva, P. A. (1988). Factors associated with arithmetic-and-reading disability and specific arithmetic disability. *Journal of Learning Disabilities*, *21*, 313-320.
- Share, D. L., & Silva, P. A. (1987). Language deficits and specific reading retardation: Cause or effect? *British Journal of Disorders of Communication*, *22*, 219-226.
- Shaw, S. F., Cullen, J. P., McGuire, J. M., & Brinckerhoff, L. C. (1995). Operationalizing a definition of learning disabilities. *Journal of Learning Disabilities*, *28*, 586-597.
- Shaywitz, B. A., Fletcher, J. M., Holahan, J. M., & Shaywitz, S. E. (1992). Discrepancy compared to low achievement definitions of reading disability: Results from the Connecticut longitudinal study. *Journal of Learning Disabilities*, *25*, 639-648.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, *326*, 145-150.
- Shaywitz, S. E., & Shaywitz, B. A. (1988). Attention-deficit disorder: Current perspectives. In J. Kavanagh & J. Truss (Eds.), *Learning disabilities: Proceedings of the National Conference* (pp. 369-567). Parkton, MD: York Press.
- Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology*, *81*, 91-100.
- Shepard, L. (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. *Journal of Special Education*, *14*, 79-91.
- Shinn, M. R., Tindal, G. A., Spira, D., & Marston, D. (1987). Practice of learning disabilities as social policy. *Learning Disability Quarterly*, *10*, 17-28.
- Short, E. J., Feagans, L., McKinney, J. D., & Applebaum, M. I. (1984). Longitudinal stability of LD subtypes based on age- and IQ-achievement discrepancies. *Learning Disability Quarterly*, *9*, 214-225.
- Short, E. J., & Ryan, E. B. (1984). Metacognitive differences between skilled and less skilled readers: Remediating deficits through story grammar and attributional training. *Journal of Educational Psychology*, *76*, 225-235.
- Short, E. J., Schatschneider, C. W., Cuddy, C. L., Evans, S. W., Dellick, D. M., & Basili, L. A. (1991). The effects of thinking aloud on the problem-solving performance of bright, average learning disabled, and developmentally handicapped students. *Contemporary Educational Psychology*, *16*, 139-153.

- Short, E. J., Schatschneider, C. W., & Friebert, S. E. (1992). The inactive learner hypothesis: Myth or reality? In N. N. Singh, & I. L. Beale (Eds.), *Learning disabilities: Nature, theory, and treatment* (pp. 302-326). New York: Springer-Verlag.
- Short, E. J., & Weissberg-Benchell, J. (1989). The triple alliance for learning: Cognition, metacognition, and motivation. In C. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: Implications for the curriculum* (pp. 33-63). New York: Springer-Verlag.
- Shuell, T. (1988). The role of the student in learning from instruction. *Contemporary Educational Psychology, 13*, 276-295.
- Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42*, 201-215.
- Siegel, L. S. (1989a). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*, 469-478, 486.
- Siegel, L. S. (1989b). Why we do not need IQ test scores in the definition and analyses of learning disability. *Journal of Learning Disabilities, 22*, 514-518.
- Siegel, L. S. (1990). IQ and learning disabilities: R.I.P. In H. Swanson, & B. Keogh (Eds.), *Learning disabilities: Theoretical and research issues* (pp. 111-128). Hillsdale, NJ: Erlbaum.
- Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25*, 618-629.
- Silva, P. A., McGee, R., & Williams, S. (1985). Some characteristics of nine year old boys with general reading backwardness or specific reading retardation. *Journal of Child Psychology and Psychiatry, 20*, 407-421.
- Silver, L. B. (1988). A review of the federal government's Interagency Committee of Learning Disabilities report to the U.S. Congress. *Learning Disabilities Focus, 3*, 73-80.
- Slife, B. D., & Weaver, C. A. (1992). Depression, cognitive skill, and metacognitive skill in problem solving. *Cognition and Emotion, 6*, 1-22.
- Slife, B. D., Weiss, J., & Bell, T. (1985). Separability of metacognition and cognition: Problem solving in learning disabled and regular students. *Journal of Educational Psychology, 77*, 437-445.
- Spedding, S., & Chan, L. K. S. (1993). Metacognition, word identification, and reading competence. *Contemporary Educational Psychology, 18*, 91-100.
- Speece, D., McKinney, J., & Applebaum, M. (1985). Classification and validation of behavioral subtypes of learning disabled children. *Journal of Educational Psychology, 77*, 67-77.
- Spreen, O., & Haaf, R. (1986). Empirically derived learning disability subtypes: A replication attempt and longitudinal patterns over 15 years. *Journal of Learning Disabilities, 19*, 170-181.
- Spring, C., & French, L. (1990). Identifying children with specific reading disabilities from listening and reading discrepancy scores. *Journal of Learning Disabilities, 23*, 53-58.

- St. George, A., van Kraayenoord, C. E., & Chapman, J. W. (1985). *Cognition and reading of learning disabled and normally achieving intermediate school children* (Research contract No. 55-22-121). Wellington, New Zealand: Massey University, Department of Education.
- St. George, R., & Chapman, J. W. (1987). WISC-R item characteristics: A study with 11-year old New Zealand children. *New Zealand Journal of Psychology*, *16*, 28-36.
- Stacy, A. W., Widaman, K. F., Hays, R., & DiMatteo, M. R. (1985). Validity of self-reports of alcohol and other drug use: A multitrait-multimethod assessment. *Journal of Personality and Social Psychology*, *49*, 219-232.
- Stanovich, K. E. (1986a). Cognitive processes and the reading problems of learning-disabled children: Evaluating the assumption of specificity. In J. K. Torgesen & B. Y. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 87-131). New York: Academic Press.
- Stanovich, K. E. (1986b). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360-407.
- Stanovich, K. E. (1986c). Commentary: New beginnings, old problems. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 1, pp. 229-238). Hillsdale, NJ: Erlbaum.
- Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*, 590-605.
- Stanovich, K. E. (1989). Has the learning disabilities field lost its intelligence? *Journal of Learning Disabilities*, *22*, 487-492.
- Stanovich, K. E. (1991a). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*, *26*, 1-29.
- Stanovich, K. E. (1991b). Conceptual and empirical problems with discrepancy definitions of reading disability. *Learning Disability Quarterly*, *14*, 269-280.
- Stanovich, K. E. (1993a). A model for studies of reading disability. *Developmental Review*, *13*, 225-245.
- Stanovich, K. E. (1993b). The construct validity of discrepancy definitions of reading disability. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 273-307). Baltimore: Paul H. Brookes.
- Stanovich, K. E., Cunningham, A., & Feeman, D. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly*, *19*, 278-303.
- Sternberg, R. (1985). Review of Meichenbaum, Burland, Gruson, and Cameron's "Metacognitive Assessment". In S. Yussen (Ed.), *The growth of reflection in children* (pp. 31-35). Wisconsin: Academic Press.
- Sternthal, B., Tybout, A. M., & Calder, B. J. (1987). Confirmatory versus comparative approaches to judging theory tests. *Journal of Consumer Research*, *14*, 114-125.

- Strawser, S. (1993). Assessment and identification practices. In W. N. Bender (Ed.), *Learning disabilities: Best practices for professionals* (pp. 71-103). Boston: Andover Medical.
- Swanson, H. L. (1984). Effects of cognitive effort and word distinctiveness of learning disabled and nondisabled readers' recall. *Journal of Educational Psychology, 76*, 894-908.
- Swanson, H. L. (1988). Toward a metatheory of learning disabilities. *Journal of Learning Disabilities, 21*, 196-209.
- Swanson, H. L. (1989). Strategy instruction: Overview of principles and procedures for effective use. *Learning Disability Quarterly, 12*, 3-14.
- Swanson, H. L. (1990). Intelligence and learning disabilities. In H. L. Swanson & B. K. Keogh (Eds.), *Learning disabilities and research issues* (pp. 97-113). Hillsdale, NJ: Erlbaum.
- Swanson, H. L. (1991a). Introduction: Issues in the assessment of learning disabilities. In H. L. Swanson (Ed.), *Handbook on the assessment of learning disabilities: Theory, research, and practice* (pp. 1-19). Austin, TX: Pro-Ed.
- Swanson, H. L. (1991b). Operational definitions and learning disabilities: An overview. *Learning Disability Quarterly, 14*, 242-254.
- Swanson, H. L. (1993). Learning disabilities from the perspective of cognitive psychology. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 199-228). Baltimore: Paul H. Brookes.
- Swanson, H. L., & Cooney, J. (1985). Strategy transformations in learning disabled children. *Learning Disability Quarterly, 8*, 221-231.
- Swanson, H. L., & Keogh, B. (1990). *Learning disabilities: Theoretical and research issues*. Hillsdale, NJ: Erlbaum.
- Swanson, H. L., & Ramalgia, J. M. (1992). The relationship between phonological codes on memory and spelling tasks for students with and without learning disabilities. *Journal of Learning Disabilities, 25*, 396-407.
- Swanson, H. L., & Rhine, B. (1985). Strategy transformation in learning disabled children's math performance: Clues to the development of expertise. *Journal of Learning Disabilities, 18*, 596-603.
- Swanson, H. L., & Trahan, M. (1995). *Learning disabled and average readers' working memory and comprehension: Do word knowledge and metacognition play a role?* Unpublished manuscript.
- Sweeney, P. D., Anderson, K., & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology, 50*, 974-991.
- Symons, S., Snyder, B. L., Cariglia-Bull, T., & Pressley, M. (1989). Why be optimistic about cognitive strategy instruction? In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications* (pp. 3-32). New York: Springer-Verlag.
- Szuskiewicz, T., & Symons, D. (1993). Letters to the editor: Comment on Rosenberg et al. (1993). *Journal of Learning Disabilities, 26*, 570-571.

- Taylor, R. L. (1984). *Assessment of exceptional students: Educational and psychological procedures*. Englewood Cliffs, NJ: Prentice-Hall.
- Taylor, R. L., & Sternberg, L. (1989). *Exceptional children: Integrating research and teaching*. New York: Springer-Verlag.
- Taylor, S. E., & Fiske, S. T. (1981). Getting inside the head: Methodologies for process analysis in attribution and social cognition. In J. H. Harvey, W. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 3, pp. 459-524). Hillsdale, NJ: Erlbaum.
- Tellegen, A., & Briggs, P. F. (1967). Old wine in new skins: Grouping Wechsler subtests into new scales. *Journal of Consulting Psychology*, 31, 499-506.
- Thomas, A. (1979). Learned helplessness and expectancy factors: Implications for research in learning disabilities. *Review of Educational Research*, 49, 208-221.
- Thomas, C. C., Englert, C. S., & Gregg, S. (1987). An analysis of errors and strategies in the expository writing of learning disabled students. *Remedial and Special Education*, 8, 21-30.
- Tindal, G., & Marston, D. (1986). Approaches to assessment. In J. K. Torgesen & B. Y. L. Wong (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 55-84). Orlando, FL: Academic Press.
- Torgesen, J. K. (1979). Factors related to poor performance on rote memory tasks in reading disabled children. *Learning Disabilities Quarterly*, 2, 17-23.
- Torgesen, J. K. (1980). Conceptual and educational implications of the use of efficient task strategies by learning disabled children. *Journal of Learning Disabilities*, 13, 364-371.
- Torgesen, J. K. (1982a). The use of rationally defined subgroups in research on learning disabilities. In J. P. Das, R. F. Mulcahy, & A. E. Wall (Eds.), *Theory and Research in Learning Disabilities*. New York: Plenum Press.
- Torgesen, J. K. (1982b). The learning-disabled child as an inactive learner: Educational implications. *Topics in Learning and Learning Disabilities*, 2, 45-52.
- Torgesen, J. K. (1986). Learning disabilities theory: Its current state and future prospects. *Journal of Learning Disabilities*, 19, 399-407.
- Torgesen, J. K. (1990). Concluding comments. In J. K. Torgesen (Ed.), *Cognitive and behavioral characteristics of children with learning disabilities*. Austin, TX: Pro-Ed.
- Torgesen, J. K. (1991). Learning disabilities: Historical and conceptual issues. In B. Y. L. Wong (Ed.), *Learning about learning disabilities* (pp. 3-37). San Diego: Academic Press.
- Torgesen, J. K. (1993). Variations on theory in learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. Krasnegor (Eds.), *Better understanding learning disabilities: New views from research and their implications for education and public policies* (pp. 153-170). Baltimore: Paul H. Brookes.
- Torgesen, J. K. (1994). Learning disabilities theory: Issues and advances. In S. Vaughn & C. Bos (Eds.), *Research issues in learning disabilities: Theory, methodology, assessment, and ethics* (pp. 3-21). New York: Springer-Verlag.

- Torgesen, J. K., & Dice, C. (1980). Characteristics of research in learning disabilities. *Journal of Learning Disabilities, 13*, 531-535.
- Torgesen, J. K., & Licht, B. (1983). The learning disabled child as an inactive learner: Retrospect and prospects. In J. D. McKinney & L. Feagans (Eds.), *Topics in Learning Disabilities* (Vol. 1, pp. 3-32). Rockville, MD: Aspen Press.
- Torgesen, J. K., Murphy, H., & Ivey, C. (1979). The effects of an orienting task on the memory performance of reading disabled children. *Journal of Learning Disabilities, 12*, 396-401.
- Torgesen, J. K., & Wong, B. Y. L. (1986). *Psychological and educational perspectives on learning disabilities*. New York: Academic Press.
- Tuck, B. F., Hanson, A. L., & Zimmerman, M. (1975). The WISC-R: A New Zealand Study of norms and validity. *New Zealand Journal of Educational Studies, 10*, 52-58.
- Tunmer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read. In P. B. Gough, L. C. Ehri, & R. C. Treiman (Eds.), *Reading acquisition* (pp. 175-214). Hillsdale, NJ: Erlbaum.
- Valencia, S. W., & Pearson, P. D. (1987). Reading assessment: Time for a change. *The Reading Teacher, 40*, 726-732.
- Valus, A. (1986). Teacher perceptions of identification criteria emphasized in initial learning disabilities placement. *Learning Disabilities Research, 2*, 21-25.
- van den Bos, K. P. (1989). Relationship between cognitive, decoding skill, and reading comprehension in learning disabled children. In P. G. Aaron & R. M. Joshi (Eds.), *Reading and writing disorders in different orthographic systems*. Dordrecht, The Netherlands: Kluwer Academic..
- van der Neut, L. (1992). *A survey of services available in Auckland to individuals with learning disabilities*. Unpublished masters thesis, University of Auckland, Auckland, New Zealand.
- van der Wissel, A., & Zegers, F. E. (1985). Reading retardation revisited. *British Journal of Developmental Psychology, 3*, 3-9.
- van Kraayenoord, C. E. (1986). *Metacognition, reading and causal attributions: A comparison of learning disabled and non learning disabled intermediate school children*. Unpublished doctoral dissertation, Massey University, Palmerston North, New Zealand.
- van Kraayenoord, C. E. & Elkins, J. (1994). Learning difficulties in regular classrooms. In A. Ashman & J. Elkins (Eds.), *Educating children with special needs* (2nd ed.) (pp. 233-289). New York: Prentice Hall.
- Wade, S. E., & Trathen, W. (1989). Effect of self-selected study methods on learning. *Journal of Educational Psychology, 81*, 40-47.
- Wade, S. E., Trathen, W., & Schraw, G. (1990). An analysis of spontaneous study strategies. *Research Reading Quarterly, 25*, 147-166.
- Wagner, D. A., Spratt, J. E., Gal, I., & Paris, S. G. (1989). Reading and believing: Beliefs, attributions, and reading achievement in Moroccan school children. *Journal of Educational Psychology, 81*, 283-293.

- Wagoner, S. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, 28, 328-346.
- Walsh, D. E. (1979). *Project CHILD: A report on the incidence of specific learning difficulties in Form 1 pupils in New Zealand primary and intermediate schools*. Wellington, New Zealand: New Zealand Council for Research.
- Wang, M. C., & Peverly, S. T. (1986). The self-instructive process in classroom learning contexts. *Contemporary Educational Psychology*, 11, 370-404.
- Watkins, D. (1984). Student learning processes: An exploratory study in the Philippines. *Human Learning*, 3, 33-42.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children - Revised*. New York: The Psychological Corporation.
- Weller, C. (1987). A multifaceted hierarchical theory of learning disabilities. In S. Vaughn & C. Bos (Eds.), *Research in learning disabilities: Issues and future directions* (pp. 35-46). Boston: College-Hill Press.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71, 3-25.
- Weiner, B. (1983). Some methodological pitfalls in attribution research. *Journal of Educational Psychology*, 75, 530-543.
- Weiner, B. (1985a). *Human motivation*. New York: Springer-Verlag.
- Weiner, B. (1985b). An attributional theory of achievement motivation. *Psychological Bulletin*, 92, 548-573.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer Verlag.
- Weinstein, C. E. (1988). Assessment and training of student learning strategies. In R. R. Schmeck (Ed.), *Learning styles and learning strategies* (pp. 291-316). New York: Plenum.
- Weinstein, C. E., Goetz, E. T., & Alexander, P. A. (Eds.) (1988). *Learning and study strategies: Issues in assessment, instruction, and evaluation*. San Diego: Academic Press.
- Weinstein, C. E., Hagen, A. S., & Meyer, D. K. (1991, April). *Work smart...not hard: The effects of combining instruction in using strategies, goal using, and executive control on attributions and academic performance*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315-327). New York: MacMillan.
- Wigfield, A., & Asher, S. R. (1984). Social and motivation influences on reading. In D. Pearson (Ed.), *Handbook of reading research* (pp. 423-452). New York: Longman.
- Williams, L. J., & Hazer, J. T. (1986). Antecedents and consequences of organizational turnover: A reanalysis using a structural equations model. *Journal of Applied Psychology*, 71, 219-231.

- Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, 49, 287-293.
- Winograd, P. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 19, 404-425.
- Winograd, P., & Hare, V. C. (1988). Direct instruction of reading comprehension strategies: The nature of teacher explanation. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 121-139). New York: Academic Press.
- Winograd, P., & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of Reading Behavior*, 14, 61-76.
- Winograd, P., Witte, R., & Smith, L. (1986). Measuring attributions for reading. A comparison of direct and indirect locus and stability scores. In J. A. Niles, & R. Lalik (Eds.), *Solving problems in literacy: Learners, teachers, and researchers. Thirty-fifth Yearbook of the National Reading Conference* (pp. 387-394). Rochester, NY: National Reading Conference.
- Wixson, K. K., & Peters, C. W. (1984). Reading redefined: A Michigan Reading Association position paper. *Michigan Reading Journal*, 17, 4-7.
- Wong, B. Y. L. (1980). Activating the inactive learner: Use of questions/prompts to enhance comprehension and retention of implied information in learning disabled children. *Learning Disability Quarterly*, 3, 29-37.
- Wong, B. Y. L. (1985). Metacognition and learning disabilities. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 2. Instructional practices* (pp. 137-180). Orlando, FL: Academic Press.
- Wong, B. Y. L. (1987). Metacognition and special education: A review of a view. *The Journal of Special Education*, 20, 9-29.
- Wong, B. Y. L. (1991). The relevance of metacognition to learning disabilities. In B. Y. L. Wong (Ed.), *Learning about learning disabilities* (pp. 231-258). San Diego, CA: Academic Press.
- Yi, Y., & Nassen, K. (1992). *Multiple comparison and cross-validation in evaluating structural equation models* (pp. 407-411). In Academy of Marketing Science (Ed.), *Developments in Marketing Science, Vol. 15*. Crittenden, MA: Author.
- Yussen, S. R. (1985). The role of metacognition in theories of cognitive development. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 253-283). Orlando, FL: Academic Press.
- Zabrocky, K., & Commander, N. E. (1993). Rereading to understand: The role of text coherence and reader proficiency. *Contemporary Educational Psychology*, 18, 442-454.
- Zabrocky, K., & Ratner, H. H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development*, 57, 1401-1418.
- Zimmerman, B. J. (1989a). Models of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theory, research, and practice* (pp. 1-25). New York: Springer-Verlag.

- Zimmerman, B. J. (1989b). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*, 329-339.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: The emergence of a social cognitive perspective. *Educational Psychology Review, 2*, 173-201.
- Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*, 614-628.
- Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284-290.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.
- Zimmerman, B. J., & Martinez-Pons, M. (1992). Perceptions of efficacy and strategy use in the self-regulation of learning. In D. H. Schunk & J. Meece (Eds.), *Student perceptions in the classroom: Causes and consequences* (pp. 185-207). Hillsdale, NJ: Erlbaum.
- Zimmerman, B. J., & Schunk, D. H. (Eds.) (1989). *Self-regulated learning and academic achievement: Theory, research, and practice*. New York: Springer-Verlag.

APPENDIX A
INFORMATION SHEET/CONSENT FORMS

[MASSEY LOGO]

RESEARCH PROJECT ON READING SKILLS
INFORMATION SHEET

What is the study about?

The basic aim of this research is to investigate students' reading skills and their motivation to use them. We believe that research of this type is extremely valuable, in that it improves our understanding of the influence of children's motivation and use of reading skills on academic achievement. The project has the support of both the principal, Mr/Mrs _____ of _____ school/college and the Massey University Human Ethics Committee. It is being run by Mrs. Joyce Laird in consultation with Lecturers, Dr. Frank Deane and Dr. Julie Bunnell of Massey University (Psychology Department), and is funded by Massey University.

What would your child have to do?

As we do not have the resources to include every Form 2 student in the school in all stages of this research, some selection has to take place to ensure that we include students with a variety of reading skills in the main study. If you are willing to allow your child to participate in this project and if your child is willing, then she/he will be required initially for a period of 30-60 minutes (depending on the child).

Should we need to include your child in the main study, she/he will be required to fill out two questionnaires and have a short interview. The questionnaires are quite straight forward and easy to fill out. In total, your child will be required for two 50 min. sessions to fill out the questionnaires and one 30 min. session for a short interview. Every effort will be made to ensure that this does not interfere too much with normal school work activities.

What can you and your child expect from the researchers?

As a parent/guardian of the participant:

1. You have the right to withdraw your child from the study at any time.
2. Your child has the right to withdraw from the study at any time.
3. The information obtained is confidential to the researchers. The questionnaires are seen only by the researchers, and the participant will be identified only by a code number. It will not be possible to identify individuals in any published reports.
4. You will expect to be informed of the results of the study.
5. You will expect the principal and the teachers to be given a summary of the findings of the study. This will provide them with useful information on how students learn reading.

If you have any questions at all about this study, please call Joyce Laird at Massey University, Palmerston North (06) 3569099, extension 7922, or 3562045.

PARENT/GUARDIAN CONSENT FORM

Please fill in the boxes and sign below to show whether you are willing to allow your child to participate in this study.

(Please tick the boxes if you agree.)

I agree to my child participating in the preliminary assessment.

I also agree to my child participating in the main part of the study,
if this is required.

Student's name: _____

Parent's/Guardian's name: _____

Parent's/Guardian's signature: _____

Date: _____

- Notes:**
1. Please keep the information sheet for your reference.
 2. Students will be asked to sign a separate consent form before being included in the study. Could you please ask your child to fill out the Student Consent Form, found on the next page.
 3. Please ask your child to hand the signed consent forms to his/her class or group teacher.

STUDENT CONSENT FORM

Dear Student,

My name is Joyce Laird and I am conducting some research at Massey University on students' feelings about reading and their use of reading skills. This is very important research because the findings will greatly improve our understanding of the way reading skills helps learning.

I need the help of students in your age group to carry out this research and would be very grateful if you would be prepared to take part.

If you are prepared to take part, you will

1. Have the right to withdraw from the study at any time.
2. Provide information on the understanding that it is confidential. The questionnaires are seen only by the researchers, and you will be identified only by a code number.
3. Have the opportunity to receive information about the findings of the study.

Please fill in the boxes and sign below to show whether you are willing to participate in this study.

.....
(Please cut along dotted line and return the bottom half)

(Please tick the boxes if you agree.)

I agree to participate in the preliminary assessment.

I agree to participate in the main part of the study, if this is required.

Student's name: _____

Student's signature: _____

Date: _____

Class: _____

School: _____

APPENDIX B

- B.1 READING QUESTIONNAIRE I
- B.2 READING QUESTIONNAIRE II
- B.3 TARSUM: MEASURES, SCORING PROCEDURES, AND DEFINITIONS OF STRATEGIES
 - B3.1 PASSAGES
 - B.3.2 INTERVIEW CODING FORM
 - B.3.3 SCORING PROCEDURES FOR METACOGNITIVE STRATEGIES
 - B3.4 DEFINITIONS OF STRATEGIES MEASURED IN THE TARSUM

B.1 READING QUESTIONNAIRE I

Please answer the following before you begin the questionnaire.

NAME:

--	--	--

DATE OF BIRTH:

--	--	--

SCHOOL:

--	--

CLASS:

--	--

MALE FEMALE (circle)

--

Do you consider yourself: MAORI EUROPEAN POLYNESIAN
ASIAN OTHER (circle)

--

If you have circled OTHER, please specify: _____

What is your first language: _____

--

--

Introduction

This questionnaire measures your learning skills and motivation for **READING**. The questionnaire is arranged in three separate sections, each with a different set of instructions. Be sure to read these instructions carefully. **THERE ARE NO RIGHT OR WRONG ANSWERS TO THIS QUESTIONNAIRE.** Your data will be treated as confidential. Nobody will see your answers apart from the researcher.

--	--	--

--	--

--	--

--	--

--	--

SECTION 1

Directions

In this section of the questionnaire, you will find a number of statements about your reading skills and your feelings about reading. For each of these statements, decide how well the statement describes you, and not in terms of how you think you should be or what others do. For each statement, use the scale from 1 (NEVER TRUE) to 7 (ALWAYS TRUE), and circle the number that best describes you. As there are no "right" or "wrong" answers, please be absolutely honest. Do not leave any statements out; make a response to every one.

To help you decide which number to circle, we would like to explain what is meant by each term.

- 1 – NEVER TRUE
- 2 – ALMOST NEVER TRUE
- 3 – SELDOM TRUE
- 4 – SOMETIMES TRUE
- 5 – OFTEN TRUE
- 6 – ALMOST ALWAYS TRUE
- 7 – ALWAYS TRUE

EXAMPLE: (Circle the number that best describes you)

	never	almost never	seldom	sometimes	often	almost always	always
1. Personal problems interfere with my studying.	1	2	3	4	5	6	7
2. I really like spinach.	1	2	3	4	5	6	7

Now turn over to the next page and begin.

SECTION 1 STARTS HERE

Read each statement carefully and answer every item, even if it is hard to decide which answer is most like you. Just circle one number for each statement.

	How true these statements are for me							office use
	never	almost never	seldom	sometimes	often	almost always	always	
1. To help me understand what I have read, I say it in my own words.	1	2	3	4	5	6	7	<input type="checkbox"/>
2. When I read, I form pictures in my mind of the things I am trying to understand.	1	2	3	4	5	6	7	<input type="checkbox"/>
3. I am not sure I will be able to understand most things I read.	1	2	3	4	5	6	7	<input type="checkbox"/>
4. After reading something, I sit and think about it for a while to check my understanding.	1	2	3	4	5	6	7	<input type="checkbox"/>
5. When reading I learn by heart difficult words and ideas without understanding them.	1	2	3	4	5	6	7	<input type="checkbox"/>
6. When some of the sentences that I am reading are hard I either give up or read only the easy words.	1	2	3	4	5	6	7	<input type="checkbox"/>
7. I read quickly through the whole passage to get the general idea before I read it thoroughly.	1	2	3	4	5	6	7	<input type="checkbox"/>
8. When reading about something I try to link it to what I already know.	1	2	3	4	5	6	7	<input type="checkbox"/>
9. I am sure I can do an excellent job on the reading tasks given to me in class.	1	2	3	4	5	6	7	<input type="checkbox"/>
10. I learn new words by picturing in my mind a situation in which they occur.	1	2	3	4	5	6	7	<input type="checkbox"/>
11. When I find that a chapter in my book is hard to understand, I slow down my reading.	1	2	3	4	5	6	7	<input type="checkbox"/>
12. I am sure I will receive a good mark for my reading comprehension test.	1	2	3	4	5	6	7	<input type="checkbox"/>
13. When reading I check how well I understand the meaning of the story by asking myself whether the ideas fit with the other information in the story.	1	2	3	4	5	6	7	<input type="checkbox"/>

How true these statements are for me

never almost never seldom sometimes often almost always always

office use

14. I don't feel confident in doing well in reading.	1	2	3	4	5	6	7	<input type="checkbox"/>	39
15. I am able to decide between more important and less important information while reading.	1	2	3	4	5	6	7	<input type="checkbox"/>	
16. When I read, I underline the main ideas.	1	2	3	4	5	6	7	<input type="checkbox"/>	
17. I learn new words by relating/linking them with words which I already know.	1	2	3	4	5	6	7	<input type="checkbox"/>	
18. I am sure that I will be able to understand most difficult words I come across.	1	2	3	4	5	6	7	<input type="checkbox"/>	
19. When I'm reading I stop once in a while and go over what I have read.	1	2	3	4	5	6	7	<input type="checkbox"/>	
20. I am certain I can understand well the books that I read.	1	2	3	4	5	6	7	<input type="checkbox"/>	
21. I decide how difficult my reading passage is and then adjust the speed of my reading accordingly.	1	2	3	4	5	6	7	<input type="checkbox"/>	
22. I stop once in a while and ask myself questions to see how well I understand what I am reading.	1	2	3	4	5	6	7	<input type="checkbox"/>	47
23. I make an outline of what I am reading.	1	2	3	4	5	6	7	<input type="checkbox"/>	
24. When I cannot read a word in the story, I skip it.	1	2	3	4	5	6	7	<input type="checkbox"/>	
25. When I get lost while reading, I go back to the point where I first had trouble.	1	2	3	4	5	6	7	<input type="checkbox"/>	
26. I feel confident that I can read anything that is given to me by my teacher.	1	2	3	4	5	6	7	<input type="checkbox"/>	
27. After I have been reading for a short time, the words stop making sense.	1	2	3	4	5	6	7	<input type="checkbox"/>	
28. I read critically or thoughtfully, that is while reading something, I judge what I am reading.	1	2	3	4	5	6	7	<input type="checkbox"/>	
29. When I find I do not understand something while reading I read it again and try to figure it out.	1	2	3	4	5	6	7	<input type="checkbox"/>	54

SECTION 2

In this section of the questionnaire there are 25 multiple choice questions. There are three parts to this section. I want you to think about what kinds of things you can do to help you understand a story *before*, *during*, and *after* you read it.

Four statements are given for each question. Read each of the four statements and decide which one of them would help you the most. There are no right answers. It is just what you think would help the most. Put a circle around the letter in front of the **BEST** answer to the question. You are asked to choose only **ONE** answer for each question. Work as quickly and as carefully as you can. Answer every question even when you are not sure of your choice. Do not spend too much time on questions you find hard.

Each of the questions in this section is set out like the example below.
We will go over an example before you begin.

EXAMPLE: (Circle the letter in front of the best answer)

I should do my school homework

- A. Because my teacher will scold me if I don't.
 - B. Because it will help me learn more about the subject.
 - C. Because it helps me grow up.
 - D. Because it will get me good marks.
-

Please turn over the next page for section 2

SECTION 2 STARTS HERE

I. In each set of the four choose the one statement which tells a good thing to do to help you understand a story better before you read it.

1. Before I begin reading, it's a good idea to:

- A. See how many pages are in the story.
- B. Look up all of the big words in the dictionary
- C. Make some guesses about what I think will happen in the story.
- D. Think about what has happened so far in the story.

office use

 55

2. Before I begin reading, it's a good idea to:

- A. Look at the pictures to see what the story is about.
- B. Decide how long it will take me to read the story.
- C. Sound out the words I don't know
- D. Check to see if the story is making sense.

3. Before I begin reading, it's a good idea to:

- A. Ask someone to read the story to me.
- B. Read the title to see what the story is about.
- C. Check to see if most of the words have long or short vowels in them.
- D. Check to see if the pictures are in order and make sense.

4. Before I begin reading, it's a good idea to:

- A. Check to see that no pages are missing.
- B. Make a list of the words I'm not sure about.
- C. Use the title and pictures to help me make guesses about what will happen in the story.
- D. Read the last sentence so I will know how the story ends.

 58

5. Before I begin reading, it's a good idea to:

- A. Decide on why I am going to read the story.
- B. Use the difficult words to help me make guesses about what will happen in the story.
- C. Reread some parts to see if I can figure out what is happening if things aren't making sense.
- D. Ask for help with the difficult words.

6. Before I begin reading, it's a good idea to:

- A. Retell all of the main points that have happened so far.
- B. Ask myself questions that I would like to have answered in the story.
- C. Think about the meanings of the words which have more than one meaning.
- D. Look through the story to find all of the words with three or more syllables.

7. Before I begin reading, it's a good idea to:

- A. Check to see if I have read this story before.
- B. Use my questions and guesses as a reason for reading the story.
- C. Make sure I can pronounce all of the words before I start.
- D. Think of a better title for the story.

 61

8. Before I begin reading, it's a good idea to:

- A. Think of what I already know about the things I see in the pictures.
- B. See how many pages are in the story.
- C. Choose the best part of the story to read again.
- D. Read the story aloud to someone.

9. Before I begin reading, it's a good idea to:

- A. Practice reading the story aloud.
- B. Retell all of the main points to make sure I can remember the story.
- C. Think of what the people in the story might be like.
- D. Decide if I have enough time to read the story.

10. Before I begin reading, it's a good idea to:

- A. Check to see if I am understanding the story so far.
- B. Check to see if the words have more than one meaning.
- C. Think about where the story might be taking place.
- D. List all of the important details.

II. In each set of four, choose the one statement which tells a good thing to do to help you understand a story better while you are reading it.

11. While I'm reading, it's a good idea to:

- A. Read the story very slowly so that I will not miss any important parts.
- B. Read the title to see what the story is about.
- C. Check to see if the pictures have anything missing.
- D. Check to see if the story is making sense by seeing if I can tell what's happened so far.

12. While I'm reading, it's a good idea to:

- A. Stop to retell the main points to see if I am understanding what has happened so far.
- B. Read the story quickly so that I can find out what happened.
- C. Read only the beginning and the end of the story to find out what it is about.
- D. Skip the parts that are too difficult for me.

13. While I'm reading, it's a good idea to:

- A. Look all of the big words up in the dictionary.
- B. Put the book away and find another one if things aren't making sense.
- C. Keep thinking about the title and the pictures to help me decide what is going to happen next.
- D. Keep track of how many pages I have left to read.

14. While I'm reading, it's a good idea to:

- A. Keep track of how long it is taking me to read the story.
- B. Check to see if I can answer any of the questions I asked before I started reading.
- C. Read the title to see what the story is going to be about.
- D. Add the missing details to the pictures.

15. While I'm reading, it's a good idea to:

- A. Have someone read the story aloud to me
- B. Keep track of how many pages I have read.
- C. List the story's main character.
- D. Check to see if my guesses are right or wrong.

 69

16. While I'm reading, it's a good idea to:

- A. Check to see that the characters are real.
- B. Make a lot of guesses about what is going to happen next.
- C. Not look at the pictures because they might confuse me.
- D. Read the story aloud to someone.

17. While I'm reading, it's a good idea to:

- A. Try to answer the questions I asked myself.
- B. Try not to confuse what I already know with what I'm reading about.
- C. Read the story silently.
- D. Check to see if I am saying the new vocabulary words correctly.

18. While I'm reading, it's a good idea to:

- A. Try to see if my guesses are going to be right or wrong.
- B. Reread to be sure I haven't missed any of the words.
- C. Decide on why I am reading the story.
- D. List what happened first, second, third, and so on.

 72

19. While I'm reading, it's a good idea to:

- A. See if I can recognize the new vocabulary words.
- B. Be careful not to skip any parts of the story.
- C. Check to see how many of the words I already know.
- D. Keep thinking of what I already know about the things and ideas in the story to help me decide what is going to happen.

20. While I'm reading, it's a good idea to:

- A. Reread some parts or read ahead to see if I can figure out what is happening if things aren't making sense.
- B. Take my time reading so that I can be sure I understand what is happening.
- C. Change the ending so that it makes sense.
- D. Check to see if there are enough pictures to help make the story ideas clear.

III. In each set of four, choose the one statement which tells a good thing to do to help you understand a story better after you have read it.

21. After I've read a story it's a good idea to:

- A. Count how many pages I read with no mistakes.
- B. Check to see if there were enough pictures to go with the story to make it interesting.
- C. Check to see if I met my purpose for reading the story.
- D. Underline the causes and effects.

 75

22. After I've read a story it's a good idea to:

- A. Underline the main idea.
- B. Retell the main points of the whole story so that I can check to see if I understood it.
- C. Read the story again to be sure I said all of the words right.
- D. Practice reading the story aloud.

office use

76

23. After I've read a story it's a good idea to:

- A. Read the title and look over the story to see what it is about.
- B. Check to see if I skipped any of the vocabulary words.
- C. Think about what made me make good or bad predictions.
- D. Make a guess about what will happen next in the story.

24. After I've read a story it's a good idea to:

- A. Look up all of the big words in the dictionary.
- B. Read the best parts aloud.
- C. Have someone read the story aloud to me.
- D. Think about how the story was like things I already knew about before I started reading.

25. After I've read a story it's a good idea to:

- A. Think about how I would have acted if I were the main character in the story.
- B. Practice reading the story silently for practice of good reading.
- C. Look over the story title and pictures to see what will happen.
- D. Make a list of the things I understood the most.

79

SECTION 3 STARTS HERE

Think of the reasons or causes why you might succeed in reading. The items below concern your opinions of these causes or reasons for your success in reading. Circle one number for each of the following scales.

If you do well in reading,

- | | | | | | | | | | |
|--|---|---|---|---|---|---|---|--|---|
| 1. Are the cause(s) something that is:
outside of you | 1 | 2 | 3 | 4 | 5 | 6 | 7 | inside of you | office use
<input type="checkbox"/> 80 |
| 2. Are the cause(s) something that:
reflects an aspect of yourself | 1 | 2 | 3 | 4 | 5 | 6 | 7 | reflects an aspect of the situation | <input type="checkbox"/> 2/1 |
| 3. Are the cause(s):
controllable by you or other people | 1 | 2 | 3 | 4 | 5 | 6 | 7 | uncontrollable by you or other people | <input type="checkbox"/> |
| 4. Are the cause(s) something that is:
permanent | 1 | 2 | 3 | 4 | 5 | 6 | 7 | temporary | <input type="checkbox"/> |
| 5. Are the cause(s) something:
intended by you or other people | 1 | 2 | 3 | 4 | 5 | 6 | 7 | not intended by you or other people | <input type="checkbox"/> 4 |
| 6. Are the cause(s) something that is:
variable over time | 1 | 2 | 3 | 4 | 5 | 6 | 7 | stable over time | <input type="checkbox"/> |
| 7. Are the cause(s):
something about you | 1 | 2 | 3 | 4 | 5 | 6 | 7 | something about others | <input type="checkbox"/> |
| 8. Are the cause(s) something that is:
changeable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | unchanging | <input type="checkbox"/> |
| 9. Are the cause(s) something for which:
no one is responsible | 1 | 2 | 3 | 4 | 5 | 6 | 7 | someone is responsible | <input type="checkbox"/> 8 |

Think of the reasons or causes why you might do badly in reading. The items below concern your opinions of these causes or reasons for your poor performance in reading. Circle one number for each of the following scales.

If you do **badly** in reading,

office use

- | | | | | | | | | | | |
|--|---|---|---|---|---|---|---|--|--------------------------|----|
| 1. Are the cause(s) something that is:
outside of you | 1 | 2 | 3 | 4 | 5 | 6 | 7 | inside of you | <input type="checkbox"/> | 9 |
| 2. Are the cause(s) something that:
reflects an aspect of yourself | 1 | 2 | 3 | 4 | 5 | 6 | 7 | reflects an aspect of the situation | <input type="checkbox"/> | |
| 3. Are the cause(s):
controllable by you or other people | 1 | 2 | 3 | 4 | 5 | 6 | 7 | uncontrollable by you or other people | <input type="checkbox"/> | |
| 4. Are the cause(s) something that is:
permanent | 1 | 2 | 3 | 4 | 5 | 6 | 7 | temporary | <input type="checkbox"/> | |
| 5. Are the cause(s) something:
intended by you or other people | 1 | 2 | 3 | 4 | 5 | 6 | 7 | not intended by you or other people | <input type="checkbox"/> | 13 |
| 6. Are the cause(s) something that is:
variable over time | 1 | 2 | 3 | 4 | 5 | 6 | 7 | stable over time | <input type="checkbox"/> | |
| 7. Are the cause(s):
something about you | 1 | 2 | 3 | 4 | 5 | 6 | 7 | something about others | <input type="checkbox"/> | |
| 8. Are the cause(s) something that is:
changeable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | unchanging | <input type="checkbox"/> | |
| 9. Are the cause(s) something for which:
no one is responsible | 1 | 2 | 3 | 4 | 5 | 6 | 7 | someone is responsible | <input type="checkbox"/> | 17 |

Note:

Items 4, 7, 11, 13, 15, 19, 21, 22, 25, 27, 28, and 29 in Section 1 measure metacognitive strategy use.

Items 1, 2, 5, 6, 8, 10, 16, 17, 23, and 24 in Section 1 measure cognitive strategy use.

Items 3, 9, 12, 14, 18, 20, and 26 in Section 1 measure self-efficacy.

B.2 READING QUESTIONNAIRE II

Please answer the following before you begin the questionnaire.

NAME:

--	--	--

DATE OF BIRTH:

SCHOOL:

CLASS:

MALE FEMALE (circle)

Do you consider yourself: MAORI EUROPEAN POLYNESIAN

ASIAN OTHER (circle)

If you have circled OTHER, please specify: _____

What is your first language: _____

Introduction

This questionnaire measures your confidence and knowledge for **READING**. The questionnaire is arranged in two separate sections, each with a different set of instructions. Be sure to read these instructions carefully. **THERE ARE NO RIGHT OR WRONG ANSWERS TO THIS QUESTIONNAIRE.**

Your data will be treated as confidential. Nobody will see your answers apart from the researcher.

SECTION 1

Directions

In this section we will be looking at how **confident** you feel in reading and understanding various things. Each item is followed by a rating scale. Circle **one number** on the scale that matches how confident you are in being able to successfully read and understand what the author is saying for each of the listed tasks. Remember that the higher the number the more sure you are, while the lower the number the less sure you are. Please be honest and mark how you really feel right now. There are no right or wrong answers.

Let's try some practice questions first. The practice questions involve how confident you feel in jumping varying distances. Each item is followed by a rating scale with numbers ranging from 0 to 100. Circle the number on the scale that matches how sure you are that you could jump the distance given.

DISTANCE JUMPING

0.5 metres

0	10	20	30	40	50	60	70	80	90	100
no chance										completely sure

1.6 metres

0	10	20	30	40	50	60	70	80	90	100
no chance										completely sure

2.4 metres

0	10	20	30	40	50	60	70	80	90	100
no chance										completely sure

Now turn over to the next page and begin.

SECTION 2 STARTS HERE

In this section of the questionnaire there are 20 multiple choice questions. You are asked to choose only **ONE** answer for each question. Put a **circle** around the letter in front of the **BEST** answer to the question. Answer every question even when you are not sure of your choice.

1. What is the hardest part about reading for you?

- a. Sounding out the hard words.
- b. When you don't understand the story.
- c. Nothing is hard about reading for you.

office use

 10

2. The best way to help you become a better reader is

- a. having more people help you when you're reading.
- b. reading easier books with shorter words.
- c. checking to make sure you understand what you read.

3. What is special about the first sentence or two in a story?

- a. They always begin with "Once upon a time..."
- b. The first sentences are the most interesting.
- c. They often tell what the story is about.

4. How are the last sentences of a story special?

- a. They are the exciting, action sentences.
- b. They tell you what happened.
- c. They are harder to read.

5. How can you tell which sentences are the most important ones in a story?

- a. They are the ones that tell the most about the characters and what happens.
- b. They're the most interesting ones.
- c. All of them are important.

 14

6. If you could only read some of the sentences in the story because you were in a hurry, which ones should you read?

- a. Read the sentences in the middle of the story.
- b. Read the sentences that tell you the most about the story.
- c. Read the interesting, exciting sentences.

7. When you tell other people about what you read, what should you tell them?

- a. What happened in the story.
- b. The number of pages in the book.
- c. Who the characters are.

8. If the teacher told you to read a story to remember the general meaning, what would be the best thing to do?

- a. Skim through the story to find the main parts.
- b. Read all of the story and try to remember everything.
- c. Read the story and remember all of the words.

 17

9. Before starting to read, what kinds of plans can you make to help you read better?
- You shouldn't make any plans. You should just start reading.
 - You should choose a comfortable place.
 - You should think about why you are reading.
10. If you had to read very fast and could only read some words, which ones should you try to read?
- Read the new vocabulary words because they are important.
 - Read the words that you could pronounce.
 - Read the words that tell the most about the story.
11. What things can be read faster than others?
- Books that are easy to read.
 - When you've read the story before.
 - Books that have a lot of pictures.
12. Why do you go back and read things over again?
- Because it is good practice.
 - Because you didn't understand it.
 - Because you forgot some words.
13. What is the best thing to do if you come to a word and you don't know what it means?
- Use the words around it to figure it out.
 - Ask someone else.
 - Go on to the next word.
14. What is the best thing to do if you don't know what a whole sentence means?
- Read it again.
 - Sound out all of the words.
 - Think about the other sentences in the paragraph.
15. What parts of the story can be skipped as you are reading?
- The hard words and parts you don't understand.
 - The unimportant parts that don't mean anything for the story.
 - Should not skip anything.
16. If you are reading a story for fun, what should you do?
- Look at the pictures to get the meaning.
 - Read the story as fast as you can.
 - Imagine the story like a movie in your mind.

office use

 18 21 25

17. If you are reading for science or social studies, what is the best thing to do if you want to remember the information?

- a. Ask yourself questions about the important ideas.
- b. Skip the parts you don't understand.
- c. Concentrate and try hard to remember it.

office use

 26

18. If you are reading for a test, which would help you the most?

- a. Read the story as many times as possible.
- b. Talk about it with somebody to make sure you understand it.
- c. Say the sentences over and over.

19. If you are reading a library book to write a book report, which would help you the most?

- a. Sound out words you don't know.
- b. Write it down in your own words.
- c. Skip the parts you don't understand.

20. Which of these is the best way to remember a story?

- a. Say every word over and over.
- b. Think about remembering it.
- c. Write it down in your own words.

 29

B.3 TARSUM: MEASURES, SCORING PROCEDURES, AND DEFINITIONS OF STRATEGIES

B3.1 PASSAGES

***THE VOLCANO - VERSION 1**

The scientists approached the crater's edge fascinated at the prospect of recording the spectacle of an inactive volcano smouldering again. Intent on their photography, they ignored an ominous rumbling. Within seconds, the subterranean cauldron exploded violently ejecting a great quantity of rocks. Fortunately these fell the in direction of the opposite slope.

Greatly alarmed by this premature explosion, the group hastily began the descent. Immediately, fiery boulders from a gigantic avalanche hurtled around. Aware that their apparatus hindered progress, they abandoned all equipment except their precious cameras. Then came an urgsy moment. As they were evading flying fragments, one of them was struck off-balance by a rebounding boulder. A lengthy halt would have been disastrous. Everyone was, therefore, very relieved when they found the injuries were superficial. They resumed their dangerous scramble to regain safety just before the surroundings were destroyed. The scientists were pleased that they did not leave any of their equipment behind.

When they finally reached the bottom of the volcano, the scientists hurried towards their cabin which nearby was. One of the young scientists was very seriously injured and was screaming because of the pain. Luckily, there was an ikran waiting near their cabin and the young scientist was whisked away quickly to hospital.

THE VOLCANO- VERSION 2

The scientists moved towards the crater's edge interested at the chance of recording the event of an inactive volcano burning quietly again. They were so interested in their photography that they ignored the rumbling sound of the volcano. Within seconds, the "underground boiling pot" exploded throwing out a great number of rocks. Luckily, these fell the in direction of the opposite slope.

Greatly frightened by this early explosion, the group quickly began to climb downwards. Immediately, burning rocks from a big landslide hurtled around. Aware that carrying their tools was slowing them down, they threw away all their tools except their cameras. Then came an urgsy moment. As they were avoiding the flying pieces of rock, one of them was struck off-balance by a rebounding rock. A long stop would have been dangerous. Everyone was, therefore, greatly relieved when they found the injuries were not very serious. They carried on their dangerous climb down the volcano to regain safety just before the place around them was destroyed. The scientists were pleased that they did not leave any of their tools behind.

When they finally reached the bottom of the volcano, the scientists ran towards their cabin which nearby was. One of the young scientists was very seriously hurt and was screaming because of the pain. Luckily, there was an ikran waiting near their cabin and the young scientist was rushed to the hospital.

THE VOLCANO- VERSION 3

The men moved towards the crater's edge happy at the chance of studying the inactive volcano burning quietly again. They came to take photographs and to record what was happening at the volcano. They were so interested in their photography that they ignored the rumbling sound of the volcano. Soon, the "underground boiling pot" exploded throwing out a large number of rocks. Luckily, these fell the in direction of the opposite slope.

Very frightened by this early explosion, the group quickly began to climb down the volcano. At once, burning rocks from a big landslide flew around them. They knew that carrying their tools was slowing them down. So they threw away all their tools except their cameras. Then came an urgsy moment. As they were avoiding the flying pieces of rock, one of them was hit by a rebounding rock. A long stop would have been dangerous. Everyone was very relieved when they found that no one was very badly hurt. They carried on climbing down the volcano to safety just before the place around them was destroyed. The men were happy that they did not leave any of their tools behind.

When they reached the bottom of the volcano, the men ran towards their cabin which closeby was. One of the men was very badly hurt and was screaming because of the pain. Luckily, there was an ikran waiting near their cabin and he was rushed to the hospital.

Note:

Notations that were used by the interviewer to record oral reading behaviours:

s^c - self-corrected successfully

s^u - self-corrected unsuccessfully

x - made an error without correcting

b - bleeped

B.3.2 INTERVIEW CODING FORM

Name: _____ School: _____

--	--	--

VOLCANO

	Strategies	Observed	Notes	
1a. Look at this page. What do you think the story will be about? (1 mark)	Scans text	_____	_____	
	Looks at title	_____	_____	
	Uses picture	_____	_____	
	b. How did you know this?	Refers to prior knowledge	_____	_____
		Main idea or topic sentence	_____	_____
Other	_____	_____		

Answer (Ans): The volcano (1)

	Strategies	Observed	Notes
2/1a. What do you think will happen next? (1 mark)	Predicts based on prior knowledge	_____	_____
	Predicts based on text cues	_____	_____
b. Why do you think that?	Rereads	_____	_____
	Other	_____	_____

Ans: The volcano is likely to explode (1)
(rumbling sound -warning sign)

	Strategies	Observed	Notes
2/2a. What had the scientists (<i>men</i>) come to do at the volcano? (2 marks)	Predicts based on prior knowledge	_____	_____
	Predicts based on text cues	_____	_____
b. How could you tell?	Rereads	_____	_____
	Other	_____	_____

Ans: To record/study/explore the activity of the volcano (1)
and To photograph the volcano (1)

3a. What do you think "subterranean cauldron" (*underground boiling pot*) refers to in the sentence you've read? (1 mark)

b. How could you tell? (How did you go about trying to get your answer?)

Ans: refers to the volcano (1)
refers to larva/boiling liquid (1/2)

Strategies	Observed	Notes
Monitors understanding	_____	_____
Uses context cues	_____	_____
Substitution looks or sounds similar	_____	_____
Mentions others as resources	_____	_____
Mentions dictionary as resource	_____	_____
Other	_____	_____

4/1. What are you thinking here? (Why did you bleep?)

corrected it? yes/no

Strategies	Observed	Notes
Monitors/indicates scrambled words	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Rereads	_____	_____
Other	_____	_____

4/2a. Where were the scientists (*men*) when the volcano exploded? (1 mark)

b. How could you tell?

Ans: Near the crater's edge/ near top of volcano (1)

Strategies	Observed	Notes
Predicts based on text/context cues	_____	_____
Rereads	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Other	_____	_____

4/3a Why wasn't the first explosion so dangerous for the climbers? (1 mark)

b. How did you get your answer?

Ans: The rocks fell on the opposite slope (1)

Strategies	Observed	Notes
Rereads	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Predicts based on text/context cues	_____	_____
Other	_____	_____

5/1a. What are you thinking here?
(Why did you bleep?)

Strategies	Observed	Notes
Monitors word (urgsy)	_____	_____
Evaluates i.e. makes a judgmental statement about text/word	_____	_____
Rereads	_____	_____
Other	_____	_____

5/2a. Is there such a word (urgsy)?
(yes/no)

What do you think "urgsy" means? (1 mark)

b. How could you tell?
(How did you go about looking for the answer?)

Ans: anxious/fearful/troubled (1)

Strategies	Observed	Notes
Monitors understanding	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Uses context cues	_____	_____
Mentions others as resources	_____	_____
Mentions dictionary as resource	_____	_____
Other	_____	_____

6/1a. What do you think "rebounding boulder (*rock*)" means? (1 mark)

b. How could you tell?

Ans: A boulder/rock that had bounced off the wall (1)

Strategies	Observed	Notes
Infers based on prior knowledge	_____	_____
Rereads	_____	_____
Uses context cues	_____	_____
Monitors understanding	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Substitution looks or sounds similar	_____	_____
Mentions others as resources	_____	_____
Mentions dictionary as resource	_____	_____
Other	_____	_____

6/2a. What accident occurred while they were climbing down? (1 mark)

b. How could you tell?
(How did you about looking for the answer?)

Ans: One of the scientists (men) was struck off-balance by a rebounding rock (1)

Strategies	Observed	Notes
Scans text	_____	_____
Rereads	_____	_____
Infers based on text clues	_____	_____
Other	_____	_____

7. What are you thinking here? (Why did you bleep?)
(If no answer) Do you agree with the sentence?

Strategies	Observed	Notes
Monitors/ Indicates inconsistency in text	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Rereads	_____	_____
Other	_____	_____

8. What are you thinking here?
(Why did you bleep?)

corrected it? yes/no

Strategies	Observed	Notes
Monitors/ Indicates scrambled words	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Rereads	_____	_____
Other	_____	_____

9/1. What are you thinking here?
(Why did you bleep?)

Strategies	Observed	Notes
Monitors/Indicates inconsistency in text	_____	_____
Monitors the word ikran (nonsense word)	_____	_____
Evaluates i.e. makes a judgmental statement about text	_____	_____
Rereads	_____	_____
Other	_____	_____

9/2a. What do you think is another word for "ikran"?
 or What do you think "ikran" may mean?
 (1 mark)

b. How could you tell?
 (How did you go about trying to find the answer?)

Ans: ambulance (1)

Strategies	Observed	Notes
Monitors the word ikran	_____	_____
Evaluates i.e. makes a judgmental statement about text/word	_____	_____
Rereads	_____	_____
Infers based on text/context clues	_____	_____
Substitution looks or sounds similar	_____	_____
Mentions others as resources	_____	_____
Mentions dictionary as resource	_____	_____
Other	_____	_____

10a. If you wanted to tell your friends about this story, what would you tell them? (4 marks)

b. How did you decide what things to tell them?

Ans: scientists (men) came to explore volcano (1)

ignored warning sign/ volcano exploded and group hurried down (1)

threw away tools except cameras to speed up descent (1)

or
 one of scientists was struck off-balance (1)

hurried down to safety where injured man was sent to hospital (1)

Strategies	Observed	Notes
Retells mostly main ideas or summarizes	_____	_____
Organizes ideas in recall	_____	_____
Expresses opinions or reactions	_____	_____
Relates to personal experiences	_____	_____
Plans summary	_____	_____
Other	_____	_____

LISTED MAIN POINTS IN AN ORGANIZED MANNER (YES/NO)

Planning strategy: _____

11a. Which paragraph did you think was hardest to read for you?

Paragraph _____

b. How did you decide which was the hardest?

Strategies	Observed	Notes
Rereads	_____	_____
Infers based on difficulty of understanding	_____	_____
Infers based on difficulty of pronouncing words	_____	_____
Other	_____	_____

(CORRESPONDENCE OF PARAGRAPH DIFFICULTY WITH STUDENT'S ACCURACY SCORE OF PARAGRAPH: YES/NO)

Evaluation strategy: _____

12. Which of these questions do you think you answered correctly? (please tick)

- ___ a. What do you think the story will be about?
- ___ b. What do you think will happen next?
- ___ c. What had the scientists (men) come to do at the volcano?
- ___ d. What do you think subterranean cauldron (underground boiling pot) refers to?
- ___ e. Where were the scientists (men) when the volcano exploded?
- ___ f. Why wasn't the first explosion so dangerous for the climbers?
- ___ g. What do you think "urgsy" means?
- ___ h. What do you think "rebounding boulder (rock)" means?
- ___ i. What accident occurred while they were climbing down?
- ___ j. What do think is another word for "ikran"?
- ___ k. If you wanted to tell your friends about this

Strategy

(CORRESPONDENCE OF STUDENT'S EVALUATION OF EACH QUESTION WITH COMPREHENSION SCORE OF EACH QUESTION: _____ OF THE 11 QUESTIONS)

Evaluation strategy: _____

B.3.3 SCORING PROCEDURES FOR METACOGNITIVE STRATEGIES

Table B1. Scoring of metacognitive strategies used in TARSUM

Questions	Type of strategy	Score
4.1	monitoring - scrambled words	2 - monitors and corrects 1 - monitors only 0 - doesn't monitor
5.1	monitoring - nonsense word	2 - monitors without prompting 1 - monitors with prompting 0 - doesn't monitor
7	monitoring - inconsistency	2 - monitors and corrects without being prompted 1 - monitors only without being prompted but cannot correct, or monitors and corrects after being prompted 0 - doesn't monitor
8	monitoring - scrambled words	2 - monitors and corrects 1 - monitors only 0 - doesn't monitor
9.1, 9/2a	monitoring - inconsistency	2 - monitors and corrects 0 - doesn't monitor
	monitoring - nonsense word	2 - monitors 0 - doesn't monitor
10	planning - summary	1 for organized main points 0 for no planning
11	evaluation - paragraph difficulty	1 for match between subjective and objective evaluation
12	evaluation - comprehension	1 for each correct evaluation (maximum score = 11)

B3.4 DEFINITIONS OF STRATEGIES MEASURED IN THE TARSUM

Table B2. Definition of cognitive and metacognitive strategies used in the TARSUM

Type of strategy	Definition
<i>Cognitive strategies</i>	
Scans text	Students states that he looked at the text, looks forward.
Main ideas/Topic sentence	Subjects mention the thematic ideas (or the important information) of the story rather than the details (or less important information). Topic sentence usually refers to the first sentence of the paragraph which contains the main theme.
Rereading	Subjects return to the text to read again orally or may state they are rereading.
Refers to prior knowledge	Subjects identify personally with characters in the story or add information related to text based on content area knowledge, or personal experience/knowledge.
Predicting based on prior knowledge	Subjects anticipate happenings by using information related to text based on content area knowledge or personal experience/knowledge.
Predicting based on text/context cues	Subjects interpret the text based on information contained within the text or uses context to figure out something.
Substitution looks or sounds similar	Uses word attack skills (e.g., prefixes and suffixes) or finds meaning based on a word they think is or sounds similar.
Summarizing	Subjects summarize the text, using their own words.
Organizes ideas	States ideas in a logical order.
Infers based on context clues	Subjects interpret the text based on information contained within the text or uses context to figure out something.
Mentions dictionary or others as resources	Finds meaning from dictionary or other people
Substitution looks or sounds similar	Uses word attack skills (e.g., prefixes and suffixes) or finds meaning based on a word they think is or sounds similar.
<i>Metacognitive strategies</i>	
Monitoring understanding	Subjects state failure to understand the meaning of a clause, sentence, word or the story. Subjects may say something to indicate that they are struggling to understand word or sentence.
Evaluating	Subjects make a judgmental statement about information in the text or subjects comment upon the author's arrangement of the story or analyze its content.
Planning	Subjects sets goals, skims a text before reading, or does a task analysis. Alternatively, subject presents key information in an organized manner

APPENDIX C

Pereira-Laird, J. A., & Deane, F. P. (1995). *Validation of a reading self-efficacy scale for adolescents*. Manuscript submitted for publication.

ABSTRACT

Theory development in the area of reading self-efficacy has been impeded by a lack of psychometrically sound instruments to measure reading self-efficacy. Two studies were designed to remedy this deficiency by developing a reading self-efficacy measure with adequate psychometric properties. The current study describes the development and preliminary validation of the task specific measure of reading self-efficacy (TRSE). In the pilot study a preliminary version of the TRSE scale consisting of 28 items was administered to a sample of 28 adolescents. Following analyses, a final version of 6 items was generated which showed acceptable internal consistency. In general, the TRSE scale correlated as expected with criterion measures. In the second study, confirmatory factor analyses with an independent sample of 139 students confirmed the unidimensional nature of the TRSE. Convergent validity was supported by its modest correlation with a general measure of reading self-efficacy (GRSE). Confirmatory factor analysis of the TRSE and GRSE items led to the interpretation of a two-factor solution, providing empirical validity for a task specific reading self-efficacy scale and a general reading self-efficacy scale. Confirmation of several predicted conceptual relationships between the TRSE and other reading related factors (i.e., achievement, metacognitive knowledge, anxiety, and perceived value) provided evidence of construct validity. The TRSE appears useful for educational and research purposes with adolescent children. Limitations and suggestions for further research are discussed.

Self-efficacy is considered to be one mechanism of cognitive self-evaluation that mediates skilled performance (Bandura, 1982, 1986). Perceived self-efficacy is defined as the self-perceptions of one's skills and competence to perform the behaviours required to obtain desired outcomes (Bandura, 1986; Schunk, 1987, 1989a). Bandura and his associates (Bandura, 1977; Bandura & Schunk, 1981) maintain that strong perceptions of self-efficacy are based upon the gradual acquisition of social, linguistic, cognitive, and physical skills through personal or socially mediated experiences. The impact of these mediated experiences on self-efficacy is dependent on how they are cognitively appraised.

It has been argued that an individual's efficacy mediates the influence of determinants like gender, prior experience, anxiety, self-concept, and attributions on subsequent performance and is the stronger predictor of that performance when those determinants are controlled (Bandura, 1986; Hackett & Betz, 1989; Pajares & Miller, 1994). Self-efficacy enhances or impairs performance through cognitive, affective, or motivational mediating processes that contribute to the prediction of future behaviour (Bandura & Schunk, 1981; Graham & Golan, 1991). In this vein, Bandura stresses that self-efficacy assessments are "not simply inert predictors of future behaviour" in that people with efficacious self-beliefs "make things happen" (1989a, p. 731).

Although it has been established that self-efficacy is a strong predictor of behaviour (Bandura, 1986), research on the relationship between self-efficacy and academic performance is still limited (Bouffard-Bouchard, 1990). A strong sense of competence facilitates cognitive processes and academic performance. For instance, expectations of self-efficacy determines whether a person will attempt a given task, how much effort will be expended, and how much persistence will be displayed on pursuing the task in the face of difficulties. Thus, self-efficacy levels can enhance or impede the motivation to act.

Empirical research has demonstrated moderate predictive correlations of self-efficacy and academic performance (Bandura & Schunk, 1981; Schunk & Gunn, 1986). There has been considerably more research on self-efficacy in the area of mathematics (e.g., Hackett & Betz, 1989; Pajares & Miller, 1994; Randhawa, Beamer, & Lundberg, 1993; Schunk, 1984, 1989; Schunk & Gunn, 1986; Seegers & Boekaerts, 1993) but relatively little in the domain of reading. It is important to study efficacy within domains, since self-efficacy has been basically conceptualized as a situation specific belief (Sherer, Maddux, Mercandante, Prentice-Dunn, Jacobs, & Rogers, 1982), and students' perceptions of self-efficacy can vary across domains (Bandura, Adams, Hardy, & Howells, 1980).

As a complex cognitive task, reading should be affected by the self-efficacy mechanism proposed by Bandura (1986). The paucity of research in perceived self-efficacy for reading (Cohen, McDonell & Osborn, 1989), is surprising considering that the construct is intimately tied to reading achievement (Henk & Melnick, 1992).

The valid interpretation of self-efficacy in reading or achievement research requires the use of reliable and valid instruments. A review of the literature on reading revealed that there are few instruments developed to measure reading efficacy, with valid and reliable measures being sparse (Henk & Melnick, 1992). In fact, many current self-perception theorists often employ measures of self-concept instead of efficacy to predict performances across situations (e.g., Chapman & Tunmer, 1995). Bandura (1986) argued that self-efficacy and self-concept represent different phenomena. Furthermore, self-concept is not measured at the same level of specificity as efficacy (Pajares & Miller, 1994).

Self-efficacy is usually measured by asking subjects to judge their capability of succeeding at specific target tasks within the domain or subdomain being tested (cf. Berry, West, & Dennehey, 1989; Schunk & Gunn, 1986). Students may, for example, be asked to judge their capability in solving a series of multiplication problems which they are shown briefly or to assess their likely level of remembering in a series of memory task situations.

Four different methodologies for obtaining reading or verbal self-efficacy can be found in the literature. Most existing measures of efficacy tend to measure achievement globally (Horn, Bruning, Schraw, Curry, & Katkanant, 1993; Pintrich, Anderman, & Klobucar, 1994; Pintrich & De Groot, 1990). Unfortunately, such global measures are only modestly correlated and impair the ability to understand and predict behaviour in particular situations by not taking into account the complexity and variation of self-efficacy perceptions (Bandura, 1986). In addition, although a number of these global measures purport to measure efficacy, they focus more on motivational/effort variables than actual feelings and beliefs regarding personal competence (e.g., Sherer et al., 1982).

A second methodology used in the investigation of reading or verbal efficacy has involved specific words in which individuals state how well they can read or define particular words (Marsh, Walker, & Debus, 1991; Zimmerman & Martinez-Pons, 1990). As reading requires the integration and application of multiple subskills (Paris & Oka, 1986), a measure of efficacy involving words only presents a rather limited view of reading efficacy.

A third methodology involves measuring antecedent factors related to efficacy. Recently, Henk and Melnick (1992, 1995) formulated the reading self-perception scale in an attempt to measure efficacy perceptions in the domain of reading. However, the items included in the

scale represented a measure of perceived sources of reading self-efficacy information. Henk & Melnick included items from four major categorical sources of information thought to be related to reader self-efficacy. These included performance, observational comparison, social feedback, and physiological states. Although these categorical sources provide the basis for the formation of efficacy, they do not provide a direct indication of a person's sense of efficacy. It can be argued that these items measure characteristics that are related to personal efficacy, but are not synonymous with efficacy (cf. Lent, Lopez, & Bieschke, 1991).

Another group of measures involves measuring confidence in specific reading tasks (Shell, Murphy, & Bruning, 1989; Seegers & Boekaerts, 1993) or in the general domain of reading (Cohen et al., 1989; McMillan, Simonetta, & Singh, 1994). However, most of these measures include items which measure the construct marginally at best (Bandura, 1989b; Cohen et al., 1989; Gordon, 1990; Seegers and Boekaerts, 1993). For example, Seegers and Boekaerts (1993) developed a three item scale to measure efficacy for three topics in the reading curricula. Similarly, Bandura (1989b) in his academic self-efficacy scale had only one item assessing reading and writing language skills. Such circumscribed measures of reading may yield a curtailed distribution of scores which would in turn lower the magnitude of correlation between efficacy and performance.

An expanded efficacy assessment in which individuals judge the strength of their efficacy to fulfil gradations of reading task demands would be more sensitive to the variation in perceived reading self-efficacy in any given sample than assessment of personal efficacy at a single level of task difficulty (Bandura, 1989a). Such an expanded assessment of efficacy is provided by Shell et al. (1989). They developed a scale for reading self-efficacy which adequately measures the domain of reading and has good psychometric properties. However, this scale was developed to measure efficacy of undergraduate university students.

The present study focused on the development of a measure of reading self-efficacy for a number of reasons. The primary rationale for developing such a scale was to provide a research tool to assist in studying the relationship between efficacy and reading achievement. Systematic theory development in this particular area has been impeded by a lack of reliable and valid instruments to measure reading self-efficacy. Furthermore, another major problem with most of the existing efficacy measures is that few have been normed in anything resembling a systematic, empirical fashion. A variety of measures have been used to measure reading self-efficacy, but they lack adequate psychometric properties, or are too global in nature.

Second, construction of the reading self-efficacy scale was further prompted by the fact that there are no such measures that are normed with New Zealand (NZ) children and that a measure of self-efficacy for a variety of reading tasks are nonexistent for preadolescent and adolescent children. Existing measures have been constructed for American children who probably have a different reading curriculum from NZ children (e.g., McMillan et al., 1994; Seegers & Boekaerts, 1993; Shell et al., 1989).

Third, since students enter school with differing levels of reading self-efficacy, they are differentially influenced by reading instruction and remediation. Hence, a reading specific self-efficacy scale might enable teachers to tailor the course and style of instruction to the student's needs. The scale might also provide a useful index of progress in intervention since expectations of efficacy should change during intervention.

Finally, the methodology of self-efficacy has a number of advantages (Berry et al., 1989). In obtaining confidence ratings, individual variability and age differences in confidence can be examined. Reading performance predictions are made on a task-by-task basis so that predictions correspond directly to performance tasks. The examination of efficacy on a variety of reading tasks provides data on the extent to which reading self-evaluation varies across specific tasks.

In the present study, attempts were made to establish construct validity of the developed scale. Self-efficacy ratings can be validated by correlating the ratings with performance, metacognitive knowledge, motivational beliefs, and persistence or effort. Numerous studies have shown that students with a high sense of academic efficacy display greater persistence, effort, and intrinsic interest in their academic learning and performance (Schunk, 1984, 1989, 1991), show less anxiety (Hackett, 1985), and are more knowledgeable about reading (Paris & Oka, 1986). Research has also indicated that self-perceptions of reading ability are significantly correlated to reading grades and reading achievement (Paris & Oka, 1986). In addition, the literature on gender differences has indicated that girls often surpass boys in their verbal efficacy and verbal achievement (e.g., Maccoby & Jaclin, 1974).

In summary, this article describes the development of the Task Specific Reading Self-efficacy scale (TRSE). Two studies using different samples provide data on the psychometric properties of the TRSE. The development of the scale is reported in detail in the pilot study. The second study provides further validation data for the TRSE scale.

STUDY 1 - PILOT STUDY: SCALE DEVELOPMENT

The pilot study was designed to construct the TRSE and to determine its reliability. In addition, construct and discriminant validity were evaluated through correlations with several achievement measures, and assessment of group difference in efficacy between good and poor readers.

METHOD

Subjects

Parent and child permission were obtained for children to participate in an investigation of the psychometric properties of the TRSE. Twenty eight students enrolled in Form 1 - Form 3 (junior-high) classes in 2 schools at Palmerston North (New Zealand) constituted the sample. Subjects (11 M, 17 F), aged 11-13 years ($M = 12.10$ years, $SD = 0.81$) were predominantly of European descent and from middle-class families. They were chosen by teachers who classified them as good ($n = 18$) or poor readers ($n = 10$) based on their classroom performance.

Measures

Progressive Achievement Tests

Students' reading achievement was assessed using two reading tests from the Progressive

Achievement Test Series (PAT). The PATs were Reading Comprehension and Reading Vocabulary (Reid & Elley, 1991). Mathematics achievement was assessed using the PAT Mathematics test (Reid, 1993). These tests are group administered, NZ normed, paper-and-pencil scales, administered by the majority of NZ primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in NZ (Reid, 1993; Reid & Elley, 1991).

Procedure

Development of the TRSE

The TRSE was constructed for use in a larger study (Pereira-Laird, 1994) to tap students' efficacy on a variety of reading tasks employed in school. It was developed on the basis of methods outlined by Bandura (1982, 1986) and Schunk (1984). Following Bandura's (1986) recommendation that level of task be varied when self-efficacy is being assessed, the scale was composed of items that varied in difficulty. This also enabled one to discriminate varying perceptions of efficacy by low and high achieving readers.

The particular reading tasks were generated on the basis of data obtained from four different sources. First, items from Shell et al.'s (1989) reading self-efficacy scale that were considered developmentally appropriate for junior high students were included. Second, items were also generated from pilot interviews with junior high school children by asking them what reading materials they read often. The most frequently identified items from the respondents were retained. Third, following the item selection method of Meier, McCarthy, & Schmek (1984), information obtained from the reading course objectives of intermediate school children was used to guide the construction of the items. Fourth, item selection was also made in consultation with reading specialists and teachers of junior high (intermediate aged) school children. A major criterion for inclusion of an item in the TRSE was that the reading-related activity referred to was relevant for young adolescent children with varying reading abilities. Unlike the scale developed by Shell et al. (1989), the items were deliberately designed to be simple and easy to read, so as to be developmentally appropriate for junior high students across a wide intellectual ability range. A pool of 28 items relating to individuals' generalized expectancies concerning their efficacy in performing a variety of reading tasks were generated. Using similar procedures to McMillan et al. (1994), the items were reviewed by three principals of intermediate schools for clarity, content validity, and face validity. All items with retained apart from minor modifications in the wording of the items. The TRSE items are provided in Table 1.

The initial version of the TRSE was administered to a sample of 28 children by the researcher. After reading the instructions aloud, students were familiarized with the question format. Efficacy was measured on a scale that ranged from zero (no chance) to 100 (complete certainty) in 10-unit intervals. Children initially received practice by judging their certainty of successfully jumping progressively longer distances. In this explicit manner, students learned the meaning of the scale's direction and the different numerical values.

Subjects were asked to rate their confidence in being able to read and understand what the author was saying for twenty eight different reading tasks. Self-efficacy scores were

computed by calculating the mean for the items. After completion of the questionnaire, subjects were asked to give their comments on the different items, particularly those that they did not understand or were ambiguous for them. The researcher noted items that required frequent clarification for children to understand.

RESULTS

Table 1. Items of the task specific reading self-efficacy instrument

Item description
1. A letter from a classmate
2. A comic book
3. Your school's newsletter or a class magazine
4. A letter from an uncle or aunt addressed to you
5. A short fiction story written by a New Zealand writer
6. A magazine such as the "RTR Countdown"
7. The TV guide
8. A selection of poems given in class
9. The daily newspaper
10. All articles in a magazine such as "The Listener"
11. A 150 page novel
12. A short play
13. Your English class assignments/work and projects
14. The questions on your class test
15. The encyclopedia (e.g., "The World Book")
16. An instruction manual for operating a ghetto blaster
17. A recipe for cooking a meal
18. A reading passage in your test or exam
19. A dictionary
20. A telephone book
21. The School Journal
22. A play by Shakespeare like "Romeo and Juliet"
23. A school library book chosen by you
24. A book on the life of someone famous
25. A non-fiction book on people or places in NZ
26. An article in a magazine called "Shakedown"
27. A non-fiction book on a topic that interests you
28. A book on New Zealand issues, e.g., "The Treaty of Waitangi"

One goal of this study was to reduce the item pool since we had specific goals for the TRSE. We wanted a restricted pool of items because the ultimate measure was to be used in conjunction with other measures in a larger study for research purposes. A scale with a relatively small number of items was considered necessary because of the time constraints. Furthermore, brevity was desirable to allow students to complete the measure quickly.

In order to distinguish meaningful items, the initial data-analytic strategy consisted of

identifying those individual items that correlated highly and positively with reading achievement measures. Items which had the highest positive correlations with both the PAT Reading Comprehension and Vocabulary measures were retained since self-efficacy theory suggests that efficacy for reading is related to reading achievement.

As a result items 1, 3, 5, 9, 10, 11, 12, 15, 19, 21, 22, 23, and 25 were retained for further analyses. The correlations between these items and the reading achievement measures were generally low and statistically nonsignificant since the sample size was small. Correlations ranged from .15 to .46 for the comprehension measure and from .20 to .62 for the vocabulary measure. On the basis of univariate statistics and feedback following administration, items were eliminated if they required frequent clarification from the students, or had a restricted response range. The result of this process was a 9 item scale: items 1, 3, 5, 9, 11, 12, 15, 23, and 25.

An exploratory analysis was conducted with the resulting list of items. This involved combining the 9 items in different groupings to yield 6 - 9 item scales. It was thought a minimum of 6 items was required to be sufficiently comprehensive. Items 1, 11, 12, and 15 were always retained in every combination as these items had the highest correlations with PAT Reading Comprehension.

For each combination, a mean efficacy score was calculated by summing the scores of the items and dividing it by the number of items in that combination. Reliability estimates, and construct and discriminant validity were obtained for each of the combinations. Construct validity was assessed by correlating the scales with PAT Reading Comprehension, Vocabulary, and Mathematics measures. Discriminant validity was assessed by performing t-tests for efficacy between good and poor readers. Combinations which yielded high reliability estimates and good discriminant and construct validity were retained for further analyses. More specifically, a scale was chosen if it met the following criteria: (a) it contained both easy and difficult reading tasks, (b) it had a high internal consistency estimate of at least .80, (c) it had statistically significant correlations with PAT Reading Comprehension and PAT Vocabulary measures, (d) it had a statistically nonsignificant correlation with PAT Mathematics, and (e) there was a statistically significant difference in the efficacy scores between good and poor readers.

Results indicated that combinations of 8 and 9 item scales did not meet all five criteria. Only two of the 7-item and one of the 6 item combinations met all five criteria. The three resulting combinations were judged by a sample of 4 reading specialists as to their viability and comprehensiveness as a measure of self-efficacy for young adolescent children. Two of the combinations were eliminated as they had an item which would not be applicable to all schools (e.g., school newsletter), resulting in the 6-item scale being chosen.

Reliability assessed with Cronbach's alpha was .81, indicating a high degree of internal consistency for the 6-item scale. Item-analysis indicated that the correlations between items and subscale scores were positive and exceeded .50 for all items, except item 1, which had a correlation of less than .20. However, the retention of item 1 was justified because it was considered necessary to include "easy" items in the scale. Means ranged from 60.71 to 93.92, suggesting that the reading tasks varied in difficulty.

Support for the convergent validity of the 6-item TRSE scale was demonstrated as it

correlated highest with its targeted criterion measure. More specifically its correlations with the PAT Comprehension ($r = .40, p = .02$) and Vocabulary ($r = .50, p = .001$) measures were higher than its correlation with the PAT mathematics measure ($r = .24, p = .20$). Evidence of discriminant validity was suggested by the generally lower correlations with the non-targeted measure of PAT Mathematics and by statistically significant differences in efficacy scores between good and poor readers ($t = 2.34, p = .03$). Results reported in the remainder of this article are based on the 6-item (final) version of the TRSE.

Although the pilot study was an initial attempt at developing the TRSE scale, it was limited by its small sample size. Consequently, a second study was conducted to replicate the findings and provide additional psychometric data for the TRSE.

STUDY 2 - VALIDATION OF THE SCALE

Study 2 reports confirmatory factor analytic data for the final TRSE scale with a random sample of 139 students. Although confirmatory factor analysis is one method for investigating the construct validity of the efficacy measure, additional techniques must be used to derive conclusions about the construct under investigation. For example, support for the construct validity of the TRSE should include concurrent validity studies that explore the relationships between the TRSE and other measures of related constructs. For instance, theory suggests modest correlations with reading anxiety and perceived value of reading measures (Hackett, 1984).

It has been argued that the most demanding criterion of construct validity can be applied only when different methods or procedures are used to measure the same construct. Thus, a second measure of self-efficacy was required to examine convergent validity. As there were few psychometrically sound measures of reading efficacy for Form 2 (junior high school) children found in the literature, a second measure of reading efficacy, the General Reading Self-efficacy (GRSE) scale, was constructed. It was adapted from Pintrich and De Groot's (1990) measure of academic efficacy; the academic efficacy scale has been shown to have good psychometric properties.

In summary, the purposes of the second study were to (a) replicate the item analysis results and the reliability estimate of the TRSE scale, (b) examine the factor structure of the TRSE and GRSE, and (c) investigate convergent and concurrent validity.

METHOD

Subjects

The sample consisted of 139 randomly selected Form 2 (9th grade) students from four intermediate (junior high) schools in the Manawatu district. Subjects (45.3% males and 54.7% females) were aged approximately 12 years ($M = 12.48$ years, $SD = 0.32$). The majority of students were of European descent (87.1%) while 12.9% constituted non-Europeans (Maoris, Asians and others).

Measures

Efficacy

Although the TRSE was measured in line with Bandura's (1977) original definition "the conviction that one can successfully execute the behaviour required to produce particular outcomes", the second measure was measured on level with his more recent definition. This recent definition emphasizes that "self-efficacy is concerned with generative capabilities, and not with component acts" (Bandura, 1986, p. 397). Thus, the first measure assessed students' confidence about outcomes in relation to specific tasks in reading while the second assessed students' aggregated reading self-efficacy.

(i) *Task specific reading self-efficacy (TRSE) scale*

The six-item Task specific Reading Self-efficacy Scale (TRSE), followed Bandura's (1977) original methodology and was more specific in terms of reading tasks. Children's self-efficacy for reading different materials (e.g., an encyclopedia, a 150 page novel, a letter from a friend, etc.) was measured on a scale that ranged from 0 to 100 in 10-unit intervals from high uncertainty to complete certitude. A task specific self-efficacy score was obtained by calculating the mean.

(ii) *General reading self-efficacy (GRSE) scale*

Items were adapted from the self-efficacy subscale devised by Pintrich and De Groot (1990) and were modified to make it reading specific (see Appendix, Table A1). An attempt was made to incorporate many aspects/facets of reading, for example, doing well in reading, ability to understand the material, and ability to understand difficult words, etcetera. The items had the following stems "I am certain", "I am sure", and "I am confident". Items which indicated comparison to other students were excluded as it was believed that they measured perceived competence rather than self-efficacy. Students were required to respond to 7 items on reading self-efficacy using the following 7-point Likert-type response format ranging from 1 - *never*, 2 - *almost never*, 3 - *seldom*, 4 - *sometimes*, 5 - *often*, 6 - *almost always*, 7 - *always*. Scores for the 7 items were summed and then divided by 7 to produce a mean score ranging from 1 to 7. Scoring was reversed on the items that were negatively worded, so that for all items a high score indicated high self-efficacy. For the present study, this scale showed very high internal consistency, yielding a Cronbach's alpha of .91. All item scale correlations were greater than .60 (range .62 to .80).

Reading Metacognitive Knowledge

The Index of Reading Awareness (IRA), developed by Jacobs and Paris (1987), was used to assess children's metacognitive knowledge in reading. This measure comprises four separate subscales: evaluation, planning, regulation and conditional knowledge of reading. It consists of 20 items, each with three alternatives representing an inappropriate response (0 points), a partially adequate answer (1 point) and a strategic response (2 points). Scores for the 20 questions on the IRA were combined to produce a total score ranging from 0 to 40 points. This measure was based on empirical research of children's responses to metacognitive questions, so that it accurately reflected children's knowledge about reading strategies, rather than the authors' beliefs about what they know. Test-retest reliability (8 month interval) yielded a Pearson product-moment correlation of $r = .55$ ($p < .001$).

In the present study, the wording of some items was slightly changed so that the item reflected a measurement of metacognitive knowledge rather than strategy use. For instance, the IRA item of 'If you could only read some of the sentences in the story because you were in a hurry, which ones would you read?' was modified to 'If you could only read some of the sentences in the story because you were in a hurry, which ones *should* you read?' Analysis of internal consistency of the 20 items yielded a Cronbach's alpha of .62 for the present study.

Anxiety

The reading anxiety scale (RAS) was adapted from Pintrich and De Groot's (1990) test anxiety scale to create an instrument which is reading-specific and appropriate to junior high school children. The RAS consists of 11 items. Scoring of the negative items was reversed so that a high score indicated low anxiety. The scale yielded a high internal consistency of .93. An anxiety score was obtained by calculating the mean of the 11 items.

Perceived value

Eleven items, similar to those used by Pintrich and De Groot (1990) in their intrinsic value scale, were used to assess students' perceived value of reading. Whereas Pintrich and De Groot asked reactions to general academic work, students in this study were questioned specifically about reading. Items concerning the value of reading assessed perceived usefulness, importance, and inherent interest. The scale yielded a high internal consistency of .89. A mean score was calculated for this scale, with high scores reflecting high perceived value for reading.

Achievement

(i) Academic achievement

Students' academic achievement was assessed using three of the seven tests of the Progressive Achievement Test Series (PAT). The PATs that were used in the present study included Reading Comprehension (Reid & Elley, 1991), Reading Vocabulary (Reid & Elley, 1991), Reading Study Skills (Reid, Croft, & Jackson, 1978), and Mathematics (Reid, 1993). These tests are group administered, NZ normed, paper and pencil scales, administered by the majority of NZ primary and intermediate schools at the beginning of each school year (Beck and St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in NZ (Reid & Elley, 1991; Reid, 1993).

(ii) Reading comprehension performance

The latent construct of reading performance was assessed using two measures. Age percentile scores of a standardized measure of reading comprehension, the PAT Reading Comprehension, served as one indicator since a child's achievement is most often assessed with a standardized test of reading ability (German, Johnson, & Schneider, 1985). Classroom marks represented a second measure because they provide a very important source of feedback to students.

Teachers provided gradings, expressed as a percentage, for students' performance on different reading tasks in the classroom. These tasks fell into three general categories a) in-class seatwork; b) quizzes and tests and c) assignments and projects. An average score for the three categories was calculated with a high score representing good reading performance. The gradings were collected from the teachers during the later part of the year.

Procedure

In the early part of the academic year, students in each school were administered the self-efficacy, metacognitive knowledge, anxiety, and perceived value measures in groups of 15 - 20, during one 30 minute session. The measures were read out to each group by the researcher to ensure standardized presentation and to minimize reading difficulties. A short training sequence preceded the administration of each measure, to familiarize students with the use of the response format. The students were continuously observed in order to ensure that they were comprehending what was being read out and that they were attentive during testing.

Data analysis

Confirmatory factor analysis using the LISREL 8 program (Jöreskog & Sörbom, 1993) constituted the principal method of data analysis. This technique enables a series of hypothesized regression equations to be solved simultaneously to generate an estimated covariance matrix. In contrast to exploratory factor analysis, confirmatory factor analysis requires a priori specification of the items to its factors. Maximum likelihood techniques were used to generate the parameter estimates.

There are a variety of indices of fit which can be obtained for a given model. These indices may be used to determine the adequacy of fit of a single model or to compare the relative fit of a number of alternative competing models of the data. In general, no single index has been endorsed the "best index" by the majority of researchers (Gerbing & Anderson, 1993), so adequacy of fit is usually assessed on the basis of observation of several indices together. In the present study assessment of fit was based on multiple criteria including chi-square (Jöreskog & Sörbom, 1993), adjusted goodness of fit index (AGFI; Jöreskog & Sörbom, 1993), normed fit index (NFI; Bentler & Bonett, 1980), parsimony normed fit index (PNFI; Mulaik et al., 1989), and standardized root mean square residual (RMR; Jöreskog & Sörbom, 1993). In general, a chi-square probability value greater than .05, AGFI and NFI indexes exceeding .90, moderate values for PNFI, and standardized RMR of less than .05, are indicative of an acceptable model fit.

In the present study, confirmatory factor analyses of each self-efficacy scale were conducted initially to determine the factorial validity of the scales. A number of alternative models were tested ranging from the addition of correlated errors and a test of a two versus one factor solution. Following this, correlations between the two efficacy scales and the criterion measures of achievement, metacognitive knowledge, perceived value, and anxiety were calculated to provide construct validity.

RESULTS

Internal consistency

Since reliability estimates may be spuriously inflated when based on the same samples as the

item analyses, coefficient alpha was also calculated for the second sample. The resulting alpha for the TRSE was .91. All corrected item-scale correlations were greater than .40 (see Table 2).

Table 2. Item means and item-scale correlations of the TRSE scale

Item	<i>M</i>	item-scale correlation
letter from classmate	95.2	.47
school library book	86.8	.69
short fiction story	78.7	.82
short play	74.3	.79
150 page novel	62.2	.75
encyclopedia	37.6	.63

Examination of the means suggested that the items varied in difficulty. Although the range of means for the items varied between the two studies, the order of difficulty of the items was replicated. The item "letter from a classmate" represented the easiest item for the students while the item "encyclopedia" presented the most difficult.

Factor structure of the self-efficacy measures

In this section, three sets of confirmatory factor analyses were conducted to examine the factorial validity or dimensionality of the two reading efficacy scales. Initially, a number of alternative confirmatory measurement models of the task specific measure of efficacy (TRSE) was examined. This was followed by an examination of a confirmatory measurement model for the general measure of reading efficacy (GRSE). Finally, a third set of analyses was conducted to determine whether the general and task specific dimensions of efficacy were distinct.

One of the assumptions underlying the use of maximum likelihood estimation of latent variables is multivariate normal data distribution. All items for the TRSE and GRSE scales met requirements for univariate normality except for Item 1 of the TRSE scale which showed a mild negative skew and moderate positive kurtosis. The moderate kurtosis of item 1 was expected since it represented one of the easiest items. However, applying a square root or logarithm transformation to item 1 indicated that correlations between items 2-6 and the transformed variable of item 1 were similar to those obtained between items 2-6 and the untransformed item. Thus, item 1 did not warrant transformation. Tables A2 and A3 (see appendix) present the intercorrelations of TRSE and GRSE items, respectively. These correlations present in standardized form the covariance data used in the LISREL analyses.

Analyses of the TRSE scale

To identify the model, the variance of the latent task specific efficacy factor was fixed at unity, and the six items were freely estimated. Two CFA models were tested. In the first

model, it was hypothesized that the items that compose the TRSE would load on one factor, the task specific reading self-efficacy factor. This model did not assume correlated errors. The second model differed from the first in that it assumed there were correlated errors. This assumption is based on Bandura's methodology, which requires the inclusion of items with varying difficulty. It is therefore plausible that efficacy items reflecting easy tasks (e.g. items 1 and 2) may be correlated with each other and items reflecting more difficult items (items 4, 5, and 6) may be correlated with each other. Fornell (1983) suggests that allowing for correlated measurement errors should not be motivated by goodness-of-fit improvement, unless (a) it is warranted on theoretical or methodological grounds, or (b) it does not significantly modify the parameter estimates (see section on sensitivity analyses, below).

The first model provided a poor fit to the data, $\chi^2(9, N = 139) = 109.96$, $p = .00$; AGFI = .54; NFI = .83; RMR = .08; and PNFI = .50. Modification indices suggested that the fit of the model could be improved by allowing 3 correlated errors. The fit of the model was substantially improved by freeing these error covariances, $\chi^2(6, N = 139) = 6.66$, $p = .35$; AGFI = .95; NFI = .99; RMR = .02; and PNFI = .40. The normed fit index for this second model was .99, indicating that the vast majority of the variation was accounted for by the model. Each factor loading was highly significant ($t > 6.60$, $p < .001$), suggesting that the six items were representative of a single underlying construct. The standardized estimates of the factor loadings and the error variances-covariances are presented in Figure 1.

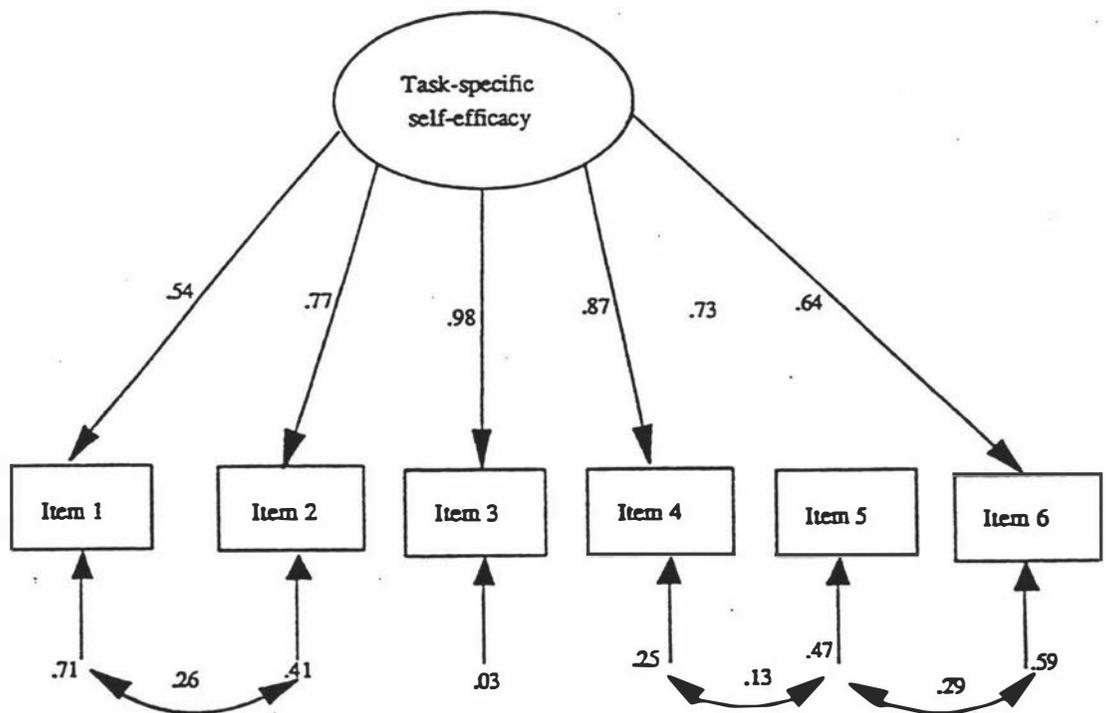


Figure 1. Completely standardized solution of a confirmatory factor model of task specific efficacy items.

Analysis of the GRSE

Even though the GRSE items were drawn from a previous study, CFA analysis of this scale

was considered necessary because the wording of the items was changed to make it reading specific. The variance of the latent general efficacy factor was fixed at unity, and the seven items were freely estimated. A model was tested where it was hypothesized that the items that compose the GRSE would load on one factor, the general reading self-efficacy factor. This model did not assume correlated errors. The model provided a good fit to the data, $\chi^2(14, N = 139) = 19.58$, $p = .14$; AGFI = .92; NFI = .97; RMR = .028; and PNFI = .64.

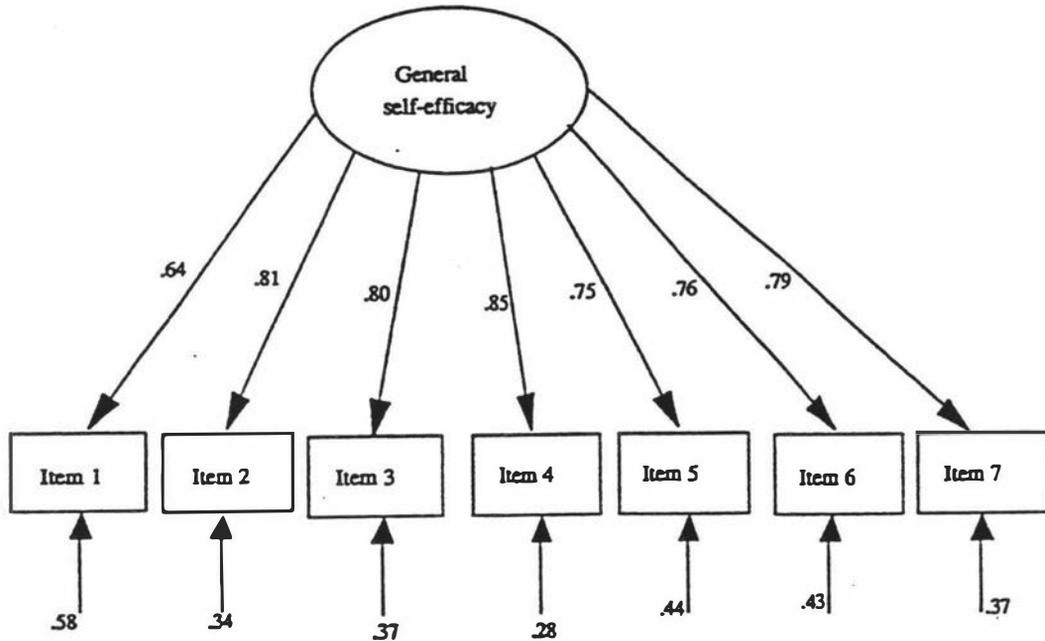


Figure 2. Completely standardized solution of a confirmatory factor model of general efficacy items.

The completely standardized solution of parameter estimates of this model is presented in Figure 2. Each factor loading was highly significant ($t > 8.21$, $p < .001$), suggesting that the seven items were representative of a single underlying construct. The moderate-to-high factor loadings of the indicators suggest that they are effective indicators in defining the latent variable. The normed fit index for this model was .97, indicating that the vast majority of the variation was accounted for by the model.

Test of a two-factor versus one-factor model

Given the above findings, that each efficacy scale is unidimensional, another model was tested to determine whether the task specific and general dimensions of reading efficacy were distinct. It was hypothesized that the TRSE and the GRSE items would fall on separate factors since different definitions were employed in devising each scale.

Following procedures adopted by Marsh and O'Neill (1984), items in each efficacy scale were combined to form item composites. Generally, item composites are formed (on a meaningful basis) based on conceptual, categorical, or task similarity. Grouping of items to test factorial validity is supported by a number of researchers (e.g., Byrne, 1989; Chapman & Tunmer, 1995; Marsh & O'Neill, 1984). Within the TRSE scale, three item pairs were formed based on their order of difficulty, that is, the first two items relating to the scale defined the first item pair, the next two items the second item pair, and so forth. For the GRSE scale, items 2, 3, and 7 which reflected "confidence on classroom tasks" formed one composite grouping.

Items 1 and 6, reflecting "confidence in understanding what was read" formed an item pair, whilst the remaining two items formed another item pair. Thus, for the GRSE scale, 2 item pairs and 1 composite grouping of 3 items were formed. As suggested by Marsh and O'Neill, this procedure was preferable to a confirmatory factor analysis of 13 items for the following reasons: (a) each variable should be more reliable and possess a smaller unique component, (b) the ratio of the number of subjects to the number of variables is increased, and (c) idiosyncratic wording of the individual items should have less affect on the factor loadings. All variables met conditions for multivariate normality. The correlations of the six composite items is provided in Table A4 in the appendix.

The 6 composite items were allowed to load freely within their respective factors. Factor variances were fixed at unity to identify the model, and the two factors were allowed to correlate freely. Again as in the case of the TRSE scale, the model assumed the presence of correlated errors. This confirmatory factor model, with the release of one correlated error, constituted a good fit to the data, $\chi^2(7, N = 139) = 8.61, p = .28$; AGFI = .94; NFI = .99; RMR = .02; and PNFI = .46. Two distinct dimensions assessing task specific and generalized efficacy for reading were identified.

Whether fewer than two factors are, in fact, adequate to account for variation among the data was evaluated by examining an alternative one factor structure, considering the high factor intercorrelation ($r = .83$). The measurement model was reestimated by specifying a perfect correlation between the task specific and general efficacy factors. This one factor model, which yielded $\chi^2(8, N = 139) = 54.13, p = .00$; AGFI = .71; NFI = .93; RMR = .06; and PNFI = .49, provided a poor fit to the data suggesting that even though the two efficacy factors are highly correlated, they represent distinct aspects/dimensions of reading efficacy. The completely standardized solution of parameter estimates of this model is presented in Figure 3.

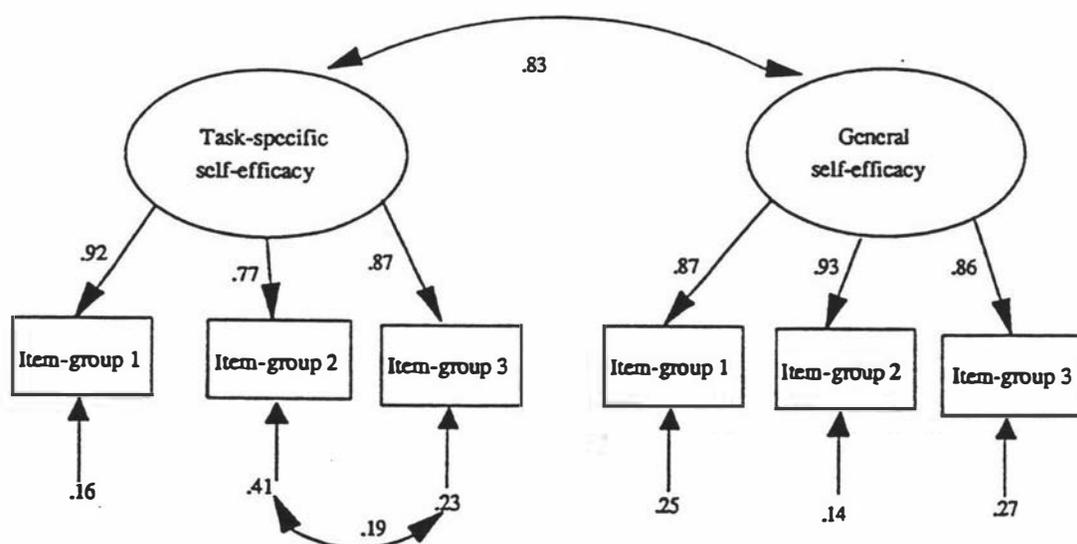


Figure 3. Completely standardized solution of a confirmatory two-factor model of task specific and general efficacy.

Sensitivity analyses

Although correlated errors of measurement were permitted for the models based on justifiable grounds, there remains the uncertainty that one may simply be capitalising on chance variation in the data. There is always the concern of overfitting the model; that is, fitting the model to trivial sample-specific artifacts in the data. Researchers have stressed the importance of investigating the sensitivity of parameter estimates to the specification of correlated errors of measurement (e.g., Hughes, Price, & Marrs, 1986; Tanaka & Huba, 1984). Sensitivity analyses (Byrne, 1989) were thus conducted to determine the practical significance of including additional parameters like correlated errors (i.e., their importance to the overall meaningfulness of the model). One way of determining this information is to test the sensitivity of the major parameters in the model to the addition of different patterns of error covariances. Results of the sensitivity analyses indicated that the imposition of correlated errors has minimal impact on the estimated parameters of the models, but does significantly improve model fit. This provides some reassurance that assumptions concerning the error structure of the data have no impact on the essential conclusions drawn from the fitted models of the TRSE and GRSE.

Construct and Discriminant Validity

Correlations with external criteria

For the remaining analyses, items were combined within each of the two self-efficacy factors. This was accomplished by forming means of all items within each factor. Correlations between these scales and other criterion measures were computed.

Validation of self-efficacy requires demonstrating theoretically consistent patterns of relationships between self-efficacy and external criteria. Consistent with a construct validation approach, this requires that a self-efficacy measure be highly correlated with criteria to which it is closely related logically (convergence) and less highly correlated with other criteria (divergence). This pattern is clearly seen in the correlations between the self-efficacy factors and the two academic achievement areas (see Table 3). Reading achievement and reading study skills were more strongly correlated with reading self-efficacy than Mathematics achievement. The significant moderate correlations between the efficacy measures with classroom marks obtained from teachers towards the end of the school year attest to the predictive validity of the efficacy scales.

Table 3. Correlations of efficacy measures with criterion measures

Criterion measures	task specific efficacy	general efficacy
metacognitive knowledge	.45	.41
anxiety	-.74	-.64
perceived value	.63	.54
PAT reading comprehension	.75	.61
PAT reading vocabulary	.65	.67
PAT reading study skills	.58	.56
PAT mathematics	.49	.47
classroom marks	.56	.47

all correlations were significant at $p < .001$ (2-tailed)

The correlations of the self-efficacy scales with metacognitive knowledge, anxiety, and perceived value for reading also define a pattern of correlations that supports the construct validity of interpretations based upon responses to the self-efficacy scales. Students with high efficacy are more knowledgeable about reading, experience low anxiety, and possess high perceived value for reading.

It is notable that as compared to the general measure of reading efficacy, the task specific measure had stronger correlations with the criterion measures. Also worthy of note is that the correlations between the two self-efficacy measures ($r = .77$) is higher than the correlation of each of the efficacy factors with the other measures.

Gender differences in the efficacy measures

Independent t tests were performed on both measures of efficacy for male ($n = 63$) and female ($n = 76$) students. Statistically significant differences were found between groups for both the task specific ($t = -2.88, p < .05$) and general ($t = -2.26, p < .05$) measures of efficacy. Interestingly, although there were no significant gender differences in their PAT reading comprehension measure, females had higher task specific ($M = 67.02, SD = 16.00$) and general ($M = 5.30, SD = .98$) scores than males (means are 58.61 ($SD = 18.69$) and 4.83 ($SD = 1.37$) for task specific and general efficacy, respectively).

DISCUSSION

This article describes the development and psychometric characteristics of the TRSE which was designed to assess self-efficacy of reading. Despite great interest in reading self-efficacy, relatively few validated tools have been developed for the assessment of this construct. The TRSE is unique in that it is derived from Bandura's self-efficacy methodology (Bandura, 1986) and is the only measure normed with young adolescent New Zealand children. This study offers some preliminary evidence for adequate reliability and construct validity of the newly developed TRSE scale.

Although a large number of items were generated, a 6-item scale was considered highly reliable and sufficiently comprehensive to measure the self-efficacy of young adolescent children. It is probable that a larger number of items (e.g., items 27, 28, see Table 1) in the originally devised scale may be relevant for children in upper grades as compared to those in lower grades. This is because older students would have developed the reading skills to read more difficult material. Another point to be noted is that although the TRSE items were developed for children in NZ, they can be used in other countries, particularly, if the classroom reading activities are consistent with those in other countries. Also items where the term NZ is used could potentially be replaced with the country of interest.

The TRSE also demonstrates good face validity because of its derivation from self-efficacy theory and methodology, which demands detailed task descriptions. The final version of the TRSE also has adequate content validity (provided by the reading specialists), although it did not sample exhaustively from the domain of reading. One of the advantages of the TRSE over most other self-efficacy measures is that it is a domain specific rather than a global indicator of self-efficacy. This is particularly important as there is evidence in educational research for the separation of academic and nonacademic self-perceptions, and even between different academic areas (Marsh, Smith, & Barnes, 1985).

The expected pattern of higher correlations of the TRSE score with the reading achievement measures (comprehension, vocabulary, and reading study skills) as compared to its correlation with the mathematics achievement measure provides support for the validity of the TRSE. Self-perception theorists (e.g., Shavelson & Bolus, 1982) argue that academic ability measures should be more highly correlated with self-perceptions in related areas than those in other areas. The significant correlation of the TRSE measure with classroom gradings suggests that the scale has promising predictive validity. Moreover, the TRSE was able to differentiate between good and poor readers, thus establishing discriminant validity of the scale.

Concurrent validation of the TRSE was supported by substantial correlations with criterion measures. The findings related to the relationships between the TRSE and other motivational-related constructs such as anxiety and perceived value were consistent with previous research (e.g., Hackett, 1985). That is, efficacy scores are moderately and positively related to perceived value but negatively related to anxiety. In addition, its significant correlation with metacognitive knowledge is consistent with previous research (Paris & Oka, 1986).

Convergent validity was supported by a statistically significant relationship between the TRSE and the general measure of reading efficacy (GRSE). Moreover, consistent with Bandura's (1986) theory, the task specific measure of efficacy had a higher correlation with the reading comprehension measure than the general measure of efficacy. When dealing with specific behaviours, more specifically worded items are likely to provide a more accurate estimation of an individual's efficacy (Sherer et al., 1982), and hence higher correlations with the achievement measure. However, even though task specific measures may be more precise, they will always require custom construction and will not be conducive to comparison across different reading curricula. In such a case, the GRSE scale may be more appropriate. In fact, it has been argued that the level of specificity of the efficacy measure should ultimately match the complexity of the performance measure with which it is being compared (Lent & Hackett, 1987).

As predicted, confirmatory factor analyses supported the hypothesis that task specific reading self-efficacy is unidimensional. These results are in accordance with other academic measures of efficacy (Pintrich & De Groot, 1990). In addition, this study demonstrates that distinctions can be made between different measures of reading self-efficacy, depending on the level of specificity of the efficacy measure. Although the correlation between the task specific and general efficacy factors was high ($r = .83$), the factors formed distinct constructs. Confirmatory factor analysis led to the interpretation of a two-factor solution, thus establishing factorial validity.

Consistent with previous research, there was a significant gender effect for efficacy; girls had higher reading self-efficacy than boys (Hay, 1994; Marsh et al., 1985). This finding is consistent with Burns' (1982) suggestion that gender differences in self-perception may be related to the frequent observation that elementary and intermediate aged boys appear to have more problems than girls. Burns also suggested that boys tend to perceive the classroom setting as being "female", a perception that is strengthened by the predominance of women teachers at the primary and junior high school levels. Consequently, boys may have greater difficulty in maintaining positive self-perceptions since they have to cope with failures in a "female" setting.

The results of the present study have important implications for education. As a valid self-report questionnaire for measuring reading efficacy in adolescents, the TRSE may be

important in intervention research. Changes in scores on the TRSE may assist in determining the effectiveness of interventions for less efficacious children in research or classroom settings. Task specific assessments of efficacy yield different information than more global measures of efficacy and this information may be useful to teachers concerned with monitoring student's efficacy across different reading tasks.

The TRSE may also be useful in a variety of future research. For instance, future research could investigate the relevance of other items in the originally compiled task specific scale for older children. Also research investigating the relationship between efficacy, metacognition, and other motivation-related variables is much needed, and may help illuminate important aspects of children's reading development including the mechanisms that lead to reading disability. Such theory based research concerning the relationships between reading efficacy, metacognition, and motivation may have been impeded by the lack of reliable and valid reading self-efficacy instruments.

There were a number of limitations to the present study which point toward several other areas of future research. First, the current study is limited by the sample size of the poor reader group and other special population diagnostic groups and thus precludes conclusive discriminant validity studies at this time. Second, the confirmatory factor analytic findings require cross validation with a second sample. Third, appropriate validity studies are needed to address the construct validity of the factor scores in special education populations (gifted, learning disabled, mildly mentally handicapped). Fourth, multitrait, multimethod studies using observer ratings (parent, teacher ratings), are essential to complement the findings reported herein. Fifth, the short- and long-term stability of the TRSE need to be assessed. Nonetheless, the results of this initial investigation support the reliability and validity of the TRSE in assessing self-efficacy of reading for adolescent children.

REFERENCES

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411-423.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1989a). Regulation of cognitive processes through perceived self-efficacy. *Developmental Psychology*, *25*, 729-735.
- Bandura, A. (1989b). *Multidimensional scales of perceived self-efficacy*. Unpublished test. Stanford, CA: Stanford University.
- Bandura, A., Adams, N. E., Hardy, A. B., & Howells, G. N. (1980). Tests of the generality of self-efficacy theory. *Cognitive Therapy and Research*, *4*, 39-66.
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, *41*, 586-598.
- Beck, L. R., & St. George, R. (1983). The alleged cultural bias of PAT Reading Comprehension and Reading Vocabulary tests. *New Zealand Journal of Educational Studies*, *18*, 32-47.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588-606.
- Berry, J. M., West, R. L., & Dennehey, D. M. (1989). Reliability and validity of the memory self-efficacy questionnaire. *Developmental Psychology*, *25*, 701-713.
- Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology*, *130*, 353-363.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and instruction* (pp. 53-92). Hillsdale, NJ: Erlbaum.
- Burns, R. (1982). *Self-concept development and education*. London: Holt, Rinehart, & Winston.
- Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Carr, M., Borkowski, J. G., & Maxwell, S. (1991). Motivational components of underachievement. *Developmental Psychology*, *27*, 108-118.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, *87*, 154-167.

- Cohen, S. G., McDonell, G., & Osborn, B. (1989). Self-perceptions of "at-risk" and high achieving readers: Beyond Reading Recovery achievement data. In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction, Thirty-eight yearbook of the National Reading Conference* (pp. 117-122). Chicago: National Reading Conference.
- Gerbing, D. W., & Anderson, J. C. (1987). Improper solutions in the analysis of covariance structures: Their interpretability and a comparison of alternative specifications. *Psychometrika*, *52*, 99-111.
- Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage Publications.
- German, D., Johnson, B., & Schneider, M. (1985). Learning disability versus reading disability: A survey of practioners' diagnostic populations and test instruments. *Learning Disability Quarterly*, *8*, 141-157.
- Gordon, C. (1990). Changes in readers' and writers' metacognitive knowledge: Some observations. *Reading Research and Instruction*, *30*, 1-14.
- Graham, S., & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology*, *83*, 187-194.
- Hackett, G. (1985). The role of mathematics self-efficacy in the choice of math-related majors of college women and men: A path analysis. *Journal of Counselling Psychology*, *32*, 47-56.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, *20*, 261-273.
- Hay, I. (1994). *The formation of self-perception and the interactions between academic self-perception and academic achievement of primary school students*. Unpublished doctoral dissertation, The university of Queensland, Queensland, Australia.
- Henk, W. A., & Melnick, S. A. (1992). The initial development of a scale to measure "Perception of self as reader". In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views from many perspectives*. Chicago: The National Reading Conference Inc.
- Henk, W. A., & Melnick, S. A. (1995). The reader Self-Perception Scale (RSPS): A new tool for measuring how children feel about themselves as readers. *The Reading Teacher*, *48*, 470-482.
- Horn, C., Bruning, R., Schraw, G., Curry, E., & Katkanant, C. (1993). Paths to success in the college classroom. *Contemporary Educational Psychology*, *18*, 467-478.
- Hughes, M. A., Price, R. L., & Marrs, D. W. (1986). Linking theory construction and theory testing: Models with multiple indicators of latent variables. *Academy of Management Review*, *11*, 128-144.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide*. Chicago: Scientific Software.
- Lent, R. W., & Hackett, G. (1987). Career self-efficacy: Empirical status and future directions. *Journal of Vocational Behavior*, *30*, 347-382.
- Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1993). Predicting mathematics-related choice and success behaviors: Test of an expanded social cognitive model. *Journal of Vocational Behavior*, *42*, 223-236.

- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Standard University Press.
- Marsh, H. W., & O'Neill, R. (1984). Self description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, 21, 153-174.
- Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multidimensional self-concepts: Relations with sex and academic achievement. *Journal of Educational Psychology*, 77, 581-596.
- Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology*, 16, 331-345.
- McMillan, J. H., Simonetta, L. G., & Singh, J. (1994) Student opinion survey: Development of measures of student motivation. *Educational and Psychological Measurement*, 54, 496-505.
- Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research*, 8, 107-120.
- Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stillwell, C. D. (1989). An evaluation of goodness of fit indices for structural equation models. *Psychological Bulletin*, 105, 430-445.
- Pajares, F. & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86, 193-203.
- Paris, S. G., & Oka, E. (1986). Children's reading strategies, metacognition and motivation. *Developmental Review*, 6, 25-56.
- Pereira-Laird, J. A. (1994). *A comparison of models of reading achievement in reading disabled versus normal achieving children using structural equation modelling techniques*. Manuscript in preparation.
- Pintrich, P. R., Anderman, E. M., & Klobucar, C. (1994). Intraindividual differences in motivation and cognition in students with and without learning disabilities. *Journal of Learning Disabilities*, 27, 360-370.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Randhawa, B. S., Beamer, J. E., & Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *Journal of Educational Psychology*, 85, 41-48.
- Reid, N. A. (1993). *Progressive Achievement Tests of Mathematics*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., Croft, A. C., & Jackson, P. F. (1978). *Progressive Achievement Tests: Study skills*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., & Elley, W. B. (1991). *Progressive Achievement Tests of Reading*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Schunk, D. H. (1984). Self-efficacy perspective on achievement behavior. *Educational Psychologist*, 19, 48-58.

- Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools, 22*, 208-223.
- Schunk, D. H. (1987). *Domain-specific measurements of students' self-regulated processes*. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review, 1*, 173-208.
- Schunk, D. H. (1991). Self-efficacy and achievement motivation. *Educational Psychologist, 26*, 207-231.
- Schunk, D. H., & Gunn, T. P. (1986). Self-efficacy and skill development: Influence of task strategies and attributions. *Journal of Educational Research, 79*, 238-244.
- Seegers, G., & Boekaerts, M. (1993). Task motivation and mathematics achievement in actual task situations. *Learning and Instruction, 3*, 133-150.
- Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology, 74*, 3-17.
- Shell, D. F., Murphy, C. C., & Bruning, R. H. (1989). Self-efficacy and outcome expectancy mechanisms in reading and writing achievement. *Journal of Educational Psychology, 81*, 91-100.
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacob, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports, 51*, 663-671.
- Spear-Swerling, L., & Sternberg, R. (1994). The road not taken: An integrative theoretical model of reading disability. *Journal of Learning Disabilities, 27*, 91-103, 122.
- Tanaka, J. S., & Huba, G. J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. *Journal of Personality and Social Psychology, 46*, 621-635.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3-25.
- Weiner, B. (1985). *Human motivation*. New York: Springer-Verlag.
- Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology, 49*, 287-293
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Author Note

This article is derived from a Doctoral dissertation completed by the first author under the supervision of the second author.

We would like to acknowledge the assistance of Mrs. Shannon Roache in her initial efforts with developing the scale.

Correspondence concerning this article should be addressed to Joyce A. Pereira-Laird, Department of Psychology, Massey University, Private Bag 11222, Palmerston North, New Zealand.

APPENDIX

Table A1. The GRSE items

-
1. I am not sure I will be able to understand most things I read.
 2. I am sure I can do an excellent job on the reading tasks given to me in class.
 3. I am sure I will receive a good mark for my reading comprehension test.
 4. I don't feel confident in doing well in reading.
 5. I am sure I will be able to understand most difficult words I come across.
 6. I am certain I can understand well the books that I read.
 7. I feel confident that I can read any book that is given to me by my teacher.
-

Table A2. Correlations between the TRSE items

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1						
Item 2	.675					
Item 3	.527	.761				
Item 4	.447	.656	.853			
Item 5	.428	.530	.708	.775		
Item 6	.391	.452	.630	.602	.773	

Table A3. Correlations between the GRSE items

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Item 1							
Item 2	.466						
Item 3	.530	.681					
Item 4	.567	.706	.658				
Item 5	.493	.626	.577	.656			
Item 6	.520	.574	.559	.635	.566		
Item 7	.492	.623	.656	.638	.567	.681	

Table A4. Correlations between the TRSE and GRSE composite items

	Composite items					
	Trse.1	Trse.2	Trse.3	Grse.1	Grse.2	Grse.3
Trse.1						
Trse.2	.665					
Trse.3	.785	.846				
Grse.1	.610	.534	.619			
Grse.2	.698	.557	.621	.767		
Grse.3	.599	.485	.561	.714	.772	

APPENDIX D

Pereira-Laird, J. A. (1996). *Development and validation of a self-report measure of the Reading Strategy Use (RSU) scale*. Manuscript submitted for publication.

ABSTRACT

This study describes the development and validation of the Reading Strategy use (RSU) scale which assesses cognitive and metacognitive strategy use. In addition, discriminant validity of the components of metacognition and cognitive strategy use was assessed. The first study describes the development of the scale. A final version of 12 metacognitive strategy use and 10 cognitive strategy use items showed acceptable internal consistency. In general, the RSU scale correlated as expected with criterion measures, providing initial support for the validity of the scale. In the second study, confirmatory factor analyses with an independent sample of 139 students confirmed the dimensional nature of the RSU scale. A high internal consistency estimate for the scale was again obtained for this sample. Confirmation of several predicted conceptual relationships between the RSU scale and other reading related factors (i.e., achievement, anxiety, perceived value, self-efficacy, and persistence) provided evidence of construct validity. In addition, it was demonstrated that metacognitive knowledge, metacognitive strategy use, and cognitive strategy use represent distinct constructs. The RSU scale shows high potential for use in classroom and research settings.

Metacognitive strategy use and metacognitive knowledge which are components of metacognition, as well as cognitive strategy use have been shown to be important predictors of academic performance and learning (Cross & Paris, 1988; Garner & Alexander, 1989; Pokay & Blumenfeld, 1990; Pressley, Borkowski, & Schneider, 1987; Paris, Wasik, & Turner, 1991). Despite the growing body of research supporting the position of metacognition and cognitive strategy use in reading comprehension, there continues to be disagreement surrounding their definition and operationalization (Kurtz, 1991; Reynolds & Wade, 1986). This could probably be due to the limited attention that has been devoted to the discriminant validity of these constructs. In addition, relatively little attention has been given to the construction and validation of strategy use questionnaires in the domain of reading, further impeding systematic theory development in the area of reading. The present study examines these two issues.

Several early definitions of metacognition have been proposed (Brown, 1980; Flavell, 1979). Two salient features emerge from these definitions and these have received wide acceptance. The first salient feature is that metacognition involves metacognitive knowledge or knowledge about thinking. Metacognitive knowledge consists of storable information about one's cognitive processes; for example, knowledge about strategies, and knowledge about when and where to use strategies. The second salient feature is that metacognition involves metacognitive strategy use, that is, it involves self-regulation or control of thinking. Metacognitive strategy use, also known as executive processes, involves planning, monitoring, and correcting one's on-line performance.

There has been considerable debate in the last decade or two about the distinction between knowledge about cognition and regulation of cognition. For instance, Brown and Palincsar (1982) commented that these two forms of metacognition are "closely related" and that "each supports the other recursively" (p. 1). They suggest that any attempt to separate the two constructs results in "oversimplification". On the other hand, Lawson (1984) argued that the term metacognition should only apply to metacognitive knowledge rather than to control processes. He sees the two dimensions as logically distinct, with metacognitive knowledge

resulting from the operation of executive control processes, but acknowledges that metacognitive knowledge is one source of influence on executive processes. Nevertheless, it is argued here that metacognitive knowledge involves executive control and as such both are interwoven components of metacognition. To argue theoretically that the components of metacognition are interwoven does not imply that research into metacognition will not distinguish between knowledge of cognition and control of cognition (van Kraayenoord, 1986). In fact, there are a number of researchers who have empirically confirmed the occurrence of both facets (Cornoldi, 1990; Cross & Paris, 1988; Schraw & Dennison, 1994).

The difficulties concerning metacognition also centre on conceptual and methodological issues relating to metacognition and establishing a clear causal relationship between the components of metacognition and cognitive performance. Kurtz (1991) suggested that a critical area in the concept of metacognition is its distinction from cognitive processes which are not metacognitive. A related issue in the literature concerns the differentiation of cognitive and metacognitive strategy use (St. George, van Kraayenoord, & Chapman, 1985). Even though these two types of strategy use have been distinguished empirically in a large number of research studies (e.g., Anthony, 1994; Pokay & Blumenfeld, 1990; Zimmerman & Martinez-Pons, 1990), some researchers still contend that they are indistinguishable (e.g., S.G. Paris, personal communication, 1995). Despite empirical evidence for the distinction of these two types of strategies in a variety of domains, research has yet to empirically distinguish these strategies in the domain of reading using confirmatory factor analytic procedures. This represented one goal of the present study.

The centrality of "strategies" for understanding text is widely acknowledged (Alexander & Judy, 1988; Brown, Bransford, Ferrara, & Campione, 1983; Garner, 1990). However, a definition of strategy is required since there is no consensus regarding its definition (Paris et al., 1991). Related discussion in the literature concerning differentiation of cognitive and metacognitive activity does little to clarify the central notion of a strategy. Gerber (1983) warned that the concept of strategy would become progressively obscured and lose its apparent explanatory power unless its meaning was clarified. The reader is referred to Pereira-Laird and Deane (1996) for more detailed discussion of this issue. For the purposes of the present study, Rowe and Rayford's (1984) definition of a strategy is adopted, where a strategy is defined as a *deliberate action* that readers take voluntarily to develop an understanding of what they read.

A variety of labels have been used by researchers across studies to describe the strategies that are used with textual material. Among other labels, these include "study strategies" (Brown et al., 1983), "learning strategies" (Alvermann & Moore, 1991; Pintrich & De Groot, 1990; Weinstein & Mayer, 1986), and "text-processing strategies" (Paris et al., 1991). Irrespective of the label, the reader uses a mixture of metacognitive strategies to monitor understanding (e.g., Baker & Brown, 1984; Brown et al., 1983; Garner, 1987; Paris et al., 1991) and cognitive strategies to learn, remember and understand text.

Although the essence of metacognition and cognitive strategy use is still being debated, there seems to be more consensus about the mental activities it constitutes (Haller, Child, & Walberg, 1988). The metacognitive strategies involve planning, monitoring, and regulation activities that take place before, during, and after any thinking act such as reading (Brown, Armbruster, & Baker, 1986; Paris & Lindauer, 1982; Pintrich & Schrauben, 1992). In contrast, cognitive strategies, refer to integrating new material with prior knowledge. Cognitive strategies that students use to acquire, learn, remember, retrieve and understand the material while reading include rehearsal, elaboration, and organizational strategies (Pintrich

& Schrauben, 1992; Weinstein & Mayer, 1986; Zimmerman & Martinez-Pons, 1986, 1988).

The purpose of the study was to examine a methodological issue which has been given scant attention in the literature: the measurement of children's reading strategies. A review of a subset of both commercially published as well as developmental or research instruments indicated that many of them concentrate on learning strategies or study skills (e.g., Di Vesta & Moreno, 1993). Hence, there is a need for psychometrically sound instruments that measure reading strategies.

Many previous investigations of strategy use have used extensive verbal interviews. Such procedures are restrictive in most research and classroom settings due to the amount of time and effort necessary to administer them (Schraw & Dennison, 1994). The purpose of the present research was to develop and validate an easily administered self-report questionnaire which measures cognitive and metacognitive strategy use independently.

Two issues were of particular importance in this study. First, the study investigated whether current conceptualizations of strategy use that discriminate between metacognitive and cognitive strategy use are valid. In line with this the present study developed a reading specific strategy use instrument. A related issue in the development of the strategy use instrument is the establishment of reliability and validity for the instrument. To this end, correlations between the strategy use constructs and construct-related criterion measures were examined. Strategy use can be partially validated by correlating strategy use with motivation, metacognitive knowledge, and achievement. Research has shown that students who are strategically oriented display less anxiety but greater persistence and intrinsic interest in their academic learning (Pintrich & De Groot, 1990). A number of studies have also shown that strategy use is significantly related to efficacy, metacognitive knowledge and achievement (Baker & Brown, 1984; Pressley et al., 1987; Pokay & Blumenfeld, 1990). In addition, research has demonstrated that good and poor (or reading disabled) readers differ in their use of cognitive and metacognitive strategies (Baker & Brown, 1984; Resnick & Klopfer, 1989; Wong, 1980).

A second issue addressed whether metacognitive/cognitive strategy use and metacognitive knowledge are discriminable constructs. Confirmatory factor analyses were performed to investigate the factorial and discriminant validity of the aforementioned constructs.

This article describes the development of the Reading Strategy Use scale (RSU). Two investigations are reported. The development of the scale is reported in detail in the first study, in addition to establishing preliminary validation of the scale. The second investigation provides further validation of the RSU scale in addition to discriminant validity of metacognitive knowledge, metacognitive strategy use, and cognitive strategy use using confirmatory factor analytic procedures.

STUDY 1: PILOT STUDY - DEVELOPMENT OF THE RSU SCALE

A preliminary pilot test was conducted to evaluate administration procedures and to begin collecting psychometric data about the items.

METHOD

Sample

The sample of 53 students comprised of 19 males and 35 females. Participants were obtained from 2 schools in Palmerston North, New Zealand. Parental and school consent were obtained for those who participated in the study. The subjects were aged between 10-13 years ($M = 12.4$ years, $SD = 1.6$ years) and consisted of 64.1% Caucasians, 27.8% Maoris and 8.1% of other racial groups. Subjects were classified as good ($n = 33$) or poor readers ($n = 21$) according to their age percentile scores on a standardised reading comprehension achievement test, described below. Good readers had cut off scores above the 35th percentile on a standardized reading comprehension test.

Instruments

Academic achievement

Students' achievement was assessed using three tests of the Progressive Achievement Test Series (PAT), namely Reading Comprehension (Reid & Elley, 1991), Reading Vocabulary (Reid & Elley, 1991), and Mathematics (Reid 1993). These tests are group administered, New Zealand normed, paper and pencil scales, administered by the majority of New Zealand primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in NZ (Reid, 1993; Reid & Elley, 1991).

Reading Metacognitive Knowledge

The Index of Reading Awareness (IRA), developed by Jacobs and Paris (1987), was used to assess children's metacognitive knowledge in reading. This measure comprises four separate subscales: evaluation, planning, regulation and conditional knowledge of reading. It consists of 20 items, each with three alternatives representing an inappropriate response (0 points), a partially adequate answer (1 point) and a strategic response (2 points). Scores for the 20 questions on the IRA were combined to produce a total score ranging from 0 to 40 points. This measure was based on empirical research of children's responses to metacognitive questions, so that it accurately reflected children's knowledge about reading strategies, rather than the authors' beliefs about what they know. Test-retest reliability (8 month interval) yielded a Pearson product-moment correlation of $r = .55$ ($p < .001$).

In the present study, the wording of some items was slightly changed so that the item reflected a measurement of metacognitive knowledge rather than strategy use. For instance, the IRA item of 'If you could only read some of the sentences in the story because you were in a hurry, which ones would you read?' was modified to 'If you could only read some of the sentences in the story because you were in a hurry, which ones *should* you read?' Analysis of internal consistency of the 20 items yielded a Cronbach's alpha of .62.

Mathematics Metacognitive Knowledge

A multiple-choice measure was devised to assess student's metacognitive knowledge about mathematics. It consists of 16 items, each with three alternatives representing an inappropriate response (0 points), a partially adequate answer (1 point) and a strategic

response (2 points). Scores for the 20 questions were combined to produce a total score ranging from 0 to 32 points. Reliability as measured by Cronbach alpha yielded an alpha of .75. Construct validity of the scale was established by a moderately strong correlation ($r = .58$) with a standardized mathematics achievement test. Test-retest reliability of the measure over a three week period was .72.

Attributions

Attributions were assessed using a modified version of the Krause Attributional Questionnaire (Krause, 1983). For the purpose of the present study, the items were made more reading specific. The attributional questionnaire assessed student's thoughts about the importance of ability, effort, luck, task interest, and task difficulty. A total of 26 questions were used, half dealing with success and half dealing with failure. The questions referred to experiences children are likely to encounter in school during reading (e.g., The teacher asked you to read a paragraph aloud to the class. You read it loudly and clearly without missing a word. Why did this happen? Each question was read aloud. The students were asked to choose the best of the six attributional explanations: competence and cleverness (which could be subsumed under ability), paying attention and trying hard (which could be subsumed under effort), task interest and difficulty of the material. The student placed a one (1) in front of the best choice and a two (2) in front of the second best choice.

The success and failure questions were scored separately. Two points were awarded for the first choice response and one point for the second choice response. A success score was obtained by using the formula effort + ability + task interest - task difficulty. Effort, ability and task interest were positively weighted because these choices suggest an internal attributional orientation. A failure score was obtained by using the formula effort + task interest - (task difficulty + ability). Attributional explanations of effort and task interest were considered to be adaptive attributional explanations for failure because of their focus on internality and controllability. Ability, on the other hand, was scored negatively because it considered to be a maladaptive attribution for failure. Students with a strong ability attribution for failure are less likely to put forth the effort necessary to learn to read. Consistent with the original scale, test-retest reliability over a three week period for the adapted scale was high ($r = .68$). The original measure has been widely used by a number of researchers (e.g., Carr, Borkowski, & Maxwell, 1991).

Effort management

An adaptation of Pintrich and De Groot's (1990) effort management scale was used to measure persistence. Whereas Pintrich and De Groot (1990) assessed students' persistence to general academic work, students in this study were questioned specifically about reading.

Students were instructed to respond to the nine items on a Likert scale in terms of their behaviour. The scale ranged from 1 to 7 with the following verbal descriptors: 1 - never, 2 - almost never, 3 - seldom, 4 - sometimes, 5 - often, 6 - almost always, 7 - always. A mean composite score was obtained by averaging the scores of the 9 items. Scoring was reversed on the four items that were negatively worded, so that for all items a high score indicated high persistence. Similar to Pintrich and De Groot's (1990) measure, the adapted measure used in this study had high internal consistency (alpha = .86) as measured by Cronbach's alpha.

Procedure

Initially an item pool of strategy use items were developed based on a literature review and items from studies using published and experimental instruments (e.g., Briggs, 1987; Wade, Trathen, & Schraw, 1990; Weinstein & Palmer, 1990). The list of cognitive strategies was based on the framework developed by Weinstein and Mayer (1986) and included rehearsal, organization, and elaboration strategies. Metacognitive strategies were based on the framework of Brown et al. (1986) and Pintrich and Schrauben (1992) and included planning, monitoring, and regulation strategies. Items were modified so they were reading specific and only included those that directly dealt with academic/study practices. A number of items were reworded to better reflect their respective conceptual definition. Care was taken to use language suitable for young adolescent children. This resulted in a pool of 30 metacognitive strategy items and 26 cognitive strategy items. The number of items was subsequently reduced to 26 metacognitive and 18 cognitive strategy items after the elimination of confusing or compound items (those containing more than one question or statement in the same item), items which were highly similar, and those that appeared to fit more than one category. Item wording and structure was modified following consultation with three principals and four heads of reading departments who checked the appropriateness of the items for use with their Form 2 (9th grade) students. The cognitive and metacognitive strategy items were randomly ordered in the questionnaire and 12 items were worded negatively to control for response set. Responses to each item were made on a 7-point Likert-type scale which ranged from "Never" (1) to "Always" (7).

Two questionnaires were administered to students in groups of approximately 17, irrespective of their reading ability. Testing took place during two 50 minute sessions with both questionnaires administered within a week. A short training session preceded each measure to familiarize students with the response format. The RSU questionnaire was administered first. Respondents were asked for feedback about clarity of instructions, item difficulty, and any confusing aspects of the questionnaire and its administration.

The second questionnaire examined children's attributions, metacognitive knowledge for reading and mathematics, and effort management strategies. Both questionnaires were read out to each group by the researcher for each session to ensure standardized presentation and to minimize reading difficulties. Students were instructed to raise their hands if they had difficulty in understanding the item or wanted a particular item repeated. The students were continuously observed in order to ensure that they were comprehending what was being read out and that they were attentive during testing. At the end of each questionnaire session, students were instructed to check that all items had been answered.

RESULTS AND DISCUSSION

Frequency means and standard deviations were calculated for each of the 44 strategy use items. Means ranged from 3.1 to 4.8 and standard deviations ranged from 1.1 to 1.8. On the basis of the univariate statistics and feedback following administration, items were eliminated if they, were considered ambiguous or difficult by students, had similar content, or had a restricted response range. This process resulted in the elimination of 10 items. The remaining 34 items were subjected to validity ratings by seven judges who were considered to be experts in the field of reading metacognition and cognition. They had either done considerable research in the area of metacognition or were involved in teaching metacognition. The seven judges were asked to categorise the strategies using the following descriptors: "C" for cognitive; "M" for metacognitive; "C?" for cognitive, but not a very good item; "M?" for metacognitive, but not a very good item; "?" for not sure what category item falls under.

Table 1 shows the mean percentage agreement data for the cognitive and metacognitive strategy use categories for those items which were retained.

Table 1. Reading Metacognitive and Cognitive Strategy Use Items - Ratings by Experts

ITEM	% agreement amongst raters		
	Item No	C	M
I read quickly through the whole passage to get the general idea before I read it thoroughly.	m1		100
I learn new words by relating/linking them with words which I already know.	c1	100	
When I find that a chapter in my book is hard to understand, I slow down my reading.	m2		100
I make an outline of what I am reading.	c2	100	
When I cannot read a word in the story, I skip it.	c3	71	14
I am able to decide between more important and less important information while reading.	m3	14	86
When I'm reading I stop once in a while and go over what I have read.	m4		100
To help me understand what I have read, I say it in my own words.	c4	86	14
I decide how difficult my reading passage is and then adjust the speed of my reading accordingly.	m5		100
I learn new words by picturing in my mind a situation which they occur.	c5	100	
I stop once in a while and ask myself questions to see how well I understand what I am reading.	m6		100
After reading something, I sit and think about it for a while to check my understanding.	m7	14	86
When I get lost while reading, I go back to the point where I first had trouble.	m8	28	72
When I read, I form pictures in my mind of the things I am trying to understand.	c6	100	
After I have been reading for a short time, the words stop making sense.	m9	14	86
When reading about something I try to link it to what I already know.	c7	86	14
I read critically or thoughtfully, that is while reading something, I judge what I am reading.	m10	14	86
When I read, I underline the main ideas.	c8	100	
When I find I do not understand something while reading I read it again and try to figure it out.	m11	28	72
When reading I learn by heart difficult words and ideas without understanding them.	c9	100	
When reading, I check how well I understand the meaning of the story by asking myself whether the ideas fit with the other information in the story.	m12		100
When some of the sentences that I am reading are hard I either give up or read only the easy words.	c10	86	14
I make simple charts, diagrams or tables to summarize information I am reading for a test.	c11*	100	
I find it hard to pay attention while reading.	m13*	14	86
I make use of diagrams or charts to help me understand what I am reading.	c12*	100	

* Items were deleted to optimize Cronbach alpha

The judges were told to base their categorisation of strategies on the theoretical definitions of cognitive and metacognitive strategy, supplied by the researcher. Only items where there was at least 70% interrater agreement regarding its category were retained. This resulted in the retention of 13 metacognitive and 12 cognitive items.

Exploratory analyses were conducted on each group of strategy use items in an effort to maximize internal reliability. As three items had poor item-total correlations (range = .10 to .15), they were deleted from further analyses resulting in 12 metacognitive and 10 cognitive items. Both subscales demonstrated reasonably high reliability; alpha coefficients were .70 and .72 for the metacognitive and cognitive subscales, respectively.

Validation of the RSU scale requires demonstrating theoretically consistent patterns of relationships between cognitive/metacognitive strategy use and external criteria. Construct validity was assessed by correlating the scales with a variety of achievement measures, mathematics and reading metacognitive knowledge measures, and effort management (persistence). Discriminant validity was assessed by performing *t* tests for strategy use between good and poor readers. Consistent with a construct validation approach, this requires that a strategy use measure be highly correlated with criteria to which it is closely related logically (convergence) and less highly correlated with other criteria (divergence). This pattern is clearly seen in the correlations between the cognitive/metacognitive strategy use measures and the three academic achievement areas (see Table 2). Reading achievement and reading vocabulary were more strongly correlated with cognitive/metacognitive strategy use than Mathematics achievement. A similar pattern was observed with the metacognitive knowledge measures.

Table 2. Correlations between RSU subscales and criterion measures

Criterion measures	RSU subscales (<i>n</i> = 54)	
	RSU-meta	RSU-cog
PAT Reading Comprehension	.30	.31
PAT Vocabulary	.38	.34
PAT Mathematics	.20 ^a	.18 ^a
Reading metacognitive knowledge	.29 ^a	.44 ^a
Mathematics metacognitive knowledge	.16	.17
Success attributions	.27	.36
Failure attributions	.23	.38
Persistence	.39	.50

All correlations were significant ($p \leq .05$) at the two-tailed level except for ^a

The correlations of the strategy use scales with attributions and persistence also define a pattern of correlations that supports the construct validity of interpretations based upon responses to the strategy use scales. Students who reported high cognitive and metacognitive strategy use adopted adaptive attributional patterns and persisted longer. Discriminant validity was demonstrated by the significant differences found between good and poor readers for cognitive ($t = 1.96, p = .05$) and metacognitive ($t = 2.27, p = .03$) strategy use.

This study offers some preliminary evidence for adequate reliability and validity of the newly

developed Reading strategy use (RSU) scale. Internal consistency estimates of reliability was high for both the cognitive and metacognitive subscales and this study provides initial support for psychometric properties of the RSU scale. Although this study is an initial attempt at developing the RSU scale, the study is limited by its small sample size. Consequently, a second study was conducted to replicate the findings and provide additional psychometric data for the RSU.

STUDY 2 - CONSTRUCT VALIDATION OF THE RSU SCALE

Study 2 reports confirmatory factor analysis for the final RSU scale using a random sample of 139 students. The purposes of the second study were to (a) replicate the item analysis results and the internal consistency estimates, (b) examine the factor structure of the RSU, (c) investigate its convergent and concurrent validity (i.e., convergence between children's ratings on the RSU and appropriate criterion measures) and (d) establish the construct and discriminant validity of metacognitive knowledge, metacognitive strategy use, and cognitive strategy use in the domain of reading.

METHOD

Sample

The sample consisted of 139 randomly selected Form 2 (9th grade) students from four junior high schools in the Manawatu district of New Zealand. Subjects (45.3% males and 54.7% females) were aged approximately 12 years ($M = 12.48$ years, $SD = 0.32$). The majority of students were of European descent (87.1%) while 12.9% constituted non-Europeans (Maoris, Asians and others).

Instruments

Achievement

Academic achievement

Students' academic achievement was assessed using four of the seven tests of the Progressive Achievement Test Series (PAT). The PATs that were used in the present study included Reading Comprehension (Reid & Elley, 1991), Reading Vocabulary (Reid & Elley, 1991), Reading Study Skills (Reid, Croft, & Jackson, 1978), and Mathematics (Reid, 1993). These tests are group administered, NZ normed, paper and pencil scales, administered by the majority of NZ primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in NZ (Reid, 1993; Reid & Elley, 1991).

Classroom reading performance

Teachers provided gradings, expressed as a percentage, for students' performance on different reading tasks in the classroom. These tasks fell into three general categories: in-class seatwork, quizzes and tests, and assignments and projects. An average score for the three

categories was calculated with high scores representing good reading performance. The gradings were collected from the teachers during the later part of the year.

Anxiety

The reading anxiety scale (RAS) was adapted from Pintrich and De Groot's (1990) test anxiety scale to create an instrument which is reading-specific and appropriate to intermediate aged children. The RAS consists of 11 items. Scoring of the negative items was reversed so that a high score indicated low anxiety. The scale yielded a high internal consistency of .93. A mean score was calculated for this scale

Perceived value

Eleven items, similar to those used by Pintrich and De Groot (1990) in their intrinsic value scale, were used to assess students' perceived value of reading. Whereas Pintrich and De Groot asked reactions to general academic work, students in this study were questioned specifically about reading. Items concerning the value of reading assessed perceived usefulness, importance, and inherent interest. Similar to the original scale, the adapted scale yielded a high internal consistency ($r = .89$). A perceived value score was obtained by calculating the mean.

Task-specific reading self-efficacy scale

The six-item Task-specific Reading Self-efficacy Scale (TRSE; Pereira-Laird & Deane, 1995), measured self-efficacy for reading. Children's self-efficacy for reading different materials (e.g., an encyclopedia, a 150 page novel, a letter from a friend, etc.) was measured on a scale that ranged from 0 to 100 in 10-unit intervals from high uncertainty to complete certainty. The scale has high internal consistency with a Cronbach's alpha of .91. The scale shows high validity as is demonstrated by its moderately high correlations with measures of reading comprehension ($r = .75$), perceived value ($r = .63$), and anxiety ($r = -.74$). More detailed psychometric properties of this scale is provided in Pereira-Laird and Deane (1995). A task-specific self-efficacy score was obtained by calculating the mean.

Metacognitive knowledge

Two measures were used to assess metacognitive knowledge for reading. The first measure, the Metacomprehension Strategy Index (MSI), developed by Schmitt (1988, cited in Schmitt, 1990) is a 25-item, four-option multiple-choice questionnaire with one option representing metacomprehension awareness (1 point) and the other three options representing inappropriate responses (0 points). Scores for the 25 items on the MSI were combined to produce a total score ranging from 0 to 25 points. It was developed by Schmitt to evaluate young adolescent students' awareness of metacomprehension strategies before, during, and after reading a narrative prose. The MSI assesses students' metacognitive awareness within six broad categories: purpose setting; predicting and verifying; previewing; using background knowledge; summarizing and applying fix-up strategies; and self-questioning. The MSI has been widely used by a number of researchers (Baumann, Seifert-Kessell, & Jones, 1987; Horrex, 1992; Lonberger, 1988).

The MSI has been reported to have an internal consistency value of .87 using the Kuder-Richardson Formula 20 (Lonberger, 1988). Statistically significant correlations were found between the MSI and the IRA ($r = .48, p \leq .001$). Furthermore, statistically significant correlations were found between the MSI and two comprehension measures commonly used to measure students' metacomprehension ability, an error detection task ($r = .50, p \leq .001$) and a cloze task ($r = .49, p \leq .001$).

Minor modifications were made to the wording of the MSI to improve the face validity of this questionnaire for Form 2 New Zealand children. The alterations acknowledged not only narrative prose reading, but also textbook/book reading. For example in item 2 (A), instead of 'Look at the pictures to see what the story is about', the item was changed to 'Look at the pictures or diagrams to see what the chapter is about'. Items 2 and 20 were deleted for the present study to improve the reliability of the scale. Internal consistency of the measure for the current study yielded a Cronbach alpha of .68.

The Index of Reading Awareness (IRA), developed by Jacobs and Paris (1987), was used as a second measure to assess children's metacognitive knowledge in reading. It is described in the pilot study.

Effort management (persistence)

As described in the pilot study.

Procedure

In the early part of the academic year, students in each school were administered a questionnaire containing self-report strategy use, metacognitive knowledge, perceived value, anxiety, self-efficacy, and effort management measures in groups of 15 - 20, during one 50 minute session. In the first section, self-report cognitive and metacognitive strategy use items were presented in a random order to avoid response set. The self-report strategy use measure was presented before the metacognitive knowledge items (section 2) so as not to cue students about potentially useful strategies. The third section contained the perceived value, anxiety, and effort management items which were presented in random order to control for potential response bias. Finally, the fourth section contained items measuring self-efficacy.

The measures were read out to each group by the researcher to ensure standardized presentation and to minimize reading difficulties. A short training sequence preceded the administration of each measure, in order to familiarize students with the use of the response format. The students were continuously observed in order to ensure that they were comprehending what was being read out and that they were attentive during testing.

A number of procedures were followed to improve the validity of the self-report data obtained from students (see Assor & Connell, 1992). Before the administration of the questionnaire, students were briefed on the reasons for conducting the study. In age appropriate terms students were informed about confidentiality of their results, use of grouped findings when results were reported, and that any answers they gave to the questions were acceptable as long as they were reporting what they really believed.

Data analyses

Latent-variable structural modelling using the LISREL 8 program (Jöreskog & Sörbom, 1993) constituted the principal method of data analysis. This technique enables a series of hypothesized regression equations to be solved simultaneously to generate an estimated correlation or covariance matrix.

In contrast to exploratory factor analysis where there are no preconceived notions regarding the structure of the data, confirmatory factor analysis (CFA) allows a priori designation of the plausible factor pattern from theoretical or previous empirical research and then explicitly tests it statistically in the observed data. CFA provides (a) the goodness-of-fit indices for testing the congruence of the factor model with the data, (b) the inferential statistics for testing the significance of factor intercorrelations and factor loadings, and (c) estimates of the uniqueness of each measure.

In the present study, all confirmatory factor analyses (CFA) were based on covariance matrices. Parameter estimates for the model specified were generated using maximum likelihood techniques. There are a variety of indices of fit which can be obtained for a given model. These indices may be used to determine the adequacy of fit of a single model or to compare the relative fit of a number of alternative competing models of the data. In general, no single index has been endorsed the "best index" by the majority of researchers (Gerbing & Anderson, 1993), so adequacy of fit is usually assessed on the basis of observation of several indices together. In the present study assessment of fit was based on multiple criteria including chi-square (Jöreskog & Sörbom, 1993), goodness of fit and adjusted goodness of fit indices (GFI, AGFI; Jöreskog & Sörbom, 1993), non-normed fit index (NNFI; Bentler, 1989), comparative fit index (CFI; Bentler, 1990), and standardized root mean square residual (RMR; Jöreskog & Sörbom, 1993). In general, a chi-square probability value greater than .05, GFI, AGFI, NNFI, and CFI indices exceeding .90, and standardized RMR of less than .05, are indicative of well fitting models (Bentler, 1990; Marsh, Balla, & McDonald, 1988).

In the present study, CFA analyses of the 22 strategy use items were conducted initially to determine the factorial validity of the scales. A number of alternative models were tested ranging from the addition of correlated errors and a test of a one factor solution. Following this, correlations between cognitive/metacognitive strategy use and the criterion measures of achievement, perceived value, anxiety, self-efficacy, and persistence were calculated. To address, the second goal of the present study, CFA analysis was conducted again to provide discriminant validity for the measures of cognitive/metacognitive strategy use and metacognitive knowledge.

RESULTS

Internal consistency

Since reliability estimates may be spuriously inflated when based on the same samples as the item analyses, coefficient alpha was also calculated for the second sample. The resulting alpha coefficients for RSU-cog ($r = .73$) and for RSU-meta ($r = .85$) were high. All corrected item-scale correlations were greater than .40.

One possible explanation for the variability in the reliability estimates for cognitive and metacognitive strategy use is that students may be more selective in their choice of cognitive

strategies than in their choice of metacognitive strategies. In other words, reliabilities for metacognitive strategies may be higher because their use would be less situation- or task-dependent than cognitive strategies.

Factor structure of RSU

A confirmatory measurement model of cognitive and metacognitive strategy use was examined. One of the assumptions underlying the use of maximum likelihood estimation of latent variables is multivariate normal data distribution. All items met the assumptions for both univariate and multivariate normality. (The correlations of the 22 items are available from the author by request).

To identify the model, the variances of the latent cognitive and metacognitive strategy use factors were fixed at unity, and the 22 strategy use items were freely estimated within their respective factors. Two CFA models were tested first. In the first model, it was hypothesized that the items that compose the RSU-meta scale would load on one factor, the metacognitive strategy use factor, and the items that compose the RSU-cog scale would load on a second factor, the cognitive strategy use factor. This model did not include correlated errors. However, it was assumed that correlated errors would be present. This assumption was based on Parkes' (1987) suggestion that data obtained from questionnaires are usually prone to being systematically distorted by generalized response biases. A second, and probably, more pertinent justification for the presence of correlated errors is based on metacognitive/cognitive theory and research. More specifically, it has been theorized and empirically demonstrated that cognitive strategy use may not be effective for performance without the concomitant use of metacognitive strategies (Borkowski et al., 1992; Pintrich & De Groot, 1990). Also, correlated errors would be expected within each category, as some items measured similar aspects. Thus, a second model was tested which allowed only significant correlated errors amongst the cognitive and metacognitive items.

As expected, the first model provided a poor fit to the data, $\chi^2(208, N = 139) = 367.86, p = .00$, GFI=.81, AGFI = .77, NNFI =.76, CFI = .79 and RMR = .08. Modification indices suggested that the fit of the model could be improved by allowing 19 correlated errors. The fit of the model was substantially improved by freeing these error covariances, $\chi^2(189, N = 139) = 205.34, p = .20$, GFI = .89, AGFI = .85, NNFI = .97, CFI = .98, and RMR = .06. The strong correlation between the strategy factors ($r = .71$) may suggest concept equivalence, so the next model tested this proposition. The measurement model was reestimated by specifying a perfect correlation between the metacognitive and cognitive strategy use factors. This one factor model, which yielded $\chi^2(189, N = 139) = 243.93, p = .01$, GFI = .86, AGFI = .82, NNFI = .91, CFI = .93, and RMR = .07, provided a poor fit to the data, suggesting that a single underlying dimension could not explain the variance in the data.

The above findings suggest that the two factor solution with correlated errors provided the best fit to the data. Although the AGFI and RMR indices were slightly lower than the expected levels, other goodness-of-fit indices suggested a good fit. The non-normed fit index for this second model was .97, indicating that the vast majority of the variation was accounted for by the model. In addition, the coefficient of determination for the observed measures was .94. This coefficient provides a generalized measure of the amount of variance in the observed measures accounted for by the measurement model. In this case the measurement model performs well, since it accounts for a high proportion of the variation in the observed

scores, suggesting that the reliability of the measurement model as a whole is very high.

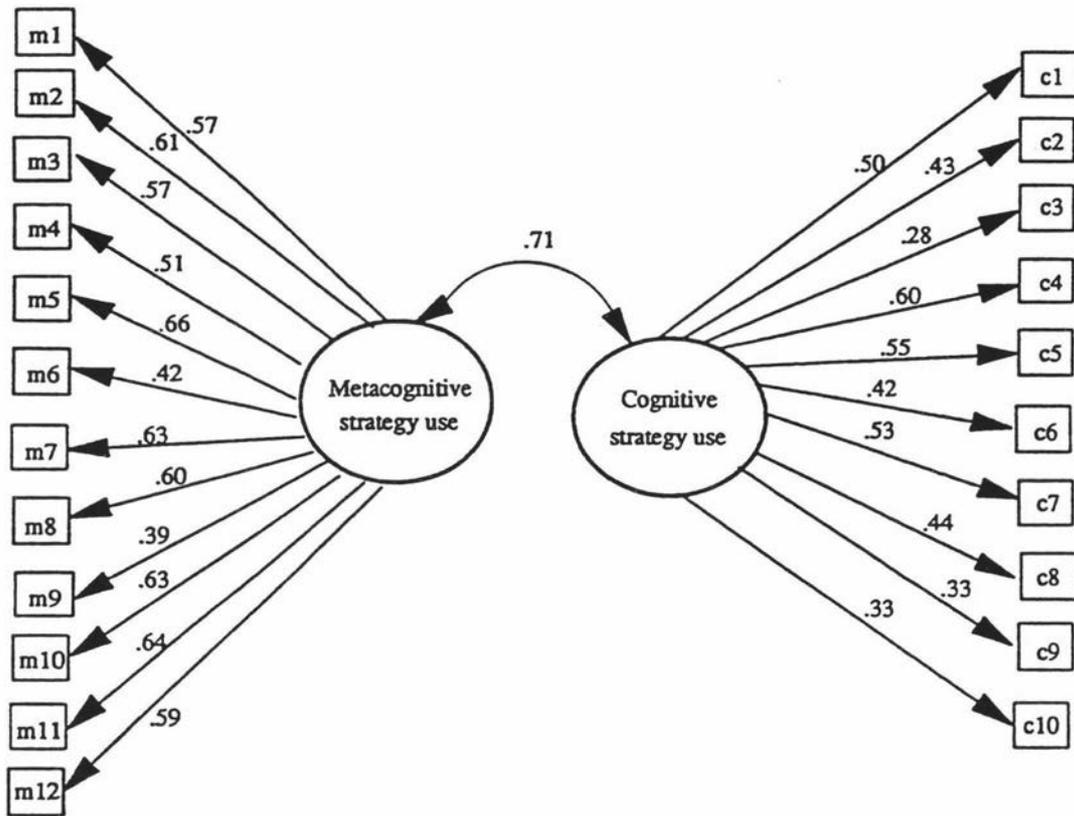


Figure 1. Completely standardized solution of a confirmatory factor model of cognitive and metacognitive strategy use items.

Note: m1 - m12 and c1 - c10 represent the metacognitive and cognitive items, respectively.

The moderate-to-high factor loadings of the indicators suggest that they are effective items in defining the latent variables. All construct loadings were statistically significant ($t > 2.24$, $p \leq .05$), and there was evidence for convergent validity. As the items did not load onto factors they were not designed to measure, they showed strong evidence of discriminant validity as well (Bollen, 1989). The completely standardized estimates of the factor loadings are presented in Figure 1. In keeping with the conventional notation of covariance structure modelling, in Figure 1, measured variables are enclosed in rectangles and latent variables in circles. Arrows relating circles to rectangles represent the latent variable's effects on indicators. Measurement errors were deleted from the figure to reduce clutter, but were estimated in the analysis. The completely standardized solution of these error variances and covariances are presented in Table A1 (see Appendix A). The t values for the variances and covariances of the measurement errors were generally considerably greater than 1.96 and so are statistically significant at $p \leq .05$ level. Thus, the observed scores in these samples were subject to measurement error, and they were also affected by generalized response tendencies, which resulted in nonrandom distribution of errors.

Although correlated errors of measurement were permitted for the second model based on justifiable grounds, there remains the uncertainty that one may simply be capitalising on

chance variation in the data. There is always the concern of overfitting the model; that is, fitting the model to trivial sample-specific artifacts in the data. Researchers have stressed the importance of investigating the sensitivity of parameter estimates to the specification of correlated errors of measurement (e.g., Hughes, Price, & Marrs, 1986; Tanaka & Huba, 1984). Sensitivity analyses (Byrne, 1989) were thus conducted to determine the practical significance of including additional parameters like correlated errors (i.e., their importance to the overall meaningfulness of the model). One way of determining this information is to test the sensitivity of the major parameters in the model to the addition of different patterns of error covariances. Results of the sensitivity analyses indicated that the imposition of correlated errors has minimal impact on the estimated parameters of the model, but does significantly improve model fit. This provides some reassurance that assumptions concerning the error structure of the data have no impact on the essential conclusions drawn from the fitted model.

Construct and Discriminant Validity

Correlations with external criteria

For the analyses in this section, the 10 cognitive and 12 metacognitive strategy use items were combined within each of the two factors. This was accomplished by forming means of all items within each factor.

The scores on the RSU-cog ($M = 11.57$, $SD = 1.98$) and the RSU-meta ($M = 17.81$, $SD = 3.84$) subscales were distributed normally. Examination of the data at the item level indicated that the mean frequency for each cognitive strategy item ranged from 3.54 to 5.41 ($SD = 1.38$ to 1.75) whilst the mean for each metacognitive strategy item ranged from 3.70 to 5.45 ($SD = 1.46$ to 1.83). The least frequently reported items were those of making an outline (item c2) and checking whether the ideas fitted with the other information in the story (item m12). Frequently reported items were giving up on difficult sentences and monitoring understanding (items c10 and m9).

Correlations between the RSU subscale scores and achievement, perceived value, anxiety, efficacy, and persistence were computed and are presented in Table 3. Due to listwise deletion, correlations between the RSU subscales and the criterion variables were based on a sample of 136. However, for PAT mathematics and PAT reading study skills, correlations were based on a sample of 118 due to some missing data. All correlations were significant, $p \leq .001$ (two-tailed).

Table 3. Correlations between RSU strategy scores and criterion constructs

Criterion measures	RSU-meta	RSU-cog
PAT Reading Comprehension	.68	.53
PAT Vocabulary	.59	.48
PAT Study Skills ^a	.57	.42
PAT Mathematics ^a	.36	.33
Classroom marks	.55	.41
perceived value	.50	.46
anxiety	-.59	-.59
self-efficacy	.63	.49
persistence	.48	.57

All correlations were highly significant ($p \leq .001$) at the two-tailed level.

Correlations between the RSU strategy use subscales was .5976.

^a Correlations were based on a sample of 118, all other correlations were based on a sample of 136.

Validation of cognitive and metacognitive strategy use requires demonstrating theoretically consistent patterns of relationships between strategy use and external criteria. Consistent with a construct validation approach, this requires that strategy use measures be highly correlated with criteria to which it is closely related logically (convergence) and less highly correlated with other criteria (divergence). This pattern is clearly seen in the correlations between the strategy use factors and the four standardized academic achievement areas (see Table 3). Reading comprehension, reading vocabulary, and reading study skills were more strongly correlated with cognitive/metacognitive strategy use than Mathematics achievement.

The correlations of the cognitive/metacognitive strategy use scales with motivational constructs, such as perceived value for reading, anxiety, self-efficacy, and persistence also define a pattern of correlations that supports the construct validity of interpretations based upon responses to the strategy use scales. Students who are more strategically engaged experience low anxiety, view reading positively, feel more efficacious, and employ more effort.

Predictive validity of the scales was obtained by correlating the RSU measures with classroom marks obtained from teachers towards the end of the school year. The significant moderate correlations attest to the predictive validity of the cognitive and metacognitive strategy subscales. Also worthy of note is that the criterion measures correlated more strongly with RSU-meta than with RSU-cog. This finding suggests that cognitive and metacognitive strategy use may be differentially related to achievement and motivation

Discriminant and construct validity of cognitive strategy use, metacognitive strategy use, and metacognitive knowledge

Conceptual distinctness of metacognitive strategy use, cognitive strategy use, and metacognitive knowledge was examined by using confirmatory factor analysis (CFA). The latent factor of metacognitive knowledge was represented by the two measures of metacognitive knowledge in reading, the MSI and the IRA.

Following procedures adopted by Marsh and O'Neill (1984), items in each subscale of the RSU were combined to form item composites. Although some researchers group items randomly (e.g., Marsh & O'Neill, 1984), item composites are usually formed on the basis of conceptual, dimensional, categorical, or task similarity. However, in the present case, due to the small number of items within each scale falling in any particular category, the items were grouped randomly. In the RSU-meta scale, the 12 items were combined to form 3 composite groupings of 4 items. However, since there were 10 items in the RSU-cog scale, items were combined to form 2 composite groupings of 3 items and 1 composite grouping of 4 items. The formation of composite grouping of items have been widely used by researchers to establish factorial and construct validity of instruments (Byrne, 1989; L.J. Horwood, personal communication, October, 1994; Marsh & O'Neill, 1984).

As suggested by Marsh and O'Neill (1984), this procedure was preferable to confirmatory factor analysing 22 items for the following reasons: 1) each variable should be more reliable and possess a smaller unique component, 2) the ratio of the number of subjects to the number of variables is increased, and 3) idiosyncratic wording of the individual items should have less affect on the factor loadings. All variables met conditions for multivariate normality. The variance-covariance matrix for the six composite strategy use variables and two

metacognitive knowledge variables was used to analyze the data. The correlations of these eight variables are presented in Table A2 (see Appendix A).

A model with three factors, viz., metacognitive strategy use, cognitive strategy use, and metacognitive knowledge, was assumed to explain the underlying structure of the eight variables. The eight variables were allowed to load freely within their respective factors. Factor variances were fixed at unity to identify the model, and the three factors were allowed to correlate freely. Again as before, the model assumed the presence of correlated errors; and only one significant error covariance arose. However, this error covariance was not included because the model fitted equally well without it. This confirmatory factor model, constituted a good fit to the data, $\chi^2(17, N = 139) = 16.36, p = .50, GFI = .97, AGFI = .94, NNFI = 1.00, CFI = 1.00,$ and $RMR = .03$. These results support a three-factor structure; that is, three distinct dimensions of metacognitive strategy use, cognitive strategy use, and metacognitive knowledge for reading were identified.

A second model was tested to investigate whether a two factor solution could explain the data better. It was demonstrated earlier that metacognitive and cognitive strategy use were distinct constructs. Thus the next model reestimated the measurement model by constraining the factor correlations between each of the strategy use factors and metacognitive knowledge to 1.0. This resulted in a significant deterioration in fit, $\chi^2(19, N = 139) = 29.54, p = .06, GFI = .95, AGFI = .91, NNFI = .96, CFI = .97,$ and $RMR = .07$. Hence, this model was rejected in favour of the three-factor model.

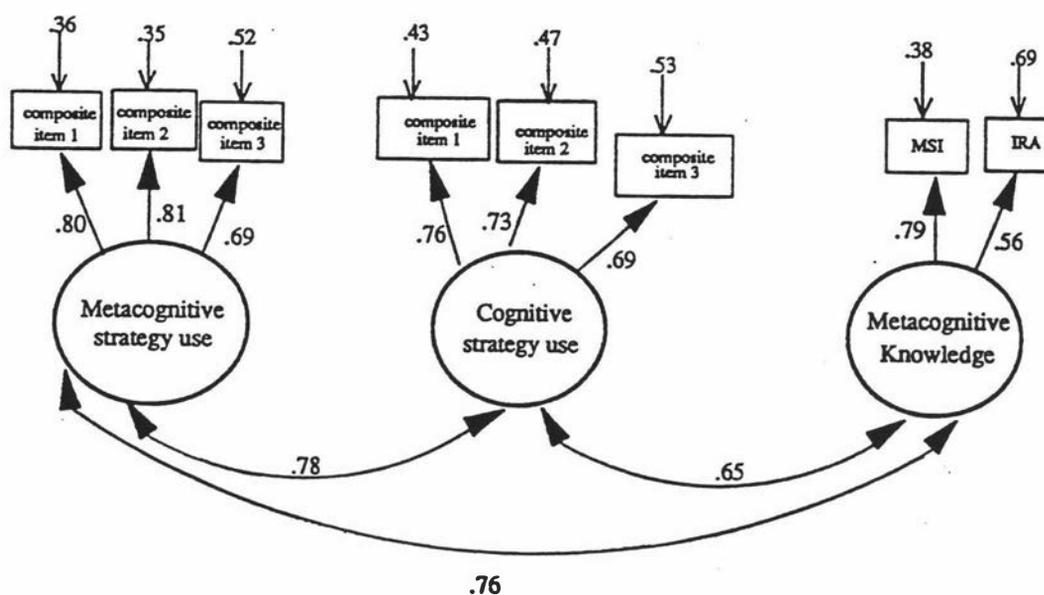


Figure 2. Completely standardized solution of a confirmatory factor model of cognitive strategy use, metacognitive strategy use, and metacognitive knowledge.

The completely standardized estimates of the three-factor model are presented in Figure 2. The coefficient of determination for the observed measures in the three-factor model was .97,

suggesting that the reliability of the measurement model as a whole is very high. The factor correlations between cognitive and metacognitive strategy use, cognitive strategy use and metacognitive knowledge, and metacognitive strategy use and metacognitive knowledge are .78, .65, and .76, respectively. The results of the alternative models suggest that while the latent factors of cognitive strategy use, metacognitive strategy use, and metacognitive knowledge are strongly and significantly correlated, they demonstrate discriminant validity in the sense that they are clearly different constructs.

DISCUSSION

This article describes the development and psychometric characteristics of the RSU scale which was designed to assess both metacognitive and cognitive strategies in the domain of reading. Despite great interest in reading strategy use, relatively few validated tools have been developed for the assessment of this construct. This study offers some preliminary evidence for adequate reliability and construct validity of the newly developed RSU measure.

Although a reasonably large number of items were generated to measure the strategy use of young adolescent children, it is probable that other types of reading strategies may be relevant for children in upper grades as compared to those in lower grades. For instance, strategies involving the use of simple charts, diagrams, or tables to summarize information may be more pertinent to older secondary school children who have more difficult textual material. It was demonstrated in the first study that although the "summarization by employing tables" strategy received complete interrater agreement as a cognitive strategy, it had to be deleted from the scale to increase reliability (because the strategy was rarely used). Intermediate (junior high) students may not find the aforementioned strategy useful as their reading material may not warrant the use of such strategies or it may be that these students may not have developed the skill to employ it. Research has demonstrated that the utilization of particular strategies comes with age (Kobasigawa, Ransom, & Holland, 1980).

One of the advantages of the RSU over most other strategy use measures is that it is a domain specific rather than a global indicator of strategy use. This is particularly important as there is evidence in educational research for the separation of strategy use between different academic areas (Alexander & Judy, 1988; Lester, 1988). In addition, many researchers and theoreticians emphasize that both generic and domain- or content-dependent strategies must be identified and taught (e.g., Gagne, 1985; Glaser, 1984). Another advantage is that the RSU is one of the few easily administered measures of reading strategy use which has been normed with young adolescent children.

CFA results suggest that strategy use is not a single trait, but is composed of two dimensions: cognitive and metacognitive strategy use. These dimensions are theoretical consistent with Borkowski et al.'s (1992) theory. These results are consistent with previous research which has used other academic measures of strategy use (Anthony, 1994; Pokay & Blumenfeld, 1990; Pressley et al., 1987, 1991; Zimmerman & Martinez-Pons, 1990).

Concurrent validation of the RSU was supported by significant correlations with criterion measures. The relationships between the RSU and motivational-related constructs such as anxiety, perceived value, self-efficacy, and persistence were consistent with previous research (Ames & Archer, 1988; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990; Tobias, 1985). That is, metacognitive and cognitive strategy use are moderately and positively related

to perceived value, efficacy, and persistence but negatively related to anxiety.

The expected pattern of higher correlations of the RSU scores with the reading achievement measures (comprehension, vocabulary, reading study skills) as compared to its correlation with the mathematics achievement measure provides support for the discriminant validity of the RSU. It can be argued that strategy use measures should be more highly correlated with achievement in related areas than those in other areas. The statistically significant relationship between metacognitive/cognitive strategy use and achievement is consistent with previous research (Brown et al., 1983; Pressley et al., 1990; Pressley, Symons, Snyder, & Cariglia-Bull, 1989).

Finally, the subscales seem to show promising predictive validity. The scales were related to classroom reading performance in the expected directions, as they were with the standardized reading achievement tests. These magnitude of the correlations with reading achievement appeared acceptable, given the multiplicity of factors that are related to reading achievement.

Discriminant validity for the components of metacognition was also demonstrated. The findings of this study reveal that although these components of metacognition are related, they form distinct constructs. In addition, the confirmatory factor analytic findings provide support for the distinction between components of metacognition and cognitive strategy use. These findings are consistent with previous research that have empirically distinguished these constructs (Anthony, 1994; Cornoldi, 1990; Cross & Paris, 1988; Pokay & Blumenfeld, 1990; Schraw & Dennison, 1994; Zimmerman & Martinez-Pons, 1990).

The results of the present study have important implications for education. First, the strong relationship between reading strategy use and motivation indicate that teachers may need to pay attention to motivation when assessing strategy use. Both theory and research predict that negative self-perceptions and expectations may hinder the deployment of strategies (Borkowski et al., 1992). Students' motivational factors are important because strategic reading depends on a positive view of one's competence and expectations for future achievements.

Second, as a valid self-report instrument for measuring reading strategy use in adolescents, the RSU scale may have high utility in intervention research. Research has shown that teaching reading disabled or poor readers to use strategies often produces positive results (Borkowski, Johnston, & Reid, 1987; Oka & Paris, 1987; Pressley, El-Dinary, & Brown, 1992). The information gleaned from the RSU could be used to identify students' level of strategic engagement and to evaluate the kinds of cognitive and metacognitive strategies students employ in reading particular texts. In addition, changes in scores on the RSU may assist in determining the effectiveness of interventions for less strategically oriented children in research or classroom settings. Responses to specific questions in the RSU may be useful for the design of individualized treatment programmes for adolescents.

There were a number of limitations to the present study. First, although the RSU can be used effectively to measure student perceptions of strategy use, it has a number of limitations. As suggested by Pokay and Blumenfeld (1990) they tap students' perceptions of what they are doing rather than the accuracy of these perceptions, and do not account for the efficiency or appropriateness of strategy use. Appropriateness of strategy use was not accounted for because students report strategies they use generally, rather than strategies they use to solve

specific tasks. Thus, responses of students in questionnaires refer only to student perceptions and not necessarily to accuracy, effectiveness, or appropriateness of strategy use.

Second, the current study may also be limited by the sample size particularly in relation to the number of items that were used. An ideal ratio would be 10 subjects to each item. Third, while the inclusion of correlated errors had minimal impact on the estimated parameters of the measurement model, it is important and necessary to cross validate the findings of the present study using a much larger sample. Fourth, while this study focuses on the quantity of cognitive and metacognitive strategies used by readers, analysing the type and quality of strategies reported by different groups of students should produce more information concerning interactions of strategy use variability with reading achievement. Finally, multitrait, multimethod studies using non-self-report strategy use measures (e.g., parent and teacher ratings or observations) are essential to complement the findings herein. Nonetheless, the results of this initial investigation provide preliminary support for the reliability and validity of the RSU scale in assessing cognitive and metacognitive strategy use for reading in adolescents.

REFERENCES

- Alexander, P. A., & Judy, J. E. (1988). The interaction of domain specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.
- Alvermann, D. E. & Moore, D. W. (1991) - Secondary school reading. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research: Volume II* (pp. 951-983). New York: Longman.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267.
- Anthony, G. (1994). Learning strategies in the mathematics classroom: What can we learn from stimulated recall interviews? *New Zealand Journal of Educational Studies, 29*, 127-140.
- Assor, A., & Connell, J. P. (1992). The validity of students' self-reports as measures of performance affecting self-appraisals. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 25-47). Hillsdale, NJ: Erlbaum.
- Baker, L. & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, M. Kamil, R. Barr & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- Baumann, J. F., Seifert-Kessell, N., & Jones, L. (1987, December). *Effects of think-aloud instruction on elementary students' ability to monitor their comprehension*. Paper presented at the National Reading Conference, St. Petersburg, FL.
- Beck, L. R., & St. George, R. (1983). The alleged cultural bias of PAT Reading Comprehension and Reading Vocabulary tests. *New Zealand Journal of Educational Studies, 18*, 32-47.
- Bentler, P. M. (1989). *EQS structural equations program manual*. Los Angeles: BMDP Statistical Software.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.
- Borkowski, J. G., Day, J. D., Saenz, D., Dietmeyer, D., Estrada, T. M., & Groteluschen, A. (1992). Expanding the boundaries of cognitive interventions. In B. Y. L. Wong (Ed.), *Contemporary intervention and research in learning disabilities: An international perspective* (pp. 1-21). New York: McGraw-Hill.
- Borkowski, J. G., Johnston, M. B., & Reid, M. K. (1987). Metacognition, motivation, and controlled performance. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol 2, pp. 147-173). Hillsdale, NJ: Erlbaum.
- Bollen, K. A. (1989). *Structural equations with latent variables*. NY: Wiley.
- Briggs, J. (1987). *Learning Process Questionnaire (LPQ): User's manual*. Melbourne: Australian Council for Educational Research.
- Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. B. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum.

- Brown, A. L., Armbruster, B. B., & Baker, L. (1986). The role of metacognition in reading and studying. In J. Oransanu (Ed.), *Reading comprehension: From research to practice*. Hillsdale, NJ: Erlbaum.
- Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of child psychology: Cognitive development* (Vol 3, pp. 77-166). New York: Wiley.
- Brown, A., & Palincsar, A. (1982). Inducing strategic knowledge from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities*, 2, 1-18
- Byrne, B. M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analytic models*. New York: Springer-Verlag.
- Carr, M., Borkowski, J. G., & Maxwell, S. E. (1991). Motivational components of underachievement. *Developmental Psychology*, 27, 108-118.
- Cornoldi, C. (1990). Metacognitive control processes and memory deficits in poor comprehenders. *Learning Disability Quarterly*, 13, 245-255.
- Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, 80, 131-142.
- Di Vesta, F. J., & Moreno, V. (1993). Cognitive control functions of study activities: A compensation model. *Contemporary Educational Psychology*, 18, 47-65.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Gagne, E. D. (1985). *The cognitive psychology of school learning*. Boston: Little Brown.
- Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research*, 60, 517-529.
- Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and Unanswered questions. *Educational Psychologist*, 24, 143-158.
- Gerber, M. M. (1983). Learning disabilities and cognitive strategies: A case for training or constraining problem solving? *Journal of Learning Disabilities*, 16, 255-260.
- Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage.
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39, 93-104.
- Haller, E. P., Child, D. A. & Walberg, H. J. (1988). Can comprehension be taught? A quantitative synthesis of "metacognitive" studies. *Educational Researcher*, 17, 5-8.
- Horrex, J. (1992). *A multicomponential intervention programme for remediating secondary students' reading comprehension difficulties* (Research Affiliateship Scheme Report No. 2). Palmerston

North, New Zealand: Massey University, Department of Education.

Hughes, M. A., Price, R. L., & Marrs, D. W. (1986). Linking theory construction and theory testing: Models with multiple indicators of latent variables. *Academy of Management Review*, *11*, 128-144.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, *22*, 255-278.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's reference guide*. Chicago: Scientific Software International.

Krause, A. (1983). *Attribution, intrinsic motivation, and metamemory: Determinants of strategic behavior*. Unpublished master's thesis. University of Notre Dame, Notre Dame, IN.

Kobasigawa, A., Ransom, C. C. & Holland, C. J. (1980). Children's knowledge about skimming. *Alberta Journal of Educational Research*, *26*, 169-182.

Kurtz, B. E. (1991). Cognitive and metacognitive aspects of text processing. In G. Denhière and J. P. Rossi (Eds.), *Text and text processing* (pp. 77-103). Amsterdam: Elsevier.

Lawson, M. (1984). Being executive about metacognition. In J. R. Kirby (Ed.), *Cognitive strategies and educational performance*. Orlando: Academic Press.

Lester, F. K. (1988). Reflections about mathematical problem-solving research. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 115-124). Hillsdale: Erlbaum.

Lonberger, R. (1988, February). *Effects of training in a self-generated learning strategy on the prose processing abilities of 4th and 6th graders*. Paper presented at the annual meeting of the Eastern Education Association, Savannah, GA.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analyses: The effect of sample size. *Psychological Bulletin*, *103*, 391-410.

Marsh, H. W., & O'Neill, R. (1984). Self description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, *21*, 153-174.

Oka, E. R., & Paris, S. G. (1987). Patterns of motivation and reading skills in underachieving children. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 115 - 145). Hillsdale, NJ: Erlbaum.

Paris, S. G., & Lindauer, B. K. (1982). The development of cognitive skills during childhood. In B. Wolman (Ed.), *Handbook of developmental psychology* (pp. 333-349). Englewood Cliffs, NJ: Prentice-Hall.

Paris, S. G., & Wasik, B. A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. Mosenthal & P. D. Pearson (Eds.), *Handbook of reading research: Volume II* (pp. 609-640). New York: Longman.

Parkes, K. R. (1987). Field dependence and the differentiation of neurotic syndromes. In P. Cuttance & R. Ecob (Eds.), *Structural modelling by example: Applications in educational, sociological, and behavioral research*. New York: Cambridge University Press.

- Pereira-Laird, J. A. & Deane, F. P. (1996). *Development of a behavioral measure of strategy use for adolescents - Think-Aloud Reading Strategy Use Measure*. Manuscript in preparation.
- Pereira-Laird, J. A. & Deane, F. P. (1995). *Validation of a Reading Self-efficacy scale for adolescents*. Manuscript submitted for publication.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.
- Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 149-183). Hillsdale, NJ: Erlbaum.
- Pokay, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology, 82*, 41-50.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 4, pp. 89-129). Greenwich, CT: JAI Press.
- Pressley, M., El-Dinary, P. B., & Brown, A. L. (1992). Skilled and not-so-skilled reading: good information processing and not-so-good information processing. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 91-127). New York: Academic Press.
- Pressley, M., Gaskins, I. W., Cunicelli, E. A., Burdick, N. J., Schaub-Matt, M., Lee, D. S., & Powell, N. (1991). Strategy instruction at benchmark school: A faculty interview study. *Learning Disability Quarterly, 14*, 19-48.
- Pressley, M., Symons, S., Snyder, B. L., & Cariglia-Bull, T. (1989). Strategy instruction research comes of age. *Learning Disability Quarterly, 12*, 16-30.
- Pressley, M., Woloshyn, V., Lysynchuk, L.M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review, 2*, 1-58.
- Reid, N. A. (1993). *Progressive Achievement Tests of Mathematics*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., Croft, A. C., & Jackson, P. F. (1978). *Progressive Achievement Tests: Study skills*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., & Elley, W. B. (1991). *Progressive Achievement Tests of Reading*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Resnick, L., & Klopfer, L. (1989). Toward the thinking curriculum: An overview. In L. Resnick & L. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 1-18). Alexandria, VA: Association for Supervision and Curriculum Development.
- Reynolds R. E. & Wade, S. E. (1986). Thinking about thinking about thinking: Reflections on metacognition. *Harvard Educational Review, 56*, 307-317.
- Rowe, D. W., & Rayford, L. (1984). *Instantiating background knowledge in reading comprehension*

- assessment*. Paper presented at the annual meeting of the National Reading Conference, St. Petersburg Beach, FL.
- Schmitt, M. C. (1990). A questionnaire to measure children's awareness of strategic reading processes. *The Reading Teacher*, 43, 454-461.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.
- St. George, A., van Kraayenoord, C. E., & Chapman, J. W. (1985). *Cognition and reading of learning disabled and normally achieving intermediate school children* (Research contract No. 55-22-121). Wellington, New Zealand: Massey University, Department of Education.
- Tanaka, J. S., & Huba, G. J. (1984). Confirmatory hierarchical factor analyses of psychological distress measures. *Journal of Personality and Social Psychology*, 46, 621-635.
- Tobias, S. (1985). Test anxiety: Inference, defective skills, and cognitive capacity. *Educational Psychologist*, 20, 135-142.
- van Kraayenoord, C. E. (1986). *Metacognition, reading and causal attributions: A comparison of learning disabled and non learning disabled intermediate school children*. Unpublished doctoral dissertation, Massey University, Palmerston North, New Zealand.
- Wade, S. E., Trathen, W., & Schraw, G. (1990). An analysis of spontaneous study strategies. *Research Reading Quarterly*, 25, 147-166.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315-327). New York: MacMillan.
- Weinstein, C. E., & Palmer, D. R. (1990). *Learning and Study Strategies Inventory - High School version: User's manual*. Clearwater, FL: H & H Publishing.
- Wong, B. Y. L. (1980). Activating the inactive learner: Use of questions/prompts to enhance comprehension and retention of implied information in learning disabled children. *Learning Disability Quarterly*, 3, 29-37.
- Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614-628.
- Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80, 284-290.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59.

Author Note

This article is derived from a Doctoral dissertation completed by the first author. The author is grateful to the participating staff and students. Correspondence concerning this article should be addressed to Joyce A. Pereira-Laird, Department of Psychology, Massey University, Private Bag 11222, Palmerston North, New Zealand.

APPENDIX A

Table A1. Estimated correlation matrix for the cognitive and metacognitive strategy items

	ITEMS																					
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12
c1	.75																					
c2		.80																				
c3			.92																			
c4				.72																		
c5					.64																	
c6						.81																
c7							.78															
c8								.80														
c9									.89													
c10		-.07	.30							.08	.20	.86										
m1											.67											
m2												.67										
m3										.14			.67									
m4														.71								
m5															.60							
m6	.18	.23														.81						
m7																	.60					
m8						.11												.60				
m9									.20	.15									.81			
m10	.15					.17														.65		
m11				-.14																.18	.59	
m12	.21																				-.14	.70

note: see Table 1 for the items.

Table A2. Estimated correlation matrix for the cognitive/metacognitive strategy and metacognitive knowledge indicators

	COMPOSITE ITEMS							
	cog.1	cog.2	cog.3	meta.1	meta.2	meta.3	MSI	IRA
cog.1								
cog.2	.566							
cog.3	.496	.510						
meta.1	.441	.476	.425					
meta.2	.553	.425	.449	.634				
meta.3	.393	.333	.337	.577	.564			
MSI	.346	.386	.411	.492	.477	.383		
IRA	.279	.191	.299	.341	.323	.347	.440	

Note: cog.1 to cog.3 and meta.1 to meta.3 represent cognitive and metacognitive composite items, respectively.

APPENDIX E

Pereira-Laird, J. A., & Deane, F. P (1996). Development of a behavioural measure of strategy use for adolescents: Think-Aloud Reading Strategy Use Measure (TARSUM). Manuscript submitted for publication.

ABSTRACT

The lack of systematic research into the strategic processes for effective reading and learning may have been hindered by a lack of psychometrically sound instruments to measure cognitive and metacognitive strategy use. In an attempt to remedy this deficiency a reading strategy use measure, called the Think Aloud Reading Strategy Use Measure (TARSUM) was developed. A pilot study of an initial application of this measure was used for measure revision. A second study was conducted to trial the procedures so that the TARSUM could be administered to children of varying reading competence. The findings of these two studies suggested a number of revisions, and provided the final development of the measure. These revisions were implemented in the third study. Results from a third study provide evidence of the revised measures' validity and reliability. The final measure showed high reliability. Convergent validity was supported by its moderate correlation with a paper-and-pencil measure of reading strategy use. Confirmation of several predicted conceptual relationships between the TARSUM and other reading related factors (i.e., achievement, anxiety, perceived value, self-efficacy, and metacognitive knowledge) provided evidence of construct validity. The TARSUM appears useful for classroom and research settings. Recommendations for further development of the measure are indicated. Limitations and suggestions for further research are discussed.

Researchers have been particularly interested in understanding processes involved in higher order academic skills such as reading comprehension (Garner, 1987; Paris, Wasik, & Van der Westhurzen, 1988). Information processing theory provides the theoretical framework for most recent research in reading, because it addresses the complex interplay of numerous cognitive and metacognitive processes and strategies individuals employ as they read. Successful readers have been shown by research to utilize a variety of cognitive and metacognitive strategies to improve their reading and learning (Garner & Alexander, 1989; Pressley, Borkowski, & Schneider, 1987).

However, systematic theory development in this area has been impeded by the lack of reliable and valid instruments to measure reading strategy use. A major problem with many existing reading strategy use measures is the lack of systematic empirical norming in addition to being too global in nature. This article describes the development of a behavioural measure of metacognitive and cognitive strategy use in the domain of reading. The primary rationale for developing such a measure was to provide a research tool to assist in studying the relationship between reading cognitive/metacognitive strategy use and reading achievement in children.

Although there is consensus that strategy use is important for understanding text (e.g., Alexander & Judy, 1988; Brown, Bransford, Ferrara, & Campione, 1983; Garner, 1990), there is little agreement on the definition of strategy (Paris, Wasik, & Turner, 1991). Related discussion in the literature concerning differentiation of cognitive and metacognitive activity does little to clarify the central notion of a strategy. Rather, the concept appears to drift in response to the conceptual needs of the writer, sometimes describing: (a) a self-verbalisation (Schunk, 1986), (b) a particular learning skill such as rehearsing and outlining (e.g., Levin, 1986), (c) more general types of self management activities such as planning and comprehension monitoring (Pressley et al., 1987), (d) complex plans that combine several

techniques (e.g., Derry & Murphy, 1986), or (e) mental activity presumed to relate to information handling and transformation (Swanson, 1989). Gerber (1983) warned that the concept of strategy would become progressively obscured and lose its apparent explanatory power unless its meaning was clarified.

In recent articles about strategies some researchers do not include a specific definition for "strategy" (Borkowski, Carr, & Pressley, 1987; Paris & Oka, 1986); while others use different definitions between studies (e.g., Brown et al., 1983; Paris et al., 1991; Wade, Trathen, & Schraw, 1990; Weinstein & Mayer, 1986). One common element in most of the definitions of "strategy" that were surveyed, is the idea that a "strategy" is a student-directed action the learner chooses consciously and deliberately as a means to an end. There are, however, others in the strategy research literature who argue for the automaticity of strategy use (Pressley, Forrest-Pressley, & Elliot-Faust, 1988). For the purposes of this article, Rowe and Rayford's (1984) definition of a strategy is adopted where it is defined as a deliberate action that readers take voluntarily to develop an understanding of what they read.

It appears that at times the distinction between cognitive and metacognitive strategies is unclear with researchers often being quick to label any strategic action as metacognitive (Brown et al., 1983). Perhaps a resolution of the conflict is that proposed by Flavell (1979) who states "Cognitive strategies are invoked to *make* cognitive progress, metacognitive strategies to *monitor* it" (p. 909). In other words, one way of viewing the relationship between them is that cognition is involved in doing, whereas metacognition is involved in choosing and planning what to do and monitoring what is being done. However, Brown (1987) notes that in some cases a strategy may be invoked for either purpose or may achieve both a cognitive and a metacognitive goal. She describes an example as follows: "Asking yourself questions about the chapter might function either to improve your knowledge (a cognitive function) or to monitor it (a metacognitive function)" (p. 67). Thus, a specific activity can represent the strategy itself (looking for main points), its monitoring role (a metacognitive activity), and a reflection of metacognitive knowledge that the strategy is appropriate for the given situation.

Metacognitive and cognitive strategies have been distinguished empirically in a large number of studies (Anthony, 1994; Pokay & Blumenfeld, 1990; Pressley et al., 1987, 1991; Zimmerman & Martinez-Pons, 1990). Pintrich and De Groot (1990) who successfully used factor analysis to differentiate these strategies. Even though there was a substantial correlation (.83) between these strategies, the metacognitive strategy use measure predicted the students' academic performance much better. The authors concluded the two types of strategies were conceptually distinct based on their finding that cognitive strategy use without the concomitant use of self-regulation was not conducive for academic performance.

Little agreement has been reached about the most effective ways of measuring metacognition and strategy use (McLain, Gridley, McIntosh, 1991). Because of the difficulty associated with measuring metacognitive aspects of reading (Jacobs & Paris, 1987), there has been wide use of techniques which describe the strategies and processes used by good and poor readers (Garner, 1987). These include verbal reports (interview, think-aloud measures), questionnaires, computer and video studies, and reaction time. In this article, the two most commonly used measures will be discussed, self-report questionnaires and verbal report.

Self-report questionnaires have been used extensively to assess metacognitive and cognitive

strategies (Chan, 1995; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990). Although these measures can be used effectively to measure student perceptions of strategy use (Ames & Archer, 1988), researchers have suggested the need to replicate results with other measures, particularly in the light of certain inherent limitations of self-report measures (Zimmerman & Martinez-Pons, 1990). These limitations were outlined by Pokay and Blumenfeld (1990) as tapping students' perceptions of what they are doing rather than the accuracy of these perceptions, and not accounting for the efficiency or appropriateness of strategy use. Appropriateness of strategy use was not accounted for because students report strategies they use generally, rather than strategies they use to solve specific tasks. Thus, responses of students in questionnaires refer only to student perceptions and not necessarily to accuracy, effectiveness, or appropriateness of strategy use.

The verbal report method includes interviews and "think-aloud" (or concurrent) procedures. The interview involves asking children open-ended questions about how they think, read, learn, and remember, what they might think and do in a hypothetical situation or while completing a task at hand. Interviews produce retrospective verbalizations, for they elicit reports of metacognitive/cognitive activity already completed. On the other hand, think-aloud protocols generate *concurrent* verbalizations about a reading activity that is temporarily interrupted at frequent predetermined and cued points for provision of the verbal report. Research seems to support the use of concurrent reporting (think-aloud protocols) in preference to retrospective reporting as being a more reliable measure of students' strategy usage (Ericsson & Simon, 1980; Meichenbaum, Burland, Gruson, & Cameron, 1985; Meyers & Lytle, 1986).

A number of criticisms have been levied at verbal-reports (Garner, 1987). One major concern is that children may lack the language and verbal facility to discuss the mental events. A second criticism is that verbal-report data may include subjects' rationalizations, fabrications, and elicited mimicry because of the demand characteristics of the situation. Moreover, verbal responses may be ambiguous and may be edited by the subjects to conform to their expectations about what the investigator wants to hear. There is also some evidence that asking questions during reading alters the way in which subjects process the text, and in addition, interacts with other variables to influence recall (Wade & Trathen, 1989). A third concern about verbal-report data is that their reliability and validity are rarely assessed. For instance, the cueing offered by instructions and probes may invalidate the data. Another difficulty with verbal reports is that they may not reflect what students are really doing; that is, they may be engaged in metacognitive strategy use but not report it, or they may report it inaccurately (Reynolds & Wade, 1986; Schneider, 1985). More extensive reviews of the problems surrounding verbal report data are available elsewhere (see Afflerbach & Johnston, 1984; Garner, 1987; Nisbett & Wilson, 1977).

Despite the aforementioned criticisms, Ericsson and Simon (1980) have asserted that verbal report data can provide a great deal of knowledge about students' metacognitive activity. They suggest that the following factors can increase the accuracy of verbal report data: (a) not requiring subjects to verbalize at the same time as they are performing the task, (b) minimizing the time interval between processing and reporting, (c) emphasizing to subjects that they should focus primarily on the task, rather than on the verbal-reporting of their processes, and (d) collecting data from other sources to check for convergence of data. They also argue that people are aware of thinking process, and that verbal reports offer information that otherwise would be unavailable to the researcher.

For the present study the behavioural or interview research instruments were reviewed and indicated that many concentrated on learning strategies or study skills. Hence there appeared a need for instruments that measured reading strategies specifically. While there is greater support for the 'think-aloud' concurrent procedure, a *structured* approach of the think-aloud procedure was thought to yield more information about strategy usage considering the verbal reticence of reading disabled students. One such procedure had been devised by Paris (1991) to measure reading strategies and was called the think-along-passage (TAP) procedure.

Although Paris (1991) illustrated how the TAP could be used for research purposes, he developed it mainly for teachers to enable them to find out how well students read or how they thought as they read. He suggested that any passage from regular curriculum materials can be modified to be a TAP, so long as questions similar to those used in his measure are asked. Essentially the TAP involved students reading aloud to teachers who asked interspersed questions that probe comprehension, cognitive strategy use, and explicit knowledge of strategies

One limitation of Paris' (1991) measure was that it did not measure metacognitive strategy use. Although the TAP procedure was devised mainly for assessing cognitive strategy use it can be modified to yield measures of metacognitive strategy use. Research has demonstrated that both metacognitive and cognitive strategies are necessary for reading success. For instance, Pintrich and De Groot's (1990) research suggests that cognitive strategy use without the concomitant use of metacognitive strategies is not conducive for academic performance. The cognitive strategies are essential tools to learn text, but the metacognitive strategies provide the monitoring and direction for their use. Consequently, it seems important to develop effective measures of strategy use that assess cognitive and metacognitive strategies independently.

This article describes the results and the conclusions of three investigations that were conducted to trial a modified version of Paris' (1991) TAP measure called the Think Aloud Reading Strategy Use Measure (TARSUM). Three studies using different samples provide data on the development and psychometric properties of the TARSUM. The first study is a pilot study and provides development of the measure. The second study investigates variations in passage difficulty so that the measure can be applied to children with a variety of reading skills. The third study provides further construct validation data for TARSUM.

STUDY 1: PILOT STUDY - PROCEDURAL DEVELOPMENT

This pilot study provides an initial effort in developing the TARSUM by modifying Paris' (1991) TAP procedure to include measures of metacognitive strategy use. Part of the motivation for developing the TARSUM was the need for a valid and reliable assessment method of cognitive and metacognitive strategies for use in both research and instruction. In particular, there was a need for an instrument that could be used to diagnose the strengths and weaknesses of achieving and underachieving readers. Items and procedures were developed for the TARSUM which would measure metacognitive and cognitive strategy use.

Essentially, the TAP involves asking students questions at predetermined points in the passage. The passage that was used to illustrate Paris' (1991) TAP measure had five

comprehension and five metacognition questions. While Paris' "metacognition" questions were designed to measure cognitive strategy use and explicit knowledge of strategies, they did not measure metacognitive strategy use as the name would suggest. For instance, the question "If you don't know, how could you tell?" was designed to reveal strategies that students could use, if they reported not knowing the answer to a comprehension question. This question was omitted from the TARSUM as it reflected a measure of metacognitive knowledge rather than strategy use. In its place the question "How did you go about looking for the answer?" was asked to reveal the cognitive strategies students used in their attempt to locate an answer.

The TARSUM included procedures as well as probes (questions indicated by asterisks at predetermined points in the passage), which not only assessed comprehension and cognitive strategy use, as in Paris' TAP measure, but also metacognitive strategy use. A number of researchers measure metacognitive processes (regulation) by examining students' oral reading behaviours such as self-correcting and altering reading rate. These behaviours were measured in the TARSUM to provide an index of metacognitive strategy use.

The scoring method for strategy use was modified for the TARSUM. For each cognition question, students were awarded 1 point for each appropriate cognitive strategy they used in arriving at the answer instead of just 1 point irrespective of the number of strategies used. In addition, subjects were not awarded any points for strategies that were observed by the interviewer but not verbalised by the subject (e.g., scanning text, looking at title, rereading silently). This avoided subjective inferences being made on the part of the interviewer.

Two passages that were appropriate for Form 2 (9th grade) New Zealand children were chosen to trial the TARSUM. In this pilot study, eight comprehension and eight cognition questions similar to those found in the TAP were generated by the researcher for each passage. Unlike the TAP, questions which assessed cognitive strategy use in the TARSUM are referred to as "cognition" questions.

In summary, as an initial effort in the development of a behavioural measure, the present study pilot tested a modified version of Paris' (1991) measure, called TARSUM. The changes involved the incorporation of measures of metacognitive strategy use, and revising the scoring procedures. It was hoped that the results would provide some directions in which the measure could be improved further. This investigation was also used to establish the amount of time taken to administer the measures, to trial instructions given to the students, to determine the appropriateness and clarity of the language of the measures, and to determine the suitability of the passages. In addition, construct validity was explored through correlations with reading achievement measures and self-report reading measures.

METHOD

Sample

Parent and child permission were obtained for children to participate in this investigation. Twenty two Form 2 children were randomly selected from an intermediate (junior high) school. Subjects constituted 13 female and 9 male students with a mean age of 12.10 years ($SD = 0.81$). The students were predominantly of European descent and from middle-class families. Teachers were asked to classify the 22 students as either good ($n = 13$) or poor readers ($n = 9$) based on their reading performance in the classroom.

Instruments

Passages

An easy and a moderately difficult passage were used. The passage, entitled "The Moa Hunters" (the easy passage) was obtained from the Analytical Reading Inventory (Woods & Moe, 1981). The second passage, entitled "The Volcano", was obtained from the Neale Analysis of Reading Ability Manual (Neale, 1989). Both passages were considered by teachers to be appropriate for Form 2 (9th grade) children. These passages have been used by other New Zealand researchers (van Kraayenoord, 1986; L. Wilkinson, personal communication, April, 1993). In addition, the passages were not used by participating schools and hence the children had not been exposed to the passages before.

TARSUM - Interview Coding Form

The interview coding form was used by trained interviewers to record behavioural data on cognitive and metacognitive strategies. Similar to Paris' (1991) TAP procedure, the TARSUM involved students reading aloud to interviewers who asked predetermined questions interspersed along the passage. The eight comprehension and eight cognition questions were asked at designated points in the passage (indicated by an asterisk).

The questions were contained on the left hand side of the "interview coding form" with several strategies listed opposite each cognition question so that the spontaneous cognitive strategies used by students could be recorded. In addition, opposite each cognition question were frequency boxes in which interviewers could record the number of times students altered their reading rate, made errors, and self-corrected their reading. (Data on these oral reading behaviours was collected for each segment read between asterisks).

Similar to the TAP, the comprehension questions assessed children's ability to identify the topic of the passage, to make a prediction within the passage, to monitor the meaning of a unique word, to make an inference, and to provide a summary of key ideas found in the passage. Students were awarded points for each comprehension question that was correctly answered. The maximum score for each comprehension question is provided in the interview coding form beside the question. The comprehension score for each question were summed to provide an additional measure of students' comprehension performance. Note that students' comprehension performance was also assessed using the PAT, which is a standardized achievement test.

The cognition questions that followed children's answers to each comprehension question were designed to reveal the **cognitive** strategies that students used. Examples of cognitive strategies that were examined included: rereading, predicting based on context clues, and using prior knowledge. The cognition questions were usually in the form "How did you know this?", "How could you tell?", or "How did you get your answer?" Students were awarded one point for each appropriate cognitive strategy that was verbalized.

Progressive Achievement Test

Students' standardized reading achievement was assessed using one of the reading tests from the Progressive Achievement Test Series (PAT), namely the PAT Reading Comprehension

(Reid & Elley, 1991). This test is a group administered, New Zealand (NZ) normed, paper and pencil scale, administered by the majority of NZ primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Its split-half reliability coefficient is above .85 and the test is described as having medium to high validity in NZ (Reid, 1993; Reid & Elley, 1991).

Reading Strategy Use Questionnaire

The Reading Strategy Use questionnaire (RSU, Pereira-Laird, 1996) is a 22 item scale containing 10 cognitive strategy items and 12 metacognitive strategy items. The metacognitive strategies subscale (RSU-meta) contained items on planning, monitoring, and evaluation of the student's own cognition; for example "I read quickly through the whole passage to get the general idea before I read it thoroughly". The cognitive strategies subscale (RSU-cog) refer to those strategies students use to integrate new material with prior knowledge, that is, strategies that students use to learn, remember, and understand the material (Zimmerman & Martinez-Pons, 1988). They include rehearsal, elaboration, and organizational strategies; for example, "I learn new words by picturing in my mind a situation in which they occur". Respondents indicate the frequency they use each strategy on a Likert-type scale ranging from "never" (1) to "always" (7).

The RSU has high internal reliability with total scale Cronbach alpha of $r = .87$. Internal reliability was moderately high for each of the subscales; $r = .73$ for the cognitive subscale and $r = .85$ for metacognitive subscale. The RSU correlates significantly with measures of reading metacognitive knowledge ($r = .46$ to $.53$), reading comprehension ($r = .53$ to $.70$), and perceived value for reading ($r = .45$ to $.53$).

A mean score for cognitive strategy use was calculated by summing the scores of the 10 cognitive strategy items on cognitive strategy use subscale (RSU-cog) and dividing it by 10. Similarly a mean score for metacognitive strategy use was obtained by summing the scores of the 12 metacognitive items (RSU-meta) and dividing it by 12.

Procedure

Individual interviews were conducted with all students and each session was tape recorded. All interviews were conducted by the primary researcher (JP) and a research assistant. The research assistant received training in the interview procedure from the primary researcher. This involved training in the notation used to record students' oral reading behaviours, training in identifying the various strategy types, reviewing interview sessions conducted by the primary researcher and conducting practice trials under the supervision of the primary researcher. Any differences in coding were discussed and clarified.

The data collection method was similar to that outlined by Paris (1991). Students completed the easy passage first before going on to the difficult passage. They were instructed to stop at predetermined points (indicated by an asterisk on the passage) where they were asked questions designed to reflect their comprehension and use of cognitive strategies.

The interviewer had a similar passage to refer to while the student was reading. While the student read each segment of the passage or an equivalent passage, the interviewer noted on

his/her copy of the passage the points where errors, successful and unsuccessful self-corrections, and alterations in reading rate occurred. Each time the student stopped at the designated points, the interviewer also asked the student the prescribed questions for the segment.

As the student answered the questions, the interviewer recorded the answer to the comprehension question and the presence of each strategy on the interview coding form. If a student used a strategy not listed on the coding form this was checked as "other" and the strategy was specified. After answering the questions at each designated point, students continued reading the next segment of text. If they failed to stop at the designated points, they were reminded to do so.

Metacognitive strategy use was assessed by measuring the following variables: (a) students' accuracy in self-correcting self-generated errors (e.g. correcting mispronounced words or errors they made while reading) and (b) frequency in alterations of reading rate. Data on frequencies of errors, self-corrections, and alteration in reading rate was collected for each segment read between asterisks. The total frequency for each of the oral reading behaviours was calculated at the end of the interview session.

RESULTS

The data from the interviews generated 3 types of dependent measures: a behavioural measure of cognitive strategy use (TARSUM-cog), a behavioural measure of metacognitive strategy use (TARSUM-meta), and a comprehension measure (TARSUM-comp). TARSUM-cog was calculated by summing the scores for all the cognition questions. TARSUM-meta was calculated by expressing the ratio of successful attempts at self-correction to total errors (uncorrected errors plus unsuccessful self-corrections) as a percentage. (A frequency count of alterations in reading rate provided a second measure of TARSUM-meta). TARSUM-comp was calculated by summing the scores of all the comprehension questions.

Data was collected for the easy and difficult passages. Observations indicated a number of the poor readers found the "Volcano" passage difficult to read and were unable to read it with more than 40% accuracy. Therefore, these children (3 males, 2 females) were excluded from further study. It was also noted that more strategic responses were generated for the moderately difficult passage, the "Volcano". Furthermore, it took about an hour to conduct interviews using the two passages. Given that excessive interview length might affect both student motivation and the utility of the behavioural measure for research purposes, only the moderately difficult passage (the Volcano) was used in future measure development. In addition, this passage elicited more strategic responses and had greater potential for use by students from other countries given the generic content. Consequently, only the results for the "Volcano" passage are elaborated.

Reliability for the TARSUM subscales were calculated for all 17 students by using point-by-point percentage agreement (Cooper, Heron, & Heward, 1987). Interrater agreement was calculated on a question-by-question basis by dividing the number of agreements by the number of agreements plus disagreements, multiplied by 100.

Interrater agreement for cognitive strategy use (TARSUM-cog) and metacognitive strategy use (TARSUM-meta) was 94.02% and 81.15% respectively. Overall interrater agreement for the

two subscales (i.e., both types of strategy use) was 94.02%. Interrater agreement for metacognitive strategy only utilized the self-correction measure because interrater reliability for determining alterations in reading rate was low (48%). Consequently, alteration in reading rate was dropped as a measure of metacognitive strategy use.

Despite the small sample size, it was anticipated that if like constructs had high positive correlations than dissimilar constructs (i.e., cognitive-metacognitive) this would provide preliminary data suggesting construct validity. Table 1 provides correlations between the behavioural (TARSUM-cog, TARSUM-meta) and self-report strategy use (RSU-cog, RSU-meta) measures of strategy use. In general, the anticipated patterns were confirmed with TARSUM-meta correlating more strongly with RSU-meta than with TARSUM-cog. Similarly, TARSUM-cog correlated more strongly with RSU-cog than with TARSUM-meta. The strongest correlation was found between self-report measures which may suggest a degree of method variance.

Table 1. Correlations among the behavioural and self-report measures of cognitive and metacognitive strategy use ($n = 17$)

	RSU-cog	TARSUM-cog	RSU-meta	TARSUM-meta	TARSUM-comp	PAT
RSU-cog		.33	.67*	.17	.72*	.63*
TARSUM-cog			.13	.24	.50*	.58*
RSU-meta				.35	.50*	.67*
TARSUM-meta					.39	.34

Note: * indicates that the correlations are significant at $p \leq .05$ (one-tailed).

Correlations between each of the comprehension measures (TARSUM-comp and PAT) and the behavioural and self-report measures of strategy use were also calculated. These correlations were statistically significant, ranging between .50 and .71, with the exception of TARSUM-meta. Of the behavioural and self-report measures, the self-report measures had the highest correlation with the comprehension measures. High correlations were also found between the comprehension measures and the behavioural measures. As expected, the correlation between both comprehension measures was high ($r = .78$).

DISCUSSION

The results provided preliminary support for the reliability and validity of TARSUM, with the exception of the use of alterations in rate of reading as an aspect of metacognitive strategy use. In addition to dropping this part of the measure, the results indicated a number of directions for continued development. A more comprehensive selection of metacognitive strategy use questions appeared necessary. Second, a monitoring device which could assess comprehension unobtrusively would acknowledge student's own awareness of monitoring, rather than an inference of monitoring made by the researcher. The latter was considered important because pilot study data suggested that interviewers found it difficult to make accurate estimations of students' monitoring or regulation processes. Third, some very minor changes to the comprehension questions and general instructions were indicated because some

students found them ambiguous.

Finally, "The Volcano" passage was too difficult for a number of poor readers, thus necessitating the creation of multiple versions of this passage to control for task difficulty. It was considered important that the passage should be of sufficient difficulty and length so as to generate more strategic responses. Research has indicated that oral reading behaviour can be altered significantly by text difficulty (Blaxall & Willows, 1984). Students who are faced with an extremely difficult passage will not be able to use context and cues to monitor meaning and make predictions. When the text is too difficult, poor readers are likely to treat the words as isolated units and fail to bring their strategies to bear on the text, making examination of their use of strategies difficult. This was observed in the pilot study where the number of strategic responses were severely hindered because some readers found the passage extremely difficult to read. Hence, generating multiple versions of the passage would enable the TARSUM measure to be applied to children of varying reading competence.

Before implementing these modifications, it was considered necessary to generate multiple versions of a passage and trial them to determine their suitability across a wide range of readers.

STUDY 2 - VARIATIONS IN PASSAGE DIFFICULTY

The results of the first investigation indicated that multiple versions of a passage were required to control for task difficulty and to enable TARSUM to be applied to children of varying reading competence. Three criteria were thought to be important in the selection of the passage. As noted, the passage needed to be of reasonable difficulty and length so as to provide students with the opportunity to generate more strategic responses. A related issue was the confounding effect of task difficulty on the generation of strategic responses. Several researchers have noted that oral reading behaviour alters significantly when children read at different levels of text difficulty (Blaxall & Willows, 1984; Ng, 1979; Williamson & Young, 1974). In particular, research has indicated that flexibility of strategy usage decreases for reading disabled or poor readers as reading material increases in difficulty (Blaxall & Willows, 1984). Thus, it was considered important to equate difficulty level by having children read at their own "moderately difficult" level. Second, the passage selected was required to be typical of the type of comprehension exercises frequently used in most reading classrooms at the second form level. Third, the passage was required to be on an unfamiliar topic. This was an important consideration as prior knowledge could negate the need for students to employ a range of strategies when attempting to comprehend the passage and answer the accompanying questions.

The passage, entitled "The Volcano", used in the pilot study was considered likely to meet the aforementioned considerations. This investigation developed three versions of "The Volcano" passage. As it was considered impractical and time consuming to determine the reading age and the appropriate version to be used for large numbers of students, it was decided to administer the versions according to preset criteria, using achievement test scores.

METHOD

Sample

The sample consisted of 33 randomly selected Form 2 (9th grade) students from two junior high schools in the Manawatu district. Fourteen were males and nineteen were females. The

majority of students were of European descent (87%) whilst 13% constituted non-Europeans (Maoris, Asians and others). This was generally consistent with the population parameters in New Zealand.

Instruments

Standardized measure of reading achievement

As in the pilot study, the PAT reading comprehension scores of each child was obtained from school records.

Passages

Three versions of the passage entitled "The Volcano" were generated and rewritten at three different readability levels, in order to exercise some control over task difficulty. Prior to generating the versions, six textual irregularities were inserted to measure metacognitive strategy use and a third paragraph was appended so as to lengthen the passage. The textual irregularities consisted of 2 textual inconsistencies, 2 nonsense words, and 2 scrambled words. In creating the versions, some of the words and sentences from the original passage were rephrased and modified to reduce difficulty of the passage. Essentially the story line was the same for all the versions. Every attempt was made to include all information in the three versions and to maintain a comparable length. As a result, length ranged from 211 to 243 words with small differences in the number of idea units expressed. The 3 versions of the passage differed only in difficulty of the vocabulary words with version 1 being the hardest and version 3 being the easiest (see Appendix A).

The versions were created so that students reading their version found it moderately difficult rather than too difficult or too easy. Previous researchers have used 95% or above in accuracy and 75% or above in comprehension for the easy passage (Ng, 1979; van Kraayenoord, 1986), and 90% or below in accuracy and 50% or below in comprehension for the difficult passage (Johns, 1981; van Kraayenoord, 1986). For this investigation, a set of criteria had to be determined for a moderately difficult passage. It seemed reasonable to set the criteria somewhere in between those set for the easy and difficult passage. Hence for this investigation the passage criteria was set at above 90% in accuracy and above 50% in comprehension to be considered moderately difficult. The accuracy score consisted of the proportion of correctly read words in a passage expressed as a percentage of the total number of words. Errors which had been successfully corrected were not included in the error count.

Procedure

One-to-one interviews were conducted by the researcher with each child and the sessions were tape recorded. For each student, the appropriate version of the passage was determined using the following criteria. The first version was used for the group of students who scored at or above the 70th percentile for PAT Reading comprehension. The second version was used for the group of students who scored at or above the 45th percentile but below the 70th percentile for PAT Reading comprehension. The third version was used for the group of students who scored below the 45th percentile for PAT Reading Comprehension.

After students had read the passage, they were asked eleven questions on the passage to assess

their understanding. The comprehension questions assessed similar aspects to those in the pilot study. In addition, students were asked to rate the difficulty of the passage, using the "Passage Difficulty Ratings". Four questions were asked of students to examine the perceived difficulty of the passage. Two questions were worded in the affirmative and two in the negative to control for potential response bias. The questions were: (a) How hard did you find this passage to read? (b) How comfortable did you feel in reading this passage? (c) How hard is it to understand this passage? and (d) How easy were the words in this passage to pronounce? Subjects responded on a 7-point Likert scale ranging from "not at all hard/uncomfortable", 1, to "very hard/uncomfortable", 7. The scales were presented in the reverse direction for negatively worded items.

RESULTS AND DISCUSSION

Following the interviews, the tapes were transcribed. The number of successful/unsuccessful self-corrections and uncorrected errors were calculated. Using this information, students' accuracy in reading was calculated by expressing the proportion of correctly read words as percentage of the total number of words in the passage. Errors which had been successfully corrected were not included in the error count. In addition, a score for comprehension accuracy was calculated by summing the scores of all the comprehension questions (maximum score = 15).

All items were scored such that higher scores represented greater difficulty or discomfort. Table 2 provides means and standard deviations for each difficulty rating item, overall ratings of difficulty, comprehension accuracy and accuracy of reading for each of the reading achievement groups.

Table 2. Difficulty ratings, reading and comprehension accuracy obtained for each passage version

Difficult/comfort items, reading accuracy, comprehension	Passage versions		
	1 (<i>n</i> = 10) <i>M</i> (<i>SD</i>)	2 (<i>n</i> = 10) <i>M</i> (<i>SD</i>)	3 (<i>n</i> = 13) <i>M</i> (<i>SD</i>)
How hard to read?	3.4 (1.1)	4.1 (0.9)	4.0 (1.3)
How uncomfortable to read?	3.5 (0.8)	3.9 (1.1)	4.1 (1.4)
How hard to understand?	3.2 (1.0)	3.5 (1.5)	3.5 (1.0)
How difficult to pronounce?	3.8 (0.8)	4.3 (1.1)	4.2 (1.3)
Overall rating of difficulty	3.47 (0.71)	3.95 (1.0)	4.00 (0.98)
Comprehension accuracy	10.59 (2.0)	9.89 (1.5)	9.75 (1.6)
Accuracy of reading	96.3% (1.4)	95.0% (2.1)	96.3% (1.8)

Note: A rating of 1 indicates "not at all hard/uncomfortable" and a rating of 7 indicates "very hard/uncomfortable".

The results indicated that most students who fell within each of the three reading achievement groups met both the accuracy and comprehension criteria for their version of the passage. In general, students in each of the groups rated their version of the passage as being moderately difficult to read, understand, and pronounce as indicated by mean scores ranging from 3.2 to 4.3 on the 7-point Likert-type items. A series of one-way ANOVA's indicated no significant differences between the three groups on all items suggesting a degree of passage equivalence for students at different reading comprehension achievement levels.

STUDY 3 - VALIDATION OF THE TARSUM

The pilot study suggested: (a) questions tapping cognitive strategy use had acceptable reliability and validity; (b) the metacognitive measure "altering reading rate" had low reliability and should be discarded; (c) a more comprehensive selection of metacognitive strategy use questions was required; (d) a monitoring device might help to monitor comprehension unobtrusively; and (e) the passage was too difficult for a number of reading disabled and poor readers, thus necessitating the creation of multiple versions of the passage to control for task difficulty. Study 2 suggested that the three versions that were created were of moderate difficulty and demonstrated that the preset criteria that were used in administering the versions to students of varying reading ability was acceptable. Consequently, Study 3 aimed to incorporate these improvements.

In formulating further measures of metacognitive strategy use, a review of the literature on error detection was undertaken. The error detection paradigm has been used by several researchers in investigating metacognitive strategy use. It is presumed that good cognitive monitoring reflects the ability to identify and correct errors and is thus a measure of metacognitive regulation. One useful approach to gathering data on metacognitive self-regulation processes involves altering texts (e.g., inconsistent text, nonsense words) to include impediments to comprehension and then seeing if students notice them (Baker, 1985; Baker & Brown, 1984; Zabrocky & Ratner, 1986). The assumption is that if students are engaging in monitoring activities, they will notice the inconsistency or irregularity and take some corrective action. Studies of awareness of deliberately inserted ambiguities, inconsistencies and errors have shown that both good and poor readers do monitor their comprehension; that is, detect inconsistencies (e.g., Bos & Filip, 1984; Winograd & Johnston, 1982). However, poor readers and reading disabled students do not monitor their reading spontaneously (e.g., Bos & Filip, 1984; Zabrocky & Commander, 1993). Thus, the present study incorporated the error detection paradigm as one method of collecting data on metacognitive strategy use. Questions designed to elicit students' monitoring of an inconsistency or irregularity in the text were termed as "metacognition" questions.

Study 3 was conducted to determine the psychometric properties of TARSUM. Construct and discriminant validity were evaluated through correlations with several achievement measures, reading strategy use, perceived value, and anxiety, efficacy, and metacognitive knowledge.

METHOD

Subjects

The sample consisted of 139 randomly selected Form 2 (9th grade) students from four

intermediate (junior high) schools in the Manawatu district. Subjects (45.3% males and 54.7% females) were aged approximately 12 years ($M = 12.48$ years, $SD = 0.32$). The majority of students were of European descent (87%) whilst 13% constituted non-Europeans (Maori, Asian and others), again consistent with NZ population parameters.

Measures

Achievement

Academic achievement

Students' academic achievement was assessed using four of the seven tests of the Progressive Achievement Test Series (PAT). The PATs that were used in the present study included Reading Comprehension (Reid & Elley, 1991), Reading Vocabulary (Reid & Elley, 1991), Reading Study Skills (Reid, Croft, & Jackson, 1978), and Mathematics (Reid, 1993). These tests are group administered, NZ normed, paper and pencil scales, administered by the majority of NZ primary and intermediate schools at the beginning of each school year (Beck & St. George, 1983). Their split-half reliability coefficients are above .85 and the tests are described as having medium to high validity in NZ (Reid & Elley, 1991; Reid, 1993).

Classroom reading performance

Teachers provided gradings, expressed as a percentage, for students' performance on different reading tasks in the classroom. These tasks fell into three general categories: in-class seatwork, quizzes and tests, and assignments and projects. An average score for the three categories was calculated with high scores representing good reading performance.

Reading Strategy Use Questionnaire

Described in the pilot study.

Perceived value

Eleven items, similar to those used by Pintrich and De Groot (1990) in their intrinsic value scale, were used to assess students' perceived value of reading. Whereas Pintrich and De Groot asked reactions to general academic work, students in this study were questioned specifically about reading. Items concerning the value of reading assessed perceived usefulness, importance, and inherent interest. In the present study, the scale had high internal consistency with a Cronbach's alpha of .89.

Anxiety

The reading anxiety scale (RAS) was adapted from Pintrich and De Groot's (1990) test anxiety scale to create an instrument which was reading-specific and appropriate to intermediate aged children. The RAS consists of 11 items of which four were negatively worded to control for response bias. Scoring of the negative items was reversed so that a high score indicated low anxiety. In the present study, Cronbach's alpha was .93.

Task-specific reading self-efficacy scale

The six-item Task-specific Reading Self-efficacy Scale (TRSE; Pereira-Laird & Deane, 1995), measured self-efficacy for reading. Children's self-efficacy for reading different materials (e.g., an encyclopedia, a 150 page novel, a letter from a friend, etc.) was measured on a scale that ranged from 0 to 100 in 10-unit intervals from high uncertainty to complete certainty. The scale has high internal consistency with Cronbach alpha of .91. The scale shows high validity as is demonstrated by its moderately high correlations with measures of reading comprehension ($r = .75$), perceived value ($r = .63$), and anxiety ($r = -.74$). The psychometric properties of this scale are provided in Pereira-Laird and Deane (1995). A task-specific self-efficacy score was obtained by calculating the mean.

Metacognitive knowledge

The Metacomprehension Strategy Index (MSI), developed by Schmitt (1988, cited in Schmitt, 1990) is a 25-item, four-option multiple-choice questionnaire with one option representing metacomprehension awareness (1 point) and the other three options representing inappropriate responses (0 points). Scores for the 25 items on the MSI were combined to produce a total score ranging from 0 to 25 points. It was developed by Schmitt to evaluate young adolescent students' awareness of metacomprehension strategies before, during, and after reading a narrative prose. The MSI assesses students' metacognitive awareness within six broad categories: purpose setting; predicting and verifying; previewing; using background knowledge; summarizing and applying fix-up strategies; and self-questioning. The MSI has been widely used by a number of researchers (Baumann, Seifert-Kessell, & Jones, 1987; Horrex, 1992; Lonberger, 1988).

The MSI has been reported to have an internal consistency value of .87 using the Kuder-Richardson Formula 20 (Lonberger, 1988). Statistically significant correlations were found between the MSI and the IRA ($r = .48, p \leq .001$). Furthermore, statistically significant correlations were found between the MSI and two comprehension measures commonly used to measure students' metacomprehension ability, an error detection task ($r = .50, p \leq .001$) and a cloze task ($r = .49, p \leq .001$).

Minor modifications were made to the wording of the MSI to improve the face validity of this questionnaire for Form 2 New Zealand children. The alterations acknowledged not only narrative prose reading, but also textbook/book reading. For example in item 2 (A), instead of 'Look at the pictures to see what the story is about', the item was changed to 'Look at the pictures or diagrams to see what the chapter is about'. Items 2 and 20 were deleted for the present study to improve the reliability of the scale. Internal consistency of the measure for the current study yielded a Cronbach's alpha of .72 for the whole sample.

TARSUM

Three instruments, namely a reading passage, a monitoring device called a "bleep", and an interview coding form, were required to conduct the taped concurrent interview which was designed to measure students' strategy usage. The three passages developed in Study 2 were used. Students were instructed to read aloud a passage (appropriate for their reading age) which was tape recorded by the interviewer. While reading, students were stopped at

designated points by the interviewer who asked questions that probed their comprehension, metacognitive strategy use, and cognitive strategy use. Interviewers recorded students' answers to the comprehension and cognitive/metacognitive strategy use questions on the interview coding form. Children used a device called a "bleep" to indicate awareness of monitoring their comprehension (by pressing a button).

Passage

Three versions of the passage "The Volcano" as used in Study 2.

Monitoring Device (bleep)

An unobtrusive measure of monitoring was sought to determine students' awareness of errors in the passage while reading with minimal disruption in the flow of reading. It was considered important that such a measure should only minimally interfere with the cognitive processing undertaken during reading. Furthermore, it should have validity in that it was an acknowledgement of the student's own awareness of monitoring rather than an inference of monitoring that the researcher made on the basis of puzzled looks or the like.

The bleep was considered a suitable instrument which might fulfill these requirements had been used previously by researchers to investigate on-line monitoring (e.g., van Kraayenoord, 1986). It consists of a red button placed on top of a small rectangular container. When the button is pressed an audible sound is heard. The students were asked to press the button or "bleep" when they came to any errors (textual inconsistencies, scrambled words, or nonsense words) in the text. This allowed for subsequent coding of monitoring behaviour with minimal disruption of students' flow of reading or cognitive processing.

Interview coding form

The interview coding form was utilised in the same manner as described in the pilot study but with the addition of several questions to better assess strategy use. Eleven comprehension, twelve cognition, and nine metacognition questions were used in this study (see Appendix B for sample questions). The comprehension questions, which usually formed 'part a' of the questions in the interview coding form, were identical to those used in Study 2. The cognition questions, which usually formed 'part b' of the questions in the interview coding form, were similar to those in the pilot study. Also found were questions that were designed to elicit metacognitive strategy use; these were termed "metacognition" questions and followed after segments of text which contained textual inconsistencies, scrambled words, or nonsense words. Thus, the interview coding form was equivalent to that used in the pilot study except that it now contained metacognition questions, with a list of strategies listed opposite it; frequency boxes for recording oral reading behaviour were omitted. Scoring procedures for comprehension questions are identical to those used in Study 2, whilst scoring procedures for cognition questions are identical to those used in the pilot study.

The metacognition questions assessed two types of activities involved in metacognitive strategy use: monitoring and planning. Evaluation strategies, which are also a form of monitoring, were also measured. Monitoring questions were asked at designated points where

the previous segment just read contained textual inconsistencies, nonsense or scrambled words. Planning and evaluation questions were asked after the passage was read.

To assess monitoring strategies, the passage was modified to include two sets each of textual inconsistencies, scrambled words, and nonsense words. Such a method has been used quite extensively by researchers to monitor comprehension (Bos & Filip, 1984; Zabrocky & Ratner, 1986). Students used a monitoring device (a bleep) to indicate any errors and disruption of understanding as they read the passage.

For the monitoring questions, students were asked "Why did you bleep?" or "What are you thinking here?" depending on whether they bleeped. (The second question was asked to take into account students who forgot to bleep but who may have been aware of the error in the text). Two of the six monitoring questions were followed by prompts. The monitoring questions assessing the first textual inconsistency and nonsense word were followed by prompts ("Do you agree with the sentence?" or "Is there such a word as *urgsy*?") for students who did not show awareness. Irrespective of whether students bleeped or not, two points were awarded for every verbalisation of awareness and correction by the student of scrambled words or textual inconsistency without prompting. Students were awarded only one point if they indicated awareness that something was wrong but was not able to correct the error or showed awareness and correction only after being prompted. One or two points were awarded for a verbalisation of awareness of the nonsense words depending on whether students were prompted.

The planning activity of metacognitive strategy use was assessed by an evaluation of the students' recall and summary of the story. A student who has planned the answer well is likely to deliver the main points and organize the ideas in recall. It is unlikely that the student is metacognitively aware if he/she knows that relating main points is a useful strategy but uses it indiscriminately by mentioning the irrelevant points and/or the details or presents the points in a disorganized manner. Hence, students were awarded 1 point if they had listed most of the main points in an organized manner. An answer was considered organized if the main points related by the student followed in a similar and logical order as that presented in the story.

To assess evaluation activities of metacognitive strategy use, two questions were generated. The first assessed students' subjective evaluation of difficulty of the passage they had read, and the second assessed students' evaluation of the accuracy of their answers to the comprehension questions. The first evaluation question asked students which of the three paragraphs they found hardest to read. An objective measure of student's difficulty of each paragraph was determined by calculating students' accuracy in reading each paragraph. One point was awarded if the paragraph mentioned by the student was read less accurately than the other paragraphs. If students had similar accuracy across paragraphs, mean comprehension scores for each paragraph were used as an alternative criterion. The second evaluation question asked students to indicate which of the 11 comprehension questions they had answered correctly. Students were awarded 1 point for each correct evaluation and 0 points for an incorrect evaluation. For the present study, this was done by the primary researcher at the end of the interview with interrater reliability subsequently assessed by an independent rater (see Results section).

In summary, the questions found on the interview coding form were designed to probe

comprehension as well as cognitive and metacognitive strategy use. The cognition questions which usually followed the comprehension questions, were designed to examine a number of cognitive strategies (e.g., predicting, using context clues). A number of indicators of metacognitive strategy use were also examined: (a) monitoring understanding when faced with textual inconsistencies, (b) monitoring errors and meaning when faced with scrambled words, (c) monitoring meaning when faced with nonsense words, (d) planning the summary of the story, (e) evaluating the difficulty of the paragraphs, and (f) evaluating how successful students were at answering the comprehension questions of the passage.

Two strategy use scores were computed for the interview measure. A cognitive strategy use score (TARSUM-cog) was computed by adding the scores of the twelve cognition questions. A metacognitive strategy use score (TARSUM-meta) was computed by adding the scores of the nine metacognition questions.

Procedure

Administration of the questionnaire and TARSUM took place in the beginning of the school year. Students in each school were first administered a questionnaire containing self-report strategy use, metacognitive knowledge, perceived value, anxiety, and self-efficacy measures in groups of 15 - 20, during one 50 minute session. In the first section, self-report cognitive and metacognitive strategy use items were presented in a random order to avoid response set. The self-report strategy use measure was presented before the metacognitive knowledge items (section 2) so as not to cue students about potentially useful strategies. The third section contained the perceived value and anxiety items which were presented in random order to control for potential response bias. Finally, the fourth section contained items measuring self-efficacy.

The measures were read out to each group by the researcher to ensure standardized presentation and to minimize reading difficulties. A short training sequence preceded the administration of each measure, in order to familiarize students with the use of the response format. Students were instructed to raise their hands if they had difficulty in understanding the item or wanted a particular item repeated. The students were continuously observed in order to ensure that they were comprehending what was being read out and that they were attentive during testing. At the end of the session, students were instructed to check that all items had been answered. After the questionnaires were returned the researcher also ensured that all items were answered by each student.

A number of procedures were followed to improve the validity of the self-report data obtained from students (see Assor & Connell, 1992). Before the administration of the questionnaire, students were briefed on the reasons for conducting the study. In age appropriate terms students were informed about confidentiality of their results, use of grouped findings when results were reported, and that any answers they gave to the questions were acceptable as long as they were reporting what they really believed.

Within a week of completing the questionnaires, the TARSUM was administered to provide a behavioural measure of students' use of strategies. The interviews with each student were conducted by the primary researcher (JP) and four trained interviewers who were all postgraduate psychology students. Interviews typically took 25 minutes to complete. The interviewers were familiarized with the procedure, notation (in recording oral reading

behaviours), and the strategy types. Detailed verbal explanations were provided to the raters and any points of confusion were clarified. The interviewers then listened to a recording of one of the interview sessions (which was pilot tested with a student not included in the study) and were asked to transcribe the session onto the interview coding forms. Points of confusion and disagreement were discussed and clarified.

Students came to the interview room individually and were given a short training session to familiarize them with the use of the monitoring device before the TARSUM was administered. The interviewer demonstrated how the bleep should be used. Students were then provided with three short sample texts and asked to bleep (press the button), while they were reading, when they came to a part of the text that had an error. In addition, they were asked what was wrong with the text at the points they bleeped. If the student did not bleep at the appropriate points, the student was told where he/she should have bleeped and why. During each trial, the interviewer observed whether or not the button was pressed at the appropriate points. If the button had not been pressed, the student was asked to consider the trial again and indicate where a bleep should have been made. The interviewer ensured that students knew when to press the bleep before proceeding to the TARSUM.

During the TARSUM procedure, while the student read each segment of the passage, the interviewer indicated on his/her copy of the passage the points where bleeps, errors, successful and unsuccessful self-corrections were made by the student. The reasons the students gave for bleeping were recorded on the coding form.

While the student answered the prescribed questions, the interviewer scored the student's answer to the comprehension question and recorded the presence of each strategy articulated by the student. If the strategy mentioned was not listed this was described in the "other" category. Students were instructed to continue reading the next segment after answering the questions at each designated point. If they failed to stop at the designated points, they were reminded to do so.

Encouragement was given for effort made and if a student was fully aware that he/she had done poorly on a question, the interviewer would say, "Many students find this one hard, don't worry". However, no indication was given as to whether their answers to the comprehension questions were right or wrong so as not to invalidate subsequent attempts to evaluate their responses to the comprehension questions. If the student's answers were unclear, the interviewer would request for more detail with prompts such as "How do you mean?" or "Tell me more about it".

All interview data were coded and scored by the researcher to obtain a total score each for comprehension, cognitive strategy use, and metacognitive strategy use.

RESULTS

Reliability

Twenty percent ($n = 28$) of each interviewer's coding forms were randomly subjected to reliability checks by the researcher. Reliabilities for occurrence and categorisation of strategies were calculated using the point-by-point percentage agreement method (Cooper et al., 1987). Interrater agreement was calculated on a question-by-question basis by dividing

the number of agreements by the number of agreements plus disagreements, multiplied by 100.

Mean interrater agreement for the occurrence and categorisation of cognitive strategies was 94.5% (range 90.3% - 100%) and 87.6% (range 85.0% - 91.4%) respectively. Mean interrater agreement for the occurrence (mean = 100%) and categorisation (mean = 97.6%, range = 95.7% - 100%) of the monitoring and planning strategies was higher than for cognitive strategies. Interrater agreement was also sought for comprehension scores, bleeps, and for oral reading behaviours. Similar to procedures used in van Kraayenoord (1986), interrater agreement on oral reading behaviours consisted of the number of uncorrected errors and the number of successful and unsuccessful self-corrections where there was agreement as to the site of the error or self-correction. Mean interrater agreement was 91.6% (range = 89.9% - 95.0%) for comprehension scores, 100% for bleeps, 91.5% (range = 87.4% - 96.3%) for successful self-corrections, 90.6% (range = 87.5% - 93.4%) for unsuccessful self-corrections, and 93.7% (range = 91.9% - 97.0%) for uncorrected errors. The percentage agreements for the oral reading behaviours compare favourably with reliability in other New Zealand studies (van Kraayenoord, 1986). In addition, the finding that more bleeps occurred at predetermined points (93.6%) provided a validity check that students were using the bleep correctly.

Reliability checks for the aforementioned categories were again made using an independent rater, a postgraduate psychology student who was not one of the original interviewers. Interrater agreement on the evaluation activities of metacognitive strategy use was also collected. The independent rater received the same training as other interviewers.

Ten percent of the 139 student's recordings ($n = 14$) were randomly selected for the reliability check. Comparisons of agreement were made on several variables using the reliability formula as used previously. Mean reliability coefficients for all variables ranged between 93.2% and 100%. The percentage agreement between raters is as follows: 95.8% and 94.2% for the occurrence and categorisation of cognitive strategies respectively, 100% and 97.3% for the occurrence and categorisation of metacognitive strategies respectively. Interrater agreement for oral reading behaviours included: 95.6% for comprehension scores, 93.2% for successful self-corrections, 94.6% for unsuccessful self-corrections, and 96.2% for uncorrected errors.

Procedural validity was also assessed by the researcher in terms of the appropriateness of the interviewer's verbalizations during the taped section of the interview. Interviewer verbalizations were categorized into correct prompts, incorrect prompts and other verbalizations. Incorrect prompts occurred when (a) student's comprehension answers were prompted or corrected, (b) the interviewer gave some indication that the student had answered the question correctly, (c) where the interviewer prompted the student as to type of strategy that had been used, or (d) the interviewer had prompted the student to bleep at the appropriate points. Other verbalizations referred to verbalizations not related to the TARSUM.

Each verbalisation was categorised and recorded using 15-s intervals in a partial-interval recording procedure (Cooper et al., 1987) for five minutes of 10% of the audiotaped sessions ($n = 14$). The five minute segments were randomly selected from each 25 minute interview session. Rater agreement was calculated on an interval-by-interval basis using point-by-point agreement ratio method (Cooper et al., 1987). High validity was indicated as 94.7% of the interviewer verbalizations were correct, 3.3% were incorrect, and 2% were other

verbalizations. There was 100% interobserver agreement (as determined by the primary researcher and the independent rater) on the categorisation of interviewer verbalizations.

Construct and Discriminant Validity

The scores on the TARSUM-cog ($M = 11.57$, $SD = 1.98$) and the TARSUM-meta ($M = 17.81$, $SD = 3.84$) scales were distributed normally. Qualitative examination of the data revealed that students used a variety of cognitive and metacognitive strategies in reading. In addition, the findings revealed that even though some readers reported using main ideas in retelling the story, they were less successful in distinguishing main ideas from important details. Thus, for effective performance, students may require more than reading-oriented cognitive strategies; they may need metacognitive strategies that oversee the use of cognitive strategies.

Correlations between the TARSUM scores and self-report strategy use, achievement, perceived value, anxiety, efficacy, and metacognitive knowledge were computed and are presented in Table 3. Due to listwise deletion, correlations between the TARSUM measures and the criterion variables were based on a sample of 136. However, for PAT mathematics and PAT reading study skills, correlations were based on a sample of 118 due to some missing data. All correlations were significant, $p \leq .001$ using two-tailed tests.

Validation of TARSUM requires demonstrating theoretically consistent patterns of relationships between the behavioural strategy use measures and external criteria. Consistent with a construct validation approach, this requires that a strategy use measure be highly correlated with criteria to which it is closely related theoretically (convergence) and less highly correlated with other criteria (divergence). This pattern is clearly seen in the correlations between the strategy use variables and the three reading achievement areas (see Table 3). The PAT measures of reading comprehension, vocabulary, and reading study skills were more strongly correlated with the behavioural measures of cognitive/metacognitive strategy use than Mathematics achievement.

Table 3. Correlations between TARSUM strategy scores and criterion constructs

Criterion measures	TARSUM-meta	TARSUM-cog
RSU-meta	.56	.53
RSU-cog	.44	.69
metacognitive knowledge	.47	.37
PAT Reading Comprehension	.53	.61
PAT Vocabulary	.49	.55
PAT Study Skills*	.51	.50
PAT Mathematics*	.36	.40
perceived value	.40	.54
anxiety	-.41	-.52
self-efficacy	.49	.50

All correlations were highly significant ($p \leq .001$) at the two-tailed level.

Correlations between the strategy use subscales for the behavioural (TARSUM-cog, TARSUM-meta) and self-report (RSU-cog, RSU-meta) measures were .46 and .60, respectively.

* Correlations were based on a sample of 118, all other correlations were based on a sample of 136.

Besides the correlations with achievement, the correlations between the TARSUM and RSU strategy use subscales suggest that the TARSUM is a valid measure of cognitive and metacognitive strategy use. The correlational patterns between the self-report and the behavioural strategy use measures were similar to those found in the pilot study. For example TARSUM-meta correlated more strongly with RSU-meta than RSU -cog or TARSUM-cog. Similarly, TARSUM-cog correlated more strongly with RSU-cog than with RSU-meta or TARSUM-meta. In general, the correlations between the TARSUM cognitive/metacognitive strategy measures and their self-report counterpart were higher than the correlation of each of the strategy scores with the other criterion measures. Again as in the pilot study, one of the strongest correlations was found between the self-report strategy use subscales, suggesting a degree of method variance.

The correlations of the cognitive/metacognitive strategy use scales with motivational-related constructs, such as anxiety, perceived value for reading, and self-efficacy, also define a pattern of correlations that supports the construct validity of interpretations based upon responses to the strategy use scales. Students who are more strategically engaged experience low anxiety, view reading positively, and feel more efficacious.

Support for construct validity of the TARSUM is also demonstrated by the statistically significant correlations between metacognitive knowledge and the strategy use measures of TARSUM. These findings suggest that students who engage in strategy use are often more knowledgeable about reading. Also worthy of note is that the criterion measures correlated more strongly with TARSUM-cog than with TARSUM-meta. This finding suggests that cognitive and metacognitive strategy use may be differentially related to achievement and motivation

Predictive validity of the scales was obtained by correlating the TARSUM measures with classroom marks obtained from teachers towards the end of the school year. Classroom marks correlated .47 with TARSUM-meta and .48 with TARSUM-cog. These significant moderate correlations attest to the predictive validity of the subscales.

DISCUSSION

This article describes the development and psychometric characteristics of the TARSUM which was designed to assess both metacognitive and cognitive strategies in the domain of reading. Despite great interest in reading strategy use, relatively few validated tools have been developed for the assessment of this construct. The TARSUM has several unique features. First, it is a reading specific measure of strategy use. Second, it is one of the few behavioural measures that assesses cognitive and metacognitive strategy use independently. Third, it is probably one of the few reading strategy use measures normed with young adolescent children.

Interrater agreement data suggests that the TARSUM can be reliably coded. Procedural validity was also demonstrated. Attempts to enhance the construct validity of the TARSUM were made by deriving the measure from cognitive and metacognitive theory. One of the advantages of the TARSUM over most other strategy use measures is that it is a domain specific rather than a global indicator of strategy use. This is particularly important as there is evidence in educational research for the separation of strategy use between different academic areas (Alexander & Judy, 1988; Lester, 1988). In addition, many researchers and theoreticians emphasize that both generic and domain- or content-dependent strategies must

be identified and taught (e.g., Gagne, 1985; Glaser, 1984).

There has been much debate over differentiating cognitive strategy use from metacognitive strategy use (St. George, van Kraayenoord, & Chapman, 1985). Even though these two types of strategy use have been distinguished empirically in a large number of research studies (e.g., Anthony, 1994; Pokay & Blumenfeld, 1990; Zimmerman & Martinez-Pons, 1990), some researchers still contend that they are indistinguishable (e.g., S. G. Paris, personal communication, 1995). The moderately high correlation between the cognitive and metacognitive strategy subscales ($r = .46$ for TARSUM and $r = .60$ for RSU) found in this study suggest that they are related but still form distinct constructs. Further indirect support is provided by the finding that cognitive strategy use is more strongly related to most of the criterion measures than metacognitive strategy use, suggesting that each strategy use component may be differentially related to reading performance and motivation. This finding is consistent with previous research (Pokay & Blumenfeld, 1990).

There has also been considerable debate in the last two decades about the distinction between the twin components of metacognition: metacognitive knowledge and metacognitive strategy use (Brown, 1987; Brown & Palincsar, 1982; Lawson, 1984). The findings of this study reveal that these components of metacognition are related but were differentially related to cognitive strategy use. As expected, metacognitive knowledge was more strongly related to metacognitive cognitive strategy use than cognitive strategy use. The pattern of findings were consistent with theory and empirical findings (Cornoldi, 1990; Cross & Paris, 1988; Schraw & Dennison, 1994).

Convergent validity for cognitive and metacognitive strategy use was supported by expected correlational patterns between the self-report and behavioural strategy use measures. This was demonstrated by the behavioural strategy use subscales having higher correlations with their self-report counterpart. Several researchers have documented that self-reported strategy use is significantly correlated with observable strategy use (e.g., Justice & Weaver-McDougall, 1989). Concurrent validation of the TARSUM was supported by substantial correlations with criterion measures. The findings relating to the relationships between the TARSUM and motivational-related constructs such as anxiety, perceived value, and self-efficacy were consistent with previous research (Ames & Archer, 1988; Pintrich & De Groot, 1990; Pokay & Blumenfeld, 1990; Tobias, 1985). That is, metacognitive and cognitive strategy use are moderately and positively related to perceived value and efficacy but negatively related to anxiety.

As expected the TARSUM scores correlated higher with the reading achievement measures (comprehension, vocabulary, reading study skills) than with the mathematics achievement measure, suggesting discriminant validity. The statistically significant relationship between metacognitive/cognitive strategy use and achievement is consistent with previous research (Brown, Bransford, Ferrara, & Campione, 1983; Pressley, Symons, Snyder, & Cariglia-Bull, 1989; Pressley et al., 1990).

Finally, the subscales seem to show promising predictive validity. The scales were related to classroom reading performance in the expected directions. These significant, albeit moderate correlations with reading achievement are satisfactory, given the multiplicity of factors that are related to reading achievement.

The present study extends previous research in two primary ways. First, it examines the

relationship between student's reported strategy use on paper-and-pencil instruments with their reported strategy use in individual interviews. Most researchers have theorized that students' actual strategy use is related to their reports in concurrent versus retrospective interviews (Garner, 1987; Pressley et al., 1990). This hypothesis was supported in the present study. Second, unlike previous research which has relied on standardized tests (e.g., Paris & Oka, 1986), this study included both standardized reading achievement test scores and classroom marks as indicators of reading performance. The use of a standardized reading test as the sole measure of comprehension in previous research has raised some concerns (DeFina, Anstendig, & DeLawler, 1991). It is argued that children are generally well aware of their classroom marks while they are often not informed of their standardized test scores. Therefore, if strategy use is presumed to influence performance, it is more likely that this relationship would be more accurately modeled when classroom marks are included as one measure of achievement.

There were a number of limitations to the present study. First, the first study was limited by the sample size of the poor reader group and other special population diagnostic groups and thus precludes conclusive discriminant validity studies at this time. The results of this study are additionally limited by the nature of the strategy use items included on the paper-and-pencil instrument (RSU) and in the TARSUM. It has been argued that the most demanding criterion of construct validity can be applied only when different methods or procedures are used to measure the same construct. To this end the questions about students' strategy use could have been better matched across the questionnaire and interview formats. Questions on the paper-and-pencil instrument related to students' general use of strategies with text, while questions in the interview related to students' use of strategies with a specific text task.

Third, in retrospect, a frequency measure of strategy use, as was employed in the TARSUM, may not provide a comprehensive measure of strategy use because strategies can be used with differing degrees of effectiveness (Reynolds, Wade, Trathen, & Lapan, 1989). In the present study, it was observed on some occasions that although students reported a cognitive strategy that was appropriate for the task, they were unable to answer the question correctly. In other words, although students may self-report appropriate strategy use, they may not use the strategy effectively. This discrepancy between reporting cognitive strategy use and using it effectively was highlighted by Reynolds et al. (1989) who stated that "There seems to be an implicit assumption that if students say they are using a strategy or exhibit behaviours that are seen as reflecting a strategy, then they must be using the strategy effectively" (p. 175).

The TARSUM provided a way to ask readers what strategies they used to obtain the answer; however, it may be that some subjects did not consciously use strategies, but merely "came up with something" to answer the strategy use questions. This is a problem inherent in using self-report data gathered through questioning or probing. In addition, it is possible that different strategies may be used when reading different types of texts or topics, an important qualification when making generalizations from this study. Finally, multitrait, multimethod studies using non-self-report strategy use measures (e.g., parent and teacher ratings or observations) are essential to complement the findings herein. Nonetheless, the results of this initial investigation provide preliminary support for the reliability and validity of the TARSUM in assessing cognitive and metacognitive strategy use for reading in adolescent children.

The TARSUM may have utility in theory based research on strategy use. Theory based research concerning the relationships between cognitive/metacognitive strategy use,

metacognitive knowledge, motivation, and performance have been impeded by the lack of reliable and valid reading strategy use instruments. The TARSUM may also have utility in intervention research which assesses whether teaching reading disabled or poor readers to use strategies produces positive results (Borkowski, Johnston, & Reid, 1987; Oka & Paris, 1987; Pressley, El-Dinary, & Brown, 1992). Assessment of strategies using the TARSUM could be used to identify students' level of strategic engagement and to evaluate the kinds of cognitive and metacognitive strategies students employ in reading particular texts. In addition, changes in scores on the TARSUM may assist in determining the effectiveness of interventions for less strategically oriented children in research or classroom settings. Responses to specific questions in the TARSUM may be useful for the design of individualized treatment programmes for adolescents. In conclusion, the TARSUM shows great promise as a measure for assessing metacognitive and cognitive strategy use in the domain of reading.

REFERENCES

- Afflerbach, P., & Johnston, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading Behavior, 16*, 307-322.
- Alexander, P. A., & Judy, J. E. (1988). The interaction of domain specific and strategic knowledge in academic performance. *Review of Educational Research, 58*, 375-404.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*, 260-267.
- Anthony, G. (1994). Learning strategies in the mathematics classroom: What can we learn from stimulated recall interviews? *New Zealand Journal of Educational Studies, 29*, 127-140.
- Assor, A., & Connell, J. P. (1992). The validity of students' self-reports as measures of performance affecting self-appraisals. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 25-47). Hillsdale, NJ: Erlbaum.
- Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D. L. Forrest-Pressley, G. E. MacKinnon, & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Theoretical perspectives* (Vol. 1, pp. 155-205). Orlando, FL: Academic Press.
- Baker, L. & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, M. Kamil, R. Barr & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- Baumann, J. F., Seifert-Kessell, N., & Jones, L. (1987, December). *Effects of think-aloud instruction on elementary students' ability to monitor their comprehension*. Paper presented at the National Reading Conference, St. Petersburg, FL.
- Beck, L. R., & St. George, R. (1983). The alleged cultural bias of PAT Reading Comprehension and Reading Vocabulary tests. *New Zealand Journal of Educational Studies, 18*, 32-47.
- Blaxall, J., & Willows, D. M. (1984). Reading ability and text difficulty as influences on second graders' oral reading errors. *Journal of Educational Psychology, 76*, 330-341.
- Borkowski, J. G., Carr, M., & Pressley, M. (1987). "Spontaneous" strategy use: Perspectives from metacognitive theory. *Intelligence, 11*, 61-75.
- Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 53-92). Hillsdale, NJ: Erlbaum.
- Borkowski, J. G., Day, J. D., Saenz, D., Dietmeyer, D., Estrada, T. M., & Groteluschen, A. (1992). Expanding the boundaries of cognitive interventions. In B. Y. L. Wong (Ed.), *Contemporary intervention and research in learning disabilities: An international perspective* (pp. 1-21). New York: McGraw-Hill.
- Borkowski, J. G., Johnston, M. B., & Reid, M. K. (1987). Metacognition, motivation, and controlled performance. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 147-173). Hillsdale, NJ: Erlbaum.
- Bos, C. S., & Filip, D. (1984). Comprehension monitoring in learning disabled and average students. *Journal of Learning Disabilities, 17*, 229-233.
- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious

- mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.
- Brown, A. L., Bransford, J.D., Ferrara, R. A., & Campione, J.C. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of child psychology: Cognitive development* (Vol 3, pp. 77-166). New York: Wiley.
- Brown, A., & Palincsar, A. (1982). Inducing strategic knowledge from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities*, 2, 1-18
- Chan, K. S. (1995). Relationship of motivation, strategic learning, and reading achievement in Grades 5, 7, and 9. *Journal of Experimental Education*, 62, 319-339.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.
- Cornoldi, C. (1990). Metacognitive control processes and memory deficits in poor comprehenders. *Learning Disability Quarterly*, 13, 245-255.
- Cross, D. R., & Paris, S. G. (1988). Developmental and instructional analyses of children's metacognition and reading comprehension. *Journal of Educational Psychology*, 80, 131-142.
- DeFina, A. A., Anstendig, L., & DeLawler, K. (1991). Alternative integrated reading/writing assessment and curriculum design. *Journal of Reading*, 34, 354-359.
- Derry, S. J. & Murphy, D. A. (1986). Designing systems that train learning ability: From theory to practice. *Review of Educational Research*, 56, 1-39.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Gagne, E. D. (1985). *The cognitive psychology of school learning*. Boston: Little Brown.
- Garner, R. (1987). *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Review of Educational Research*, 60, 517-529.
- Garner, R., & Alexander, P.A. (1989). Metacognition: Answered and Unanswered questions. *Educational Psychologist*, 24, 143-158.
- Gerber, M. M. (1983). Learning disabilities and cognitive strategies: A case for training or constraining problem solving? *Journal of Learning Disabilities*, 16, 255-260.
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39, 93-104.
- Horrex, J. (1992). *A multicomponential intervention programme for remediating secondary students' reading comprehension difficulties* (Research Affiliateship Scheme Report No. 2). Palmerston North, New Zealand: Massey University, Department of Education.
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22, 255-278.
- Johns, J. L. (1981). *Basic Reading Inventory: Pre-primer to grade eight* (2nd ed.). Dubuque, IA:

Kendall-Hunt.

- Justice, E. M., & Weaver-McDougall, R. G. (1989). Adults' knowledge about memory: Awareness and use of memory strategies across tasks. *Journal of Educational Psychology, 81*, 214-219.
- Lawson, M. (1984). Being executive about metacognition. In J. R. Kirby (Ed.), *Cognitive strategies and educational performance*. Orlando: Academic Press.
- Lester, F. K. (1988). Reflections about mathematical problem-solving research. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (pp. 115-124). Hillsdale: Erlbaum.
- Levin, J. R. (1986). Four cognitive principles of learning-strategy instruction. *Educational Psychologist, 21*, 3-17.
- Lonberger, R. (1988, February). *Effects of training in a self-generated learning strategy on the prose processing abilities of 4th and 6th graders*. Paper presented at the annual meeting of the Eastern Education Association, Savannah, GA.
- McLain, K.V.M, Gridley, B. E., & McIntosh, D (1991). Value of a scale used to measure metacognitive reading awareness. *Journal of Educational Psychology, 85*, 81-87.
- Meichenbaum, D., Burland, S., Gruson, L., & Cameron, R. (1985). Metacognitive assessment. In S. Yussen (Ed.), *The growth of reflection in children* (pp. 3-30). Wisconsin: Academic Press.
- Meyers, J., & Lytle, S. (1986). Assessment of the learning process. *Exceptional Children, 53*, 138-144.
- Neale, M. D. (1989). *Neale Analysis of Reading Ability - Revised: Manual*. Windsor: NFER-Nelson.
- Ng, S. (1979). *Error behavior and self correction in reading and oral language*. Unpublished doctoral dissertation, University of Auckland, Auckland, New Zealand.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
- Oka, E. R., & Paris, S. G. (1987). Patterns of motivation and reading skills in underachieving children. In S. J. Ceci (Ed.), *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (Vol. 2, pp. 115 - 145). Hillsdale, NJ: Erlbaum.
- Paris, S. G. (1991). Assessment and remediation of metacognitive aspects of children's reading comprehension. *Topics in Language Disorders, 12*, 32-50.
- Paris, S.G., & Wasik, B.A., & Turner, J. C. (1991). The development of strategic readers. In R. Barr, M.L. Kamil, P. Mosenthal & P.D. Pearson (Eds.), *Handbook of reading research: Volume II* (pp. 609-640). New York: Longman.
- Paris, S. G., Wasik, B., & Van der Westhuizen, G. (1988). Meta-metacognition: A review of research on metacognition and reading. In J. Readance & S. Baldwin (Eds.), *Dialogues in literacy research* (p. 143-166). Chicago: National Reading Conference.
- Pereira-Laird, J. A. (1996). *Development and validation of a self-report measure of the reading strategy use (RSU) scale*. Manuscript submitted for publication.
- Pereira-Laird, J. A. & Deane, F. P. (1995). *Validation of a Reading Self-efficacy scale for adolescents*. Manuscript submitted for publication.

- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Pokay, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology*, 82, 41-50.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive strategies: Good strategy users coordinate metacognition and knowledge. In R. Vasta & G. Whitehurst (Eds.), *Annals of child development* (Vol. 4, pp. 89-129). Greenwich, CT: JAI Press.
- Pressley, M., El-Dinary, P. B., & Brown, A.L. (1992). Skilled and not-so-skilled reading: good information processing and not-so-good information processing. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 91-127). New York: Academic Press.
- Pressley, M., Forrest-Pressley, D., & Elliot-Faust, D. (1988). What is strategy instructional enrichment and how to study it: Illustrations from research on children's prose memory and comprehension. In F. Weiner & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 101-130). Hillsdale, NJ: Erlbaum.
- Pressley, M., Gaskins, I. W., Cunicelli, E. A., Burdick, N. J., Schaub-Matt, M., Lee, D. S., & Powell, N. (1991). Strategy instruction at benchmark school: A faculty interview study. *Learning Disability Quarterly*, 14, 19-48.
- Pressley, M., Symons, S., Snyder, B. L., & Cariglia-Bull, T. (1989). Strategy instruction research comes of age. *Learning Disability Quarterly*, 12, 16-30.
- Pressley, M., Woloshyn, V., Lysynchuk, L. M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review*, 2, 1-58.
- Reid, N. A. (1993). *Progressive Achievement Tests of Mathematics*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., Croft, A. C., & Jackson, P. F. (1978). *Progressive Achievement Tests: Study skills*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reid, N. A., & Elley, W. B. (1991). *Progressive Achievement Tests of Reading*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Reynolds R. E. & Wade, S. E. (1986). Thinking about thinking about thinking: Reflections on metacognition. *Harvard Educational Review*, 56, 307-317.
- Reynolds, R. E., Wade, S. E., Trathen, W., & Lapan, R. (1989). The selective attention strategy and prose learning. In C. B. McCormick, G. Miller, & M. Pressley (Eds.), *Cognitive strategy research: From basic research to educational applications* (pp. 159-190). New York: Springer-Verlag.
- Rowe, D. W., & Rayford, L. (1984). *Instantiating background knowledge in reading comprehension assessment*. Paper presented at the annual meeting of the National Reading Conference, St. Petersburg Beach, FL.
- Schmitt, M. C. (1990). A questionnaire to measure children's awareness of strategic reading processes. *The Reading Teacher*, 43, 454-461.
- Schneider, W. (1985). Developmental trends in the metamemory-memory behavior relationship: An integrative review. In D. L. Forrest-Pressley, G. E. MacKinnon & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol. 1. Theoretical perspectives* (pp. 57-109). New York: Academic Press.

- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460-475.
- Schunk, D. H. (1986). Verbalization and children's self-regulated learning. *Contemporary Educational Psychology, 11*, 347-369.
- St. George, A., van Kraayenoord, C. E., & Chapman, J. W. (1985). *Cognition and reading of learning disabled and normally achieving intermediate school children* (Research contract No. 55-22-121). Wellington, New Zealand: Massey University, Department of Education.
- Swanson, H.L. (1989). Strategy instruction: Overview of principles and procedures for effective use. *Learning Disability Quarterly, 12*, 3-14.
- Tobias, S. (1985). Test anxiety: Inference, defective skills, and cognitive capacity. *Educational Psychologist, 20*, 135-142.
- van Kraayenoord, C. E. (1986). *Metacognition, reading and causal attributions: A comparison of learning disabled and non learning disabled intermediate school children*. Unpublished doctoral dissertation, Massey University, Palmerston North, New Zealand.
- Wade, S.E. & Trathen, W. (1989). Effect of self-selected study methods on learning. *Journal of Educational Psychology, 81*, 40-47.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315-327). New York: MacMillan.
- Wade, S. E., Trathen, W., & Schraw, G. (1990). An analysis of spontaneous study strategies. *Research Reading Quarterly, 25*, 147-166.
- Williamson, I. E., & Young, F. (1974). The IRI and RMI diagnostic concepts should be synthesized. *Journal of Reading Behavior, 6*, 183-194.
- Winograd, P., & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of Reading Behavior, 14*, 61-76.
- Woods, M. L., & Moe, A. J. (1981). *Analytical Reading Inventory* (2nd ed.). Columbus, Ohio: Bell & Howell.
- Zabucky, K., & Commander, N. E. (1993). Rereading to understand: The role of text coherence and reader proficiency. *Contemporary Educational Psychology, 18*, 442-454.
- Zabucky, K. & Ratner, H. H. (1986). Children's comprehension monitoring and recall of inconsistent stories. *Child Development, 57*, 1401-1418.
- Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology, 80*, 284-290.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Author Note

This article is derived from a Doctoral dissertation completed by the first author under the supervision of the second author. We would like to thank the all the intermediate schools in the Manawatu-Wanganui region for their participation in this study. Correspondence concerning this article should be addressed to Joyce A. Pereira-Laird, Department of Psychology, Massey University, Private Bag 11222, Palmerston North, New Zealand.

APPENDICES

Appendix A: Sample Passage

THE VOLCANO - VERSION 1

The scientists approached the crater's edge fascinated at the prospect of recording the spectacle of an inactive volcano smouldering again. Intent on their photography, they ignored an ominous rumbling. Within seconds, the subterranean cauldron exploded violently ejecting a great quantity of rocks. Fortunately these fell in the direction of the opposite slope.

Greatly alarmed by this premature explosion, the group hastily began the descent. Immediately, fiery boulders from a gigantic avalanche hurtled around. Aware that their apparatus hindered progress, they abandoned all equipment except their precious cameras. Then came an urgent moment. As they were evading flying fragments, one of them was struck off-balance by a rebounding boulder. A lengthy halt would have been disastrous. Everyone was, therefore, very relieved when they found the injuries were superficial. They resumed their dangerous scramble to regain safety just before the surroundings were destroyed. The scientists were pleased that they did not leave any of their equipment behind.

When they finally reached the bottom of the volcano, the scientists hurried towards their cabin which nearby was. One of the young scientists was very seriously injured and was screaming because of the pain. Luckily, there was an ikran waiting near their cabin and the young scientist was whisked away quickly to hospital.

Note:

Notations that were used by the interviewer to record oral reading behaviours:

s^c - self-corrected successfully

s^u - self-corrected unsuccessfully

x - made an error without correcting

b - bleeped

Appendix B: Sample of Interview Coding Form in Investigation 3

Sample cognition question

	Strategies	Observed	Notes
3a. What do you think "subterranean cauldron" (<i>underground boiling pot</i>) refers to in the sentence you've read? (1 mark)	Monitors understanding	_____	_____
	Uses context cues	_____	_____
	Substitution looks or sounds similar	_____	_____
	Mentions others as resources	_____	_____
b. How could you tell? (How did you go about trying to get your answer?)	Mentions dictionary as resource	_____	_____
	Other	_____	_____

Ans: refers to the volcano (1)
refers to larva/boiling liquid (1/2)

Sample monitoring question

	Strategies	Observed	Notes
8. What are you thinking here? (Why did you bleep?) corrected it? yes/no ----- ----- -----	Monitors/ Indicates scrambled words	_____	_____
	Evaluates i.e. makes a judgmental statement about text	_____	_____
	Rereads	_____	_____
	Other	_____	_____

Sample evaluation question

	Strategies	Observed	Notes
11 a. Which paragraph did you think was hardest to read for you? Paragraph _____	Rereads	_____	_____
	Infers based on difficulty of understanding	_____	_____
b. How did you decide which was the hardest?	Infers based on difficulty of pronouncing words	_____	_____
	Other	_____	_____

(CORRESPONDENCE OF PARAGRAPH DIFFICULTY WITH STUDENT'S ACCURACY SCORE OF PARAGRAPH: YES/NO)

Evaluation strategy _____

APPENDIX F

CORRELATION MATRICES FOR THE FOUR HYPOTHESIZED MODELS

Table F1. Correlation matrix of NA readers ($n = 203$) and RD readers ($n = 202$) for Model 1

	inters	contros	stabs	controf	stabf	effica	efficb	msi	ira	pat	class	iq	ethnicity ^a	gender ^a
<i>NA readers</i>														
inters														
contros	.733													
stabs	.731	.755												
controf	.544	.570	.596											
stabf	.518	.522	.537	.599										
effica	.469	.469	.538	.424	.470									
efficb	.506	.469	.503	.348	.323	.602								
msi	.452	.412	.461	.299	.460	.436	.419							
ira	.500	.473	.412	.347	.408	.369	.419	.498						
pat	.349	.400	.420	.332	.338	.408	.605	.422	.363					
class	.262	.316	.337	.281	.248	.320	.482	.356	.364	.683				
iq	.239	.259	.306	.258	.241	.298	.290	.250	.251	.400	.320			
ethnicity ^a	.139	.132 ^b	.159	.134 ^b	.113 ^b	.123 ^b	.243	.159	-.038 ^b	.182	.115 ^b	.101 ^b		
gender ^a	.146	.200	.115 ^b	.045 ^b	.043 ^b	.110 ^b	.189	.086 ^b	.119 ^b	.017 ^b	.094 ^b	-.279	-.029 ^b	
<i>SD</i>	2.789	3.207	3.231	2.570	2.592	.933	16.056	3.665	3.629	18.493	15.097	9.081	.374	.498
<i>RD readers</i>														
inters														
contros	.406													
stabs	.419	.599												
controf	.358	.550	.528											
stabf	.478	.586	.548	.728										
effica	.246	.308	.347	.352	.403									
efficb	.223	.373	.349	.359	.482	.456								
msi	.298	.296	.292	.396	.482	.305	.311							
ira	.240	.197	.286	.326	.407	.222	.285	.445						
pat	.377	.433	.421	.504	.555	.481	.470	.448	.333					
class	.261	.178	.152	.294	.316	.255	.149	.255	.283	.592				
iq	.053 ^b	.140	.122 ^b	.021 ^b	.107 ^b	.222	.197	.144	.078 ^b	.209	.005 ^b			
ethnicity ^a	.034 ^b	.001 ^b	-.073 ^b	-.022 ^b	-.058 ^b	.003 ^b	.039 ^b	-.033 ^b	-.168	-.118 ^b	-.064 ^b	.124 ^b		
gender ^a	.073 ^b	.075 ^b	.069 ^b	.089 ^b	.155	.043 ^b	.100 ^b	.224	.109 ^b	.049 ^b	.107 ^b	-.223	.059 ^b	
<i>SD</i>	3.035	2.683	2.550	2.949	3.161	.631	11.817	3.438	4.285	12.730	15.241	8.162	.501	.494

Key of variables names: inters = internality (locus) dimension of success attributions; contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; controf = controllability dimension of failure attributions; stabf = stability dimension of failure attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R. All correlations statistically significant at $p \leq .05$ except those indicated by ^b.

^a Variables included only in LISREL analyses of models which includes demographic variables.

Table F2. Correlation matrix of NA readers ($n = 199$) and RD readers ($n = 196$) for Model 2

	inters	contros	stabs	controf	stabf	effica	efficb	msi	ira	strat1	strat2	pat	class	iq	ethnicity ^a	gender ^a
<i>NA readers</i>																
inters																
contros	.738															
stabs	.731	.767														
controf	.540	.585	.595													
stabf	.523	.537	.545	.601												
effica	.462	.478	.534	.412	.469											
efficb	.504	.473	.500	.343	.323	.600										
msi	.454	.417	.466	.298	.460	.438	.419									
ira	.495	.481	.410	.336	.403	.358	.418	.503								
strat1	.407	.377	.445	.372	.356	.407	.477	.472	.379							
strat2	.329	.384	.346	.299	.323	.469	.524	.485	.327	.662						
pat	.350	.404	.417	.337	.337	.409	.607	.427	.359	.610	.590					
class	.257	.329	.326	.271	.235	.309	.482	.361	.349	.387	.436	.680				
iq	.243	.262	.311	.262	.240	.301	.291	.250	.256	.295	.297	.403	.326			
ethnicity ^a	.153	.137 ^b	.185	.147	.113 ^b	.138 ^b	.256	.159	-.027 ^b	.202	.183	.201	.142	.099 ^b		
gender ^a	.145	.208	.122 ^b	.032 ^b	.032 ^b	.105 ^b	.190	.081 ^b	.115 ^b	-.007 ^b	.051 ^b	.026 ^b	.098 ^b	-.284	-.043 ^b	
SD	2.802	3.230	3.230	2.560	2.590	.933	16.176	3.697	3.612	1.447	4.848	18.526	14.925	9.166	.373	.498
<i>RD readers</i>																
inters																
contros	.421															
stabs	.423	.599														
controf	.361	.550	.528													
stabf	.493	.581	.545	.730												
effica	.258	.304	.340	.354	.400											
efficb	.242	.385	.357	.367	.484	.465										
msi	.310	.288	.286	.394	.477	.294	.311									
ira	.231	.208	.287	.331	.418	.237	.299	.458								
strat1	.323	.366	.402	.321	.438	.457	.400	.455	.408							
strat2	.317	.272	.316	.315	.418	.348	.307	.455	.370	.654						
pat	.402	.428	.422	.506	.555	.483	.481	.444	.351	.588	.515					
class	.289	.200	.163	.308	.339	.259	.161	.266	.298	.385	.403	.604				
iq	.032	.138	.117	.013	.105	.229	.199	.143	.081	.132	.142	.206	.009			
ethnicity ^a	.008 ^b	-.013 ^b	-.084 ^b	-.032 ^b	-.068 ^b	.002 ^b	.046 ^b	-.039 ^b	-.176	-.091 ^b	-.001 ^b	-.123 ^b	-.044 ^b	.101 ^b		
gender ^a	.074 ^b	.069 ^b	.071 ^b	.086 ^b	.161	.035 ^b	.121 ^b	.224	.122 ^b	.164	.169	.040 ^b	.101 ^b	-.241	.052 ^b	
SD	3.021	2.692	2.570	2.988	3.185	.630	11.765	3.469	4.320	1.281	5.249	12.794	15.168	8.117	.500	.495

Key of variables names: inters = internality (locus) dimension of success attributions; contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; controf = controllability dimension of failure attributions; stabf = stability dimension of failure attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; strat1 = self-report measure of combined strategy use; strat2 = behavioral measure of combined strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R. All correlations statistically significant at $p \leq .05$ except those indicated by ^b.

^a Variables included only in LISREL analyses of models which includes demographic variables.

Table F3. Correlation matrix of NA (*n* = 199) and RD (*n* = 196) readers for Model 3 and Model 4

	inters	contros	stabs	controf	stabf	effica	efficb	knowa	knowb	meta	coga	metb	cogb	pat	class	iq	ethnicity ^a	gender ^a
<i>NA readers</i>																		
inters																		
contros	.738																	
stabs	.731	.767																
controf	.540	.585	.595															
stabf	.523	.537	.545	.601														
effica	.462	.478	.534	.412	.469													
efficb	.504	.473	.500	.343	.323	.600												
knowa	.454	.417	.466	.298	.460	.438	.419											
knowb	.495	.481	.410	.336	.403	.358	.418	.503										
meta	.350	.383	.425	.360	.320	.324	.428	.395	.309									
coga	.355	.263	.340	.279	.295	.386	.396	.425	.351	.499								
metb.	.324	.377	.348	.296	.303	.420	.470	.437	.254	.581	.405							
cogb	.200	.235	.193	.178	.233	.384	.425	.389	.357	.407	.616	.429						
pat	.350	.404	.417	.337	.337	.409	.607	.427	.359	.571	.480	.496	.558					
class	.257	.329	.326	.271	.235	.309	.482	.361	.349	.348	.321	.377	.388	.680				
iq	.243	.262	.311	.262	.240	.301	.291	.250	.256	.261	.250	.250	.279	.403	.326			
ethnicity ^a	.153	.137 ^b	.185	.147	.113 ^b	.138 ^b	.256	.159	-.027 ^b	.138 ^b	.217	.157	.167	.201	.142	.099 ^b		
gender ^a	.145 ^b	.208	.122 ^b	.032 ^b	.032 ^b	.105 ^b	.190	.081 ^b	.115 ^b	.019 ^b	-.034 ^b	.067 ^b	-.006 ^b	.026 ^b	.098 ^b	-.284	-.043 ^b	
SD	2.802	3.230	3.230	2.560	2.590	.933	16.176	3.697	3.612	.889	.781	3.894	1.665	18.526	14.925	9.166	.373	.498
<i>RD readers</i>																		
inters																		
contros	.421																	
stabs	.423	.599																
controf	.361	.550	.528															
stabf	.493	.581	.545	.730														
effica	.258	.304	.340	.354	.400													
efficb	.242	.385	.357	.367	.484	.465												
knowa	.310	.288	.286	.394	.477	.294	.311											
knowb	.231	.208	.287	.331	.418	.237	.299	.458										
meta	.391	.414	.441	.396	.497	.426	.398	.396	.395									
coga	.183	.233	.268	.176	.278	.368	.301	.390	.316	.483								
metb	.289	.225	.260	.279	.382	.325	.269	.417	.308	.550	.438							
cogb	.250	.257	.302	.264	.327	.261	.262	.354	.345	.387	.574	.460						
pat	.402	.428	.422	.506	.555	.483	.481	.444	.351	.550	.472	.459	.424					
class	.289	.200	.163	.308	.339	.259	.161	.266	.298	.347	.319	.390	.275	.604				
iq	.032 ^b	.138 ^b	.117 ^b	.013 ^b	.105 ^b	.229	.199	.143	.081 ^b	.106 ^b	.120 ^b	.102 ^b	.163	.206	.009 ^b			
ethnicity ^a	.008 ^b	-.013 ^b	-.084 ^b	-.032 ^b	-.068 ^b	.002 ^b	.046 ^b	-.039 ^b	-.176	-.035 ^b	-.115 ^b	.021 ^b	-.041 ^b	-.123 ^b	-.044 ^b	.101 ^b		
gender ^a	.074 ^b	.069 ^b	.071 ^b	.086 ^b	.161	.035 ^b	.121 ^b	.224	.122 ^b	.200	.093 ^b	.204	.040 ^b	.040 ^b	.101 ^b	-.241	.052 ^b	
SD	3.021	2.692	2.570	2.988	3.185	.630	11.765	3.469	4.320	.679	.807	3.920	2.126	12.794	15.168	8.117	.500	.495

Key of variables names: inters = internality (locus) dimension of success attributions; contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; controf = controllability dimension of failure attributions; stabf = stability dimension of failure attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; meta = self-report measure of metacognitive strategy use; coga = self-report measure of cognitive strategy use; metb = behavioral measure of metacognitive strategy use; cogb = behavioral measure of cognitive strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R. All correlations statistically significant at $p \leq .05$ except those indicated by ^b.
^a Variables included only in LISREL analyses of models which includes demographic variables.

APPENDIX G

LISREL SPECIFICATION OF HYPOTHESIZED MODELS 1 THROUGH 4

G.1 Optimum Model 1

G.1.1 Notation

G.1.1.1 Observed variables

Let

- x_1 = score on WISC-R measure of intellectual ability
- y_1 = score on measure of controllability dimension of success attributional style
- y_2 = score on measure of stability dimension of success attributional style
- y_3 = score on measure of stability dimension of failure attributional style
- y_4 = score on measure of controllability dimension of failure attributional style
- y_5 = score on measure of internality (locus) dimension of success attributional style
- y_6 = score on general measure of self-efficacy
- y_7 = score on task specific measure of self-efficacy
- y_8 = score on MSI measure of metacognitive knowledge
- y_9 = score on IRA measure of metacognitive knowledge
- y_{10} = score on standardized PAT measure of reading performance
- y_{11} = score on classroom performance measure of reading performance

G.1.1.2 Latent variables

Let

- ξ_1 denote the ability factor
- η_1 denote the attributional style factor
- η_2 denote the self-efficacy factor
- η_3 denote the metacognitive knowledge factor
- η_4 denote the reading comprehension/performance factor

G.1.2 Model specification

G.1.2.1 Measurement model

In LISREL notation the measurement model may be expressed in matrix notation as follows:

$$\begin{matrix}
 y & [11] & = & \Lambda_y & [5] & \eta & [4] & + & \epsilon & [11]
 \end{matrix}$$

$$\begin{bmatrix}
 y_1 \\
 y_2 \\
 y_3 \\
 y_4 \\
 y_5 \\
 y_6 \\
 y_7 \\
 y_8 \\
 y_9 \\
 y_{10} \\
 y_{11}
 \end{bmatrix}
 [11]
 =
 \begin{bmatrix}
 1 & 0 & 0 & 0 \\
 \lambda_{21} & 0 & 0 & 0 \\
 \lambda_{31} & 0 & 0 & 0 \\
 \lambda_{41} & 0 & 0 & 0 \\
 \lambda_{51} & 0 & 0 & 0 \\
 0 & \lambda_{62} & 0 & 0 \\
 0 & 1 & 0 & 0 \\
 0 & 0 & 1 & 0 \\
 0 & 0 & \lambda_{93} & 0 \\
 0 & 0 & 0 & 1 \\
 0 & 0 & 0 & \lambda_{11,4}
 \end{bmatrix}
 [5]
 \begin{bmatrix}
 \eta_1 \\
 \eta_2 \\
 \eta_3 \\
 \eta_4
 \end{bmatrix}
 [4]
 +
 \begin{bmatrix}
 \epsilon_1 \\
 \epsilon_2 \\
 \epsilon_3 \\
 \epsilon_4 \\
 \epsilon_5 \\
 \epsilon_6 \\
 \epsilon_7 \\
 \epsilon_8 \\
 \epsilon_9 \\
 \epsilon_{10} \\
 \epsilon_{11}
 \end{bmatrix}
 [11]$$

$$\mathbf{x}^{[g]} = \Lambda_x^{[g]} \xi^{[g]} + \delta^{[g]}$$

$$\boxed{\mathbf{x}_1}^{[g]} = \boxed{\lambda_{11}}^{[g]} \xi_1^{[g]} + \boxed{\delta_1}^{[g]}$$

where $\lambda_{11} = 1$

The equations for the equality constraints on the measurement parameters are specified as follows:

$$\begin{aligned} \lambda_{11}^{[g1]} &= \lambda_{11}^{[g2]} \\ \lambda_{21}^{[g1]} &= \lambda_{21}^{[g2]} \\ \lambda_{51}^{[g1]} &= \lambda_{51}^{[g2]} \\ \lambda_{62}^{[g1]} &= \lambda_{62}^{[g2]} \\ \lambda_{72}^{[g1]} &= \lambda_{72}^{[g2]} \\ \lambda_{83}^{[g1]} &= \lambda_{83}^{[g2]} \\ \lambda_{93}^{[g1]} &= \lambda_{93}^{[g2]} \\ \lambda_{10,4}^{[g1]} &= \lambda_{10,4}^{[g2]} \\ \lambda_{11,4}^{[g1]} &= \lambda_{11,4}^{[g2]} \end{aligned}$$

where g1 represents the NA group and g2 represents the RD group

G.1.2.2 Structural equation model

$$\begin{aligned} \eta_1^{[g]} &= \gamma_{11}^{[g]} \xi_1^{[g]} + \zeta_1^{[g]} \\ \eta_2^{[g]} &= \gamma_{21}^{[g]} \xi_1^{[g]} + \beta_{21}^{[g]} \eta_1^{[g]} + \zeta_2^{[g]} \\ \eta_3^{[g]} &= \gamma_{31}^{[g]} \xi_1^{[g]} + \beta_{31}^{[g]} \eta_1^{[g]} + \beta_{32}^{[g]} \eta_2^{[g]} + \zeta_3^{[g]} \\ \eta_4^{[g]} &= \gamma_{41}^{[g]} \xi_1^{[g]} + \beta_{43}^{[g]} \eta_3^{[g]} + \beta_{42}^{[g]} \eta_2^{[g]} + \zeta_4^{[g]} \end{aligned}$$

In LISREL notation these equations may be written in matrix form as follows:

$$\boldsymbol{\eta}^{[g]} = \boldsymbol{\Gamma}^{[g]} \boldsymbol{\xi}^{[g]} + \mathbf{B}^{[g]} \boldsymbol{\eta}^{[g]} + \boldsymbol{\zeta}^{[g]}$$

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}^{[g]} = \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ 0 \\ \gamma_{41} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 \\ 0 & \beta_{42} & \beta_{43} & 0 \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix}^{[g]} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \end{bmatrix}^{[g]}$$

The equations for the equality constraints on the structural parameters are specified as follows:

$$\begin{aligned} \gamma_{21}^{[g1]} &= \gamma_{21}^{[g2]} \\ \beta_{21}^{[g1]} &= \beta_{21}^{[g2]} \\ \beta_{31}^{[g1]} &= \beta_{31}^{[g2]} \\ \beta_{32}^{[g1]} &= \beta_{32}^{[g2]} \\ \beta_{43}^{[g1]} &= \beta_{43}^{[g2]} \\ \beta_{42}^{[g1]} &= \beta_{42}^{[g2]} \end{aligned}$$

where g1 represents the NA group and g2 represents the RD group

Since there is only a single exogenous variable,

$$\Phi^{[g]} = \boxed{\phi_{11}}^{[g]}$$

where ϕ_{11} represents the variance of the exogenous variable ξ_1

G.1.2.4 Identification of model

In employing the above constraints between NA and RD readers, the proposed model can be shown to be overidentified. This can be seen from the fact that over the two groups there are 156 observed variances-covariances and 60 model parameters to be estimated, leaving a model with 96 degrees of freedom. The 60 model parameters to be estimated include: (a) 9 coefficients λ_i , (b) 5 coefficients γ_i , (c) 6 coefficients β_i , (d) 2 coefficients of ϕ , (e) 21 variances and 9 covariances of the error terms ϵ_i , and (f) 8 variances of the disturbances ζ_i . Evidence also comes from the literature that this type of model is identified. The measurement component of this model is of the form of a simple confirmatory factor model; in the presence of correlated errors, such models have been shown to be identified by Long (1983a) and Horwood (1987). The structural model component is in the form of a recursive path model and such models have been shown to be identified by Duncan (1975).

G.2 Optimum Model 2

G.2.1 Notation

G.2.1.1 Observed variables

Let

- x_1 = score on WISC-R measure of intellectual ability
- y_1 = score on measure of controllability dimension of success attributional style
- y_2 = score on measure of stability dimension of success attributional style
- y_3 = score on measure of stability dimension of failure attributional style
- y_4 = score on measure of controllability dimension of failure attributional style
- y_5 = score on measure of internality (locus) dimension of success attributional style
- y_6 = score on general measure of self-efficacy
- y_7 = score on task specific measure of self-efficacy
- y_8 = score on MSI measure of metacognitive knowledge
- y_9 = score on IRA measure of metacognitive knowledge
- y_{10} = score on self-report measure of combined strategy use
- y_{11} = score on behavioural measure of combined strategy use
- y_{12} = score on standardized PAT measure of reading performance
- y_{13} = score on classroom performance measure of reading performance

G.2.1.2 Latent variables

Let

- ξ_1 denote the ability factor
- η_1 denote the attributional style factor
- η_2 denote the efficacy factor
- η_3 denote the metacognitive knowledge factor
- η_4 denote the combined strategy use factor
- η_5 denote the reading comprehension/performance factor

G.2.2 Model specification

G.2.2.1 Measurement model

In LISREL notation the measurement model may be expressed in matrix notation as follows:

$$y^{[g]} = \Lambda_y^{[g]} \eta^{[g]} + \epsilon^{[g]}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \\ y_{13} \end{bmatrix}^{[g]} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 & 0 \\ \lambda_{41} & 0 & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda_{93} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda_{11,4} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \lambda_{13,5} \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \end{bmatrix}^{[g]} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \end{bmatrix}^{[g]}$$

$$x^{[g]} = \Lambda_x^{[g]} \xi^{[g]} + \delta^{[g]}$$

$$\begin{bmatrix} x_1 \end{bmatrix}^{[g]} = \begin{bmatrix} \lambda_{11} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} \delta_1 \end{bmatrix}^{[g]}$$

where $\lambda_{11}^{[g]} = 1$

The equations for the equality constraints on the measurement parameters are specified as follows:

$$\lambda_{11}^{[g1]} = \lambda_{11}^{[g2]}$$

$$\lambda_{21}^{[g1]} = \lambda_{21}^{[g2]}$$

$$\lambda_{51}^{[g1]} = \lambda_{51}^{[g2]}$$

$$\lambda_{62}^{[g1]} = \lambda_{62}^{[g2]}$$

$$\lambda_{72}^{[g1]} = \lambda_{72}^{[g2]}$$

$$\lambda_{83}^{[g1]} = \lambda_{83}^{[g2]}$$

$$\lambda_{93}^{[g1]} = \lambda_{93}^{[g2]}$$

$$\lambda_{10,4}^{[g1]} = \lambda_{10,4}^{[g2]}$$

$$\lambda_{11,4}^{[g1]} = \lambda_{11,4}^{[g2]}$$

$$\lambda_{12,5}^{[g1]} = \lambda_{12,5}^{[g2]}$$

$$\lambda_{13,5}^{[g1]} = \lambda_{13,5}^{[g2]}$$

where g1 represents the NA group and g2 represents the RD group

G.2.2.2 Structural equation model

$$\eta_1^{[g]} = \gamma_{11}^{[g]} \xi_1^{[g]} + \zeta_1^{[g]}$$

$$\eta_2^{[g]} = \gamma_{21}^{[g]} \xi_1^{[g]} + \beta_{21}^{[g]} \eta_2^{[g]} + \zeta_2^{[g]}$$

$$\eta_3^{[g]} = \beta_{31}^{[g]} \eta_1^{[g]} + \beta_{32}^{[g]} \eta_2^{[g]} + \zeta_3^{[g]}$$

$$\eta_4^{[g]} = \beta_{42}^{[g]} \eta_2^{[g]} + \beta_{43}^{[g]} \eta_3^{[g]} + \zeta_4^{[g]}$$

$$\eta_5^{[g]} = \beta_{52}^{[g]} \eta_2^{[g]} + \beta_{54}^{[g]} \eta_4^{[g]} + \zeta_5^{[g]}$$

- y_2 = score on measure of stability dimension of success attributional style
 y_3 = score on measure of controllability dimension of failure attributional style
 y_4 = score on measure of stability dimension of failure attributional style
 y_5 = score on measure of internality (locus) dimension of success attributional style
 y_6 = score on general measure of self-efficacy
 y_7 = score on task specific measure of self-efficacy
 y_8 = score on MSI measure of metacognitive knowledge
 y_9 = score on IRA measure of metacognitive knowledge
 y_{10} = score on self-report measure of metacognitive strategy use
 y_{11} = score on self-report measure of cognitive strategy use
 y_{12} = score on behavioural measure of metacognitive strategy use
 y_{13} = score on behavioural measure of cognitive strategy use
 y_{14} = score on standardized PAT measure of reading performance
 y_{15} = score on classroom performance measure of reading performance

G.3.1.2 Latent variables

Let

- ξ_1 denote the ability factor
 η_1 denote the attributional style factor
 η_2 denote the efficacy factor
 η_3 denote the metacognitive knowledge factor
 η_4 denote the metacognitive strategy use factor
 η_5 denote the cognitive strategy use factor
 η_6 denote the reading comprehension/performance factor

G.3.2 Model specification

G.3.2.1 Measurement model

In LISREL notation the measurement model may be expressed in matrix notation as follows:

$$\begin{matrix} y^{[s]} \\ \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \end{matrix} \end{matrix} = \begin{matrix} \Lambda_y^{[s]} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 & 0 & 0 \\ \lambda_{41} & 0 & 0 & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{62} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{93} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda_{12,4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{13,5} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{15,6} \end{bmatrix} \end{matrix} \begin{matrix} \eta^{[s]} \\ \begin{matrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \end{matrix} \end{matrix} + \begin{matrix} \epsilon^{[s]} \\ \begin{matrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{15} \end{matrix} \end{matrix}$$

$$\mathbf{x}^{[s]} = \Lambda_x^{[s]} \boldsymbol{\xi}^{[s]} + \boldsymbol{\delta}^{[s]}$$

$$\begin{bmatrix} X_1 \end{bmatrix}^{[g]} = \begin{bmatrix} \lambda_{11} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} \delta_1 \end{bmatrix}^{[g]}$$

where $\lambda_{11}^{[g]} = 1$

The equations for the equality constraints on the measurement parameters are specified as follows:

$$\begin{aligned} \lambda_{11}^{[g1]} &= \lambda_{11}^{[g2]} \\ \lambda_{21}^{[g1]} &= \lambda_{21}^{[g2]} \\ \lambda_{51}^{[g1]} &= \lambda_{51}^{[g2]} \\ \lambda_{62}^{[g1]} &= \lambda_{62}^{[g2]} \\ \lambda_{72}^{[g1]} &= \lambda_{72}^{[g2]} \\ \lambda_{83}^{[g1]} &= \lambda_{83}^{[g2]} \\ \lambda_{93}^{[g1]} &= \lambda_{93}^{[g2]} \\ \lambda_{10,4}^{[g1]} &= \lambda_{10,4}^{[g2]} \\ \lambda_{11,4}^{[g1]} &= \lambda_{11,4}^{[g2]} \\ \lambda_{12,5}^{[g1]} &= \lambda_{12,5}^{[g2]} \\ \lambda_{13,5}^{[g1]} &= \lambda_{13,5}^{[g2]} \\ \lambda_{14,6}^{[g1]} &= \lambda_{14,6}^{[g2]} \\ \lambda_{15,6}^{[g1]} &= \lambda_{15,6}^{[g2]} \end{aligned}$$

where g1 represents the NA group and g2 represents the RD group

G.3.2.2 Structural equation model

$$\begin{aligned} \eta_1^{[g]} &= \gamma_{11}^{[g]} \xi_1^{[g]} + \zeta_1^{[g]} \\ \eta_2^{[g]} &= \gamma_{21}^{[g]} \xi_1^{[g]} + \beta_{21}^{[g]} \eta_1^{[g]} + \zeta_2^{[g]} \\ \eta_3^{[g]} &= \beta_{31}^{[g]} \eta_1^{[g]} + \beta_{32}^{[g]} \eta_2^{[g]} + \zeta_3^{[g]} \\ \eta_4^{[g]} &= \beta_{42}^{[g]} \eta_2^{[g]} + \beta_{43}^{[g]} \eta_3^{[g]} + \zeta_4^{[g]} \\ \eta_5^{[g]} &= \beta_{52}^{[g]} \eta_2^{[g]} + \beta_{53}^{[g]} \eta_3^{[g]} + \zeta_5^{[g]} \\ \eta_6^{[g]} &= \gamma_{61}^{[g]} \xi_1^{[g]} + \beta_{62}^{[g]} \eta_2^{[g]} + \beta_{64}^{[g]} \eta_4^{[g]} + \beta_{65}^{[g]} \eta_5^{[g]} + \zeta_6^{[g]} \end{aligned}$$

In LISREL notation these equations may be written in matrix form as follows:

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \end{bmatrix}^{[g]} = \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ 0 \\ 0 \\ 0 \\ \gamma_{61} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 & 0 & 0 \\ 0 & \beta_{42} & \beta_{43} & 0 & 0 & 0 \\ 0 & \beta_{52} & \beta_{53} & 0 & 0 & 0 \\ 0 & \beta_{62} & 0 & \beta_{64} & \beta_{65} & 0 \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \end{bmatrix}^{[g]} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \end{bmatrix}^{[g]}$$

The equations for the equality constraints are specified as follows:

$$\begin{aligned} \gamma_{21}^{[g1]} &= \gamma_{21}^{[g2]} \\ \beta_{21}^{[g1]} &= \beta_{21}^{[g2]} \\ \beta_{31}^{[g1]} &= \beta_{31}^{[g2]} \\ \beta_{32}^{[g1]} &= \beta_{32}^{[g2]} \\ \beta_{42}^{[g1]} &= \beta_{42}^{[g2]} \\ \beta_{43}^{[g1]} &= \beta_{43}^{[g2]} \\ \beta_{52}^{[g1]} &= \beta_{52}^{[g2]} \\ \beta_{53}^{[g1]} &= \beta_{53}^{[g2]} \end{aligned}$$

$\Theta_{\delta} [g]$, the matrix representing the variances and covariances of the measurement errors δ_i is given by:

$$\Theta_{\delta} [g] = [\theta_{11}] [g]$$

where $\theta_{11} [g^1] = 33.605$ and $\theta_{11} [g^2] = 32.9$

The variance covariance matrix of the disturbances is given by:

$$\Psi [g] = \begin{bmatrix} \psi_{11} & & & & & & \\ 0 & \psi_{22} & & & & & \\ 0 & 0 & \psi_{33} & & & & \\ 0 & 0 & 0 & \psi_{44} & & & \\ 0 & 0 & 0 & \psi_{54} & \psi_{55} & & \\ 0 & 0 & 0 & 0 & 0 & \psi_{66} & \end{bmatrix} [g]$$

G.3.2.4 Identification of model

The model is overidentified. This can be seen from the fact that there are 272 observed variances-covariances and 94 model parameters to be estimated, leaving a model with 178 degrees of freedom. The 94 model parameters to be estimated include: (a) 11 coefficients λ_i , (b) 5 coefficients γ_i , (c) 12 coefficients β_i , (d) 2 coefficients of ϕ , (e) 29 variances and 21 covariances of the error terms ϵ_i , (f) 12 variances and 2 covariances of the disturbances ζ_i . Evidence comes from the literature that this type of model is identified. The measurement component of this model is of the form of a simple confirmatory factor model; in the presence of correlated errors for the observed x and y variables, such models have been shown to be identified by Long (1983a) and Horwood (1987). Since the matrix Ψ is not diagonal, the structural component of the model is no longer recursive (Long, 1983a). However, in this instance the estimation of the one off diagonal element of the Ψ matrix does not affect the identifiability of the model. See for example, Hanushek and Jackson (1977) for a discussion of the identifiability of similar non-recursive models.

G.4 Optimum Model 4

G.4.1 Notation

G.4.1.1 Observed variables

Let

x_1 = score on WISC-R measure of intellectual ability

y_1 = score on measure of controllability dimension of success attributional style

y_2 = score on measure of stability dimension of success attributional style

y_3 = score on measure of stability dimension of failure attributional style

y_4 = score on measure of controllability dimension of failure attributional style

y_5 = score on measure of internality dimension of success attributional style

y_6 = score on general measure of self-efficacy

y_7 = score on task specific measure of self-efficacy

y_8 = score on MSI measure of metacognitive knowledge

y_9 = score on IRA measures of metacognitive knowledge

y_{10} = score on self-report measure of metacognitive strategy use

y_{11} = score on self-report measure of cognitive strategy use

y_{12} = score on behavioural measure of metacognitive strategy use

- y_{13} = score on behavioural measure of cognitive strategy use
 y_{14} = score on standardized PAT measure of reading performance
 y_{15} = score on classroom performance measure of reading performance

G.4.1.2 Latent variables

Let

- ξ_1 denote the ability factor
 η_1 denote the success attributional style factor
 η_2 denote the failure attributional style factor
 η_3 denote the efficacy factor
 η_4 denote the metacognitive knowledge factor
 η_5 denote the metacognitive strategy use factor
 η_6 denote the cognitive strategy use factor
 η_7 denote the reading comprehension/performance factor

G.4.2 Model specification

G.4.2.1 Measurement model

In LISREL notation the measurement model may be expressed in matrix notation as follows:

$$y^{[g]} = \Lambda_y^{[g]} \eta^{[g]} + \epsilon^{[g]}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \end{bmatrix}^{[g]} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{73} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{94} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{12,5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{13,6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{15,7} \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix}^{[g]} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \\ \epsilon_8 \\ \epsilon_9 \\ \epsilon_{10} \\ \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{15} \end{bmatrix}^{[g]}$$

$$x^{[g]} = \Lambda_x^{[g]} \xi^{[g]} + \delta^{[g]}$$

$$\begin{bmatrix} x_1 \end{bmatrix}^{[g]} = \begin{bmatrix} \lambda_{11} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} \delta_1 \end{bmatrix}^{[g]}$$

where $\lambda_{11}^{[g]} = 1$

The equations for the equality constraints on the measurement parameters are specified as follows:

$$\lambda_{11}^{[g^1]} = \lambda_{11}^{[g^2]}$$

$$\begin{aligned}
\lambda_{21}^{[g1]} &= \lambda_{21}^{[g2]} \\
\lambda_{51}^{[g1]} &= \lambda_{51}^{[g2]} \\
\lambda_{32}^{[g1]} &= \lambda_{32}^{[g2]} \\
\lambda_{42}^{[g1]} &= \lambda_{42}^{[g2]} \\
\lambda_{63}^{[g1]} &= \lambda_{63}^{[g2]} \\
\lambda_{73}^{[g1]} &= \lambda_{73}^{[g2]} \\
\lambda_{84}^{[g1]} &= \lambda_{84}^{[g2]} \\
\lambda_{94}^{[g1]} &= \lambda_{94}^{[g2]} \\
\lambda_{10,5}^{[g1]} &= \lambda_{10,5}^{[g2]} \\
\lambda_{11,5}^{[g1]} &= \lambda_{11,5}^{[g2]} \\
\lambda_{12,6}^{[g1]} &= \lambda_{12,6}^{[g2]} \\
\lambda_{13,6}^{[g1]} &= \lambda_{13,6}^{[g2]} \\
\lambda_{14,6}^{[g1]} &= \lambda_{14,6}^{[g2]} \\
\lambda_{15,6}^{[g1]} &= \lambda_{15,6}^{[g2]}
\end{aligned}$$

where g1 represents the NA group and g2 represents the RD group

G.4.2.2 Structural equation model

$$\begin{aligned}
\eta_1^{[g]} &= \gamma_{11}^{[g]} \xi_1^{[g]} + \zeta_1^{[g]} \\
\eta_2^{[g]} &= \gamma_{21}^{[g]} \xi_1^{[g]} + \zeta_2^{[g]} \\
\eta_3^{[g]} &= \gamma_{31}^{[g]} \xi_1^{[g]} + \beta_{31}^{[g]} \eta_1^{[g]} + \beta_{32}^{[g]} \eta_2^{[g]} + \zeta_3^{[g]} \\
\eta_4^{[g]} &= \gamma_{41}^{[g]} \xi_1^{[g]} + \beta_{42}^{[g]} \eta_2^{[g]} + \beta_{43}^{[g]} \eta_3^{[g]} + \zeta_4^{[g]} \\
\eta_5^{[g]} &= \beta_{53}^{[g]} \eta_3^{[g]} + \beta_{54}^{[g]} \eta_4^{[g]} + \zeta_5^{[g]} \\
\eta_6^{[g]} &= \beta_{63}^{[g]} \eta_3^{[g]} + \beta_{64}^{[g]} \eta_4^{[g]} + \zeta_6^{[g]} \\
\eta_7^{[g]} &= \gamma_{71}^{[g]} \xi_1^{[g]} + \beta_{73}^{[g]} \eta_3^{[g]} + \beta_{75}^{[g]} \eta_5^{[g]} + \beta_{76}^{[g]} \eta_6^{[g]} + \zeta_7^{[g]}
\end{aligned}$$

In LISREL notation these equations may be written in matrix form as follows:

$$\eta^{[g]} = \Gamma^{[g]} \xi^{[g]} + B^{[g]} \eta^{[g]} + \zeta^{[g]}$$

$$\begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix}^{[g]} = \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \\ 0 \\ 0 \\ 0 \\ \gamma_{71} \end{bmatrix}^{[g]} \begin{bmatrix} \xi_1 \end{bmatrix}^{[g]} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_{42} & \beta_{43} & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_{53} & \beta_{54} & 0 & 0 & 0 \\ 0 & 0 & \beta_{63} & \beta_{64} & 0 & 0 & 0 \\ 0 & 0 & \beta_{73} & 0 & \beta_{75} & \beta_{76} & 0 \end{bmatrix}^{[g]} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \\ \eta_5 \\ \eta_6 \\ \eta_7 \end{bmatrix}^{[g]} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \\ \zeta_7 \end{bmatrix}^{[g]}$$

The equations for the equality constraints are specified as follows:

$$\begin{aligned}
\gamma_{31}^{[g1]} &= \gamma_{31}^{[g2]} \\
\beta_{42}^{[g1]} &= \beta_{42}^{[g2]} \\
\beta_{43}^{[g1]} &= \beta_{43}^{[g2]} \\
\beta_{53}^{[g1]} &= \beta_{53}^{[g2]} \\
\beta_{54}^{[g1]} &= \beta_{54}^{[g2]} \\
\beta_{63}^{[g1]} &= \beta_{63}^{[g2]}
\end{aligned}$$

where $\theta_{14,14}$, the variance of the measurement error for the PAT performance indicator = 8.18

$\Theta_{\delta} [g]$, the matrix representing the variances and covariances of the measurement errors δ_i is given by:

$$\Theta_{\delta} [g] = [\theta_{11}] [g]$$

where $\theta_{11} [g^1] = 33.605$ and $\theta_{11} [g^2] = 32.9$

The variance covariance matrix of the disturbances is given by:

$$\Psi [g] = \begin{bmatrix} \psi_{11} & & & & & & & \\ \psi_{21} & \psi_{22} & & & & & & \\ 0 & 0 & \psi_{33} & & & & & \\ 0 & 0 & 0 & \psi_{44} & & & & \\ 0 & 0 & 0 & 0 & \psi_{55} & & & \\ 0 & 0 & 0 & 0 & \psi_{65} & \psi_{66} & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \psi_{77} & \end{bmatrix} [g]$$

G.4.2.4 Identification of model

The model is overidentified. This can be seen from the fact that there are 272 observed variances-covariances and 96 model parameters to be estimated, leaving a model with 176 degrees of freedom. The 96 model parameters to be estimated include: (a) 8 coefficients $\lambda_i \gamma$, (b) 7 coefficients γ_i , (c) 15 coefficients β_i , (d) 2 coefficients of ϕ , (e) 29 variances and 17 covariances of the error terms ϵ_i , (f) 12 variances and 2 covariances of the disturbances ζ_i . Evidence comes from the literature that this type of model is identified. The measurement component of this model is of the form of a simple confirmatory factor model; in the presence of correlated errors for the observed x and y variables, such models have been shown to be identified by Long (1983a) and Horwood (1987). Since the matrix Ψ is not diagonal, the structural component of the model is no longer recursive (Long, 1983a). However, in this instance the estimation of the two off diagonal elements of the Ψ matrix does not affect the identifiability of the model. See for example, Hanushek and Jackson (1977) for a discussion of the identifiability of similar non-recursive models.

APPENDIX H

DESCRIPTION OF FINDINGS FOR NESTED SEQUENCE OF MODELS FOR HYPOTHESIZED MODELS 1 THROUGH 4

H.1 Model 1

This section presents a description of a series of competing or nested models that were tested for both the measurement and structural parts of Model 1.

H.1.1 Test of competing models for the measurement model

A series of distinct hierarchical measurement models were tested for hypothesized Model 1. The chi-square values, associated degrees of freedom and probability levels for the models that were evaluated are presented in Table H1. Also reported are four measures of overall adequacy of each model: goodness-of-fit index (GFI), the normed fit index (NFI), the parsimonious normed fit index (PNFI), and the standardized root mean square residual (RMR). The chi-square difference and its associated degrees of freedom and probability level are also presented.

Model 1.1 allowed six errors to correlate in the measurement model for NA readers and three errors to correlate in the measurement model for RD readers. A similar factor loading matrix across groups was expected, since the same indicators were used in assessing the latent variables in both groups. More specifically, it was predicted that, in general, the construct pattern for NA and RD readers is similar. If true, this hypothesis would indicate that unit increases in true scores led to the same increments in measured variables for RD readers as for NA readers. This hypothesis was tested in Model 1.1, by constraining all the construct loadings for NA and RD readers to be equal. If these constraints do not significantly erode the fit of the model to the data, one may conclude that NA and RD readers have a common construct pattern for these variables. However, it was noted in the preliminary analyses that the failure attribution loadings differed across the groups; loadings were higher for RD readers suggesting that the association between the observed failure attributions and the latent construct of general attributional style is stronger for RD readers than for NA readers. Consequently, it seems more probable that most of the construct patterns would be similar across groups except for failure attributions. This hypothesis was also tested (Model 1.2). It should be noted that the fit statistics and parameter estimates are based on the analysis of covariance matrices, except the standardised root mean square residual (RMR), which was based on the analysis of correlation matrices (see Marsh, 1990, for more elaboration on this point).

The chi-square of Model 1.1 was highly significant and thus this model was rejected. The difference in chi-square coefficients for Models 1.1 and 1.2 is itself distributed as chi-square. The decrease in chi-square was significant, $\Delta\chi^2(2, N= 405) = 92.31, p < .01$, indicating that freeing the construct loadings of the failure attributions resulted in a substantial improvement in the fit of the model. As a result, Model 1.2 which yielded $\chi^2(87, N= 405) = 76.29$, GFI, NFI = .96, PNFI = .64, RMR = .04, was accepted on statistical grounds ($p = .79$) as producing a good fit for these data. From this information it can be concluded that, in general, NA and RD readers have similar factor patterns but differ in their report of failure

attributions. These results show some departure from the original specification of total factorial invariance for factor loadings. An examination of the modification indices for Model 1.2 indicated that two error covariances may also be nonzero. These covariances were not freed to avoid overfitting of the model.

The next alternative model (Model 1.3) tested the relationships among the five latent variables (the true-score constructs) in the NA and RD reader groups. In LISREL, relationships among the latent variables are represented by the factor covariance/correlation matrix. In developing the measurement model above, the values of the factor covariance matrices had not been equated across groups. Thus, the relationships among the underlying true scores could take any value and vary between groups. In Model 1.3, in addition to the constraints in Model 1.2, the values of the phi matrix were constrained to be equal across groups. This specification tested the assumption that the true scores had the same pattern of variances and covariances in the NA and RD reader groups. As seen in Table H1, this model provided a poor fit to the data, with chi-square being highly significant. This finding implies that the groups may vary somewhat in the relationships of cognitive, metacognitive, and motivational variables in predicting reading achievement.

In conclusion, Model 1.2 provided the best fit to the data.

Table H1. Simultaneous tests for the measurement part of Model 1

Competing models	χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR	$\Delta\chi^2$	Δdf	<i>p</i>	Accept?
1.1 Model 1 with all factor loadings invariant across groups, and with 6 and 3 error covariances freed in the NA and RD reader groups, respectively	168.6	89	.00	.92	.92	.62	.10	-	-	-	No
1.2 Model 1.1 with loadings of failure attributions free across groups	76.29	87	.79	.96	.96	.64	.04	92.31	2	< .01	Yes
1.3 Model 1.2 with all factor variances and covariances invariant across groups	144.43	102	.00	.94	.93	.72	.16	-	-	-	No

Tests were carried out hierarchically; each model was tested against the last accepted model.

H.1.2 Test of competing models for the structural model

The approach taken here was to fit an initial base model (with structural paths unconstrained across groups) and then to compare the model with several alternative models which could be considered nested in this base model. Each model was tested against the last accepted model. The chi-square values, associated degrees of freedom and probability levels for the models that were evaluated in this section are presented in Table H2. The summary of goodness-of-fit statistics (GFI, NFI, PNFI, RMR) together with the chi-square difference and its associated degrees of freedom and probability level are also presented. In general, the nested models presented in this section provided a test of the hypotheses that were presented in section 6.2.1.1 of Chapter 6.

Before testing the hypothesized model presented in Figure 4 (see p. 86), an initial model was estimated to test the causal relationships proposed in Figure 4, but with the structural parameters allowed to vary across the groups. This model tested the hypothesis that all structural relationships among the variables will vary in magnitude across groups. The fit statistics listed in Table H2 suggest that the initial model provided a good fit to the data. Such fit statistics compare the residual differences between the fitted covariance matrix (the matrix implied by the model) and the sample covariance matrix (the matrix used to analyze the data). Chi-square was small in relation to the degrees of freedom, $\chi^2 (89, N= 405) = 77.58, p = .80$, suggesting that this model provided a plausible explanation to the data. Specifically, the probability above .80 suggests that the model can be accepted as a reasonable explanation of the data.

The results of this initial model (Model 1.1) confirmed the preliminary findings that ability does not directly predict metacognitive knowledge in both groups and hence, this path was fixed to zero in all subsequent analyses. The influence of ability on metacognitive knowledge was not direct (as hypothesized), but was significantly mediated by attributional style and self-efficacy. As hypothesized, ability did not significantly predict attributions or performance in the RD reader group. Contrary to predictions, the path efficacy to performance differed across groups. This path was significant for RD readers but not for NA readers, and suggests that the efficacy to performance path should not be constrained equal across groups as hypothesized in Figure 4. Examination of the structural parameters indicated that there were no substantial differences between these parameters across groups except for the following paths: ability to attributions, ability to performance, and efficacy to performance.

Next, the hypothesized model, as depicted in Figure 4, was estimated. This model (Model 1.2) differed from Model 1.1 in that: (a) equality constraints were placed on the structural parameters of the following paths: ability to efficacy, attributions to efficacy, attributions/efficacy to metacognitive knowledge, metacognitive knowledge to performance, and efficacy to performance; and (b) the path ability to metacognitive knowledge was fixed to zero for both groups. Model 1.2 also fitted the data well but the change in chi-square statistic suggested Model 1.2 resulted in a significant deterioration in fit, which led to its rejection. Hence, it was necessary to consider an alternative model specification. A variation on the previous model hypothesized that the path between efficacy and performance would differ across groups for the following reasons: (a) it has been suggested by researchers that motivational variables may vary in importance in determining performance across age and skill levels (e.g., Oka & Paris, 1987); (b) the covariances of the latent constructs between efficacy and performance are markedly different for the two groups; (c) examination of the modification indices in Model 1.2 indicates that an improved fit would be attained if this path were unconstrained; and (d) the parameter estimates are substantially different across groups for this path in Model 1.1. Thus, model 1.3 tested this hypothesis.

Table H2. Fit indices for nested sequence of Theoretical Model 1

Competing models	χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR	$\Delta\chi^2$	Δdf	<i>p</i>	Accept?
1.1 Hypothesized Model 1 with no constraints on structural parameters across groups	77.58	89	.80	.96	.96	.65	.040	-	-	-	Yes
1.2 Hypothesized Model 1.1 with ability to metacognitive knowledge path fixed to zero, and equality constraints placed on six of the structural parameters across groups	99.47	97	.41	.95	.95	.70	.064	21.89	8	<.01	No
1.3 Model 1.2 with efficacy to performance path unconstrained across groups	86.15	96	.75	.96	.96	.70	.051	8.57	7	>.2	Yes
1.4 Model 1.3 with all structural parameters constrained to be equal across groups	103.71	99	.35	.95	.95	.71	.061	17.56	3	<.01	No
1.5 Model 1.3 with ability to attributions path constrained to be equal	91.75	97	.63	.96	.96	.70	.056	5.6	1	<.02	No
1.6 Model 1.3 with ability to performance path constrained to be equal	94.85	97	.54	.96	.96	.68	.050	0.98	2	>.5	No
1.7 Model 1.3 with efficacy to metacognitive knowledge path unconstrained across groups	82.40	95	.82	.96	.96	.69	.040	3.75	1	>.05	No
1.8 Model 1.3 with metacognitive knowledge to performance path unconstrained across groups	84.37	95	.77	.96	.96	.69	.052	1.78	1	>.1	No
1.9 Model 1.3 with attributions to efficacy path unconstrained across groups	86.12	95	.73	.96	.96	.69	.050	0.03	1	>.5	No
1.10 Model 1.3 with ability to efficacy path fixed to zero	101.28	97	.36	.95	.95	.70	.056	15.13	1	<.01	No

Tests were carried out hierarchically; each model was tested against the last accepted model. For all models except Model 1.1, the path, ability to metacognitive knowledge was fixed to zero.

Relaxing the constraint on the path between efficacy and performance across groups resulted in a marked improvement in the model's fit. The chi-square measure of this revised model's fit (Model 1.3) was 86.15 with 96 degrees of freedom, indicating a very good fit of the model to the data ($p = .75$). Additionally, the goodness-of-fit indices (GFI = .96; NFI = .96, PNFI = .70), and in particular the RMR (of .051), showed a marked improvement of the revised model over the hypothesized model. The difference in chi-square values between this model and Model 1.1 yielded $\Delta\chi^2(7, N= 405) = 8.57$, which was statistically nonsignificant ($p > .2$) at the .05 level of confidence. Consequently, Model 1.3 provided a significantly better fit to the data than the unconstrained model, implying that only the paths between ability and attributions, ability and performance, and efficacy and performance, differentiated NA from RD readers, with all other paths being equal.

Although Model 1.3 provided a good fit to the data, an even more constrained model might also provide an adequate fit. In particular, a model was tested that specified no differences between the structural models for the two groups (antithesis of hypothesis no. 6). The only differences in paths between Model 1.3 and Model 1.4 are the unconstrained paths of efficacy to performance, ability to attributions, and ability to performance. The chi-square difference, $\Delta\chi^2(3, N= 405) = 17.56$, $p < .01$, between the two models is significant, leading to the rejection of Model 1.4. The results support the hypothesis that not all structural effects are similar across groups.

Several other competing models were also tested. These were elaborations of Model 1.3 resulting from the constraining or freeing of one or more causal paths across groups. Model 1.5 allowed for the path between ability and attributions to be constrained equal across groups. This constraint was added to test the antithetical hypothesis that RD readers may not differ quantitatively in the relationship between ability and attributions. This model produced a significant deterioration in fit and was thus rejected. As hypothesized in the present study, this finding suggests that the intellectual ability of RD readers is not predictive of adaptive attributional style.

The next competing model (Model 1.6) tested the hypothesis that the relationship between ability and performance is similar for both NA and RD readers, by constraining this path to be equal across groups. Carr et al. (1991) assumed that ability predicted reading performance for both NA and RD readers. However, according to the definition of reading disability, there should be a discrepancy between ability and performance for RD readers. The results of Model 1.9 produced a significant deterioration in fit, $\Delta\chi^2(2, N= 405) = 0.98$, $p > .5$, resulting in it being rejected in favour of Model 1.3. These results indicate that the relationship between ability and performance differs across groups.

In each of the next three competing models, a constraint on one of the causal paths was freed in Model 1.3, to test whether the relationship between the variables for RD readers was weaker than that for NA readers. For instance, Model 1.7 allowed for the path between efficacy and metacognitive knowledge (as well as the original paths from ability to attributional belief, ability to performance, and efficacy to performance) to be unconstrained to test the hypothesis that the efficacy expectations of RD readers are weakly related to metacognitive knowledge (antithesis of hypothesis no. 10). This hypothesis was tested in light of recent empirical findings (Pintrich et al., 1994) that suggest efficacy may not predict metacognitive knowledge for RD readers. For this model, $\chi^2(95, N = 405) = 82.40$, $p = .82$,

and the goodness-of-fit indices, GFI, and NFI are similar to Model 1.3. Although the RMR for this model indicates that it is a slightly better fit, when compared to Model 1.3, the PNFI for this model indicates it is less parsimonious than Model 1.3. Given that there were no significant differences in the chi-square statistic between these models, Model 1.7 was rejected in favour of the more parsimonious model (Model 1.3). The results support the hypothesis that the relationship between efficacy and metacognitive knowledge is similar across groups.

Model 1.8 allowed for the path between metacognitive knowledge and performance to be unconstrained. This path was freed to test the hypothesis that the metacognitive knowledge of RD readers is weakly related to performance. Some theorists have suggested that RD readers possess metacognitive knowledge, but being inactive learners, they do not spontaneously apply this metacognitive knowledge to improve performance (Short & Ryan, 1984; Torgesen, 1979). For this model, the difference in chi-square (between Model 1.8 and Model 1.3), $\Delta\chi^2(1, N=405) = 1.78, p > .1$, was statistically nonsignificant, leading to the acceptance of the more parsimonious model which is Model 1.3.

Model 1.9 allowed for the path between attributions and efficacy to be free to test the hypothesis that the relationship between attributional beliefs and self-efficacy varies across groups (antithesis of hypothesis no. 12). Conflicting empirical findings have been found on the attributions of RD readers. The difference in the chi-square statistic was 0.03 with one degree of freedom which was statistically nonsignificant ($p > .5$), thus leading to the rejection of the model.

In the last competing model (Model 1.10), the path between ability and efficacy was fixed to zero. Carr et al. (1991) suggested that the link between ability and the expectancy construct (efficacy, or self esteem in the case of Carr et al.'s study) is indirect and is mediated by attributional beliefs. In contrast, theoretical and empirical findings suggest that ability directly predicts efficacy (Bandura, 1986; Earley & Lituchy, 1991). These competing models were tested. Model 1.10 yielded a highly significant chi-square difference statistic. Thus, the additional constraint led to a significant deterioration in the fit of the model, $\Delta\chi^2(1, N=405) = 15.13, p < .01$, which resulted in its rejection. These findings suggest that, in addition to the mediating effect of attributions, ability also has a direct and significant effect on efficacy.

In conclusion, Model 1.3 provided the best fit to the data.

H.2 Model 2

This section provides a description of the sequence of nested models that were tested for the structural part of hypothesized Model 2. In general these nested models provided a test of the hypotheses that were presented in section 6.2.2.1 of Chapter 6. Competing models for the measurement part of Model 2 were not tested as the construct patterns of the factors have been established to a large extent in Measurement Model 1.

Before testing the hypothesized model, the least restrictive model was estimated to test the causal relationships proposed in Figure 5 (see p. 89), where the structural parameters were allowed to vary across the groups. The results of Model 1 suggested that the path ability to

Table H3. Fit indices for nested sequence of Theoretical Model 2

Competing models	χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR	$\Delta\chi^2$	Δdf	<i>p</i>	Accept?
2.1 Hypothesized Model 2 with no constraints on structural parameters across groups	124.24	131	.65	.95	.95	.69	.044	-	-	-	Yes
2.2 Hypothesized Model 2.1 with equality constraints placed on six of the structural parameters across groups	134.71	137	.54	.95	.95	.72	.071	10.47	6	>.05	Yes
2.3 Model 2.2 with path from metacognitive knowledge to combined strategy use unconstrained across groups	130.36	136	.62	.95	.95	.71	.057	4.35	1	<.05	No, borderline
2.4 Model 2.2 with path from combined strategy use to performance constrained to be equal across groups	140.09	138	.43	.95	.95	.72	.070	5.38	1	<.05	No
2.5 Model 2.2 with path from attributions to combined strategy use set free in both groups	126.81	135	.68	.95	.95	.71	.057	7.90	2	<.02	No, theoretically not interpretable

Tests were carried out hierarchically; each model was tested against the last accepted model.

metacognitive knowledge should be fixed to zero for Model 2. The fit statistics listed in Table H3 suggest that Model 2.1 provided a good fit to the data. Chi-square was small in relation to the degrees of freedom, $\chi^2(131, N=395) = 124.24$, GFI, NFI = .95, PNFI = .69, RMR = .044, and was not statistically significant ($p = .65$), suggesting that this model provided a plausible explanation to the data. This base model is guaranteed to yield the lowest value of chi-square because it has the fewest restrictions. All other models are nested within this base model, because one can form all other models by placing or removing one or more restrictions on the base model (i.e., fixing equality constraints on one or more of the structural paths).

Next, the hypothesized model as depicted in Figure 5 (see p. 89) was tested, with some minor modifications. The findings of Model 1 suggested that hypothesized Model 2 required respecification and the following a posteriori modifications were made: (a) the ability to metacognitive knowledge path was fixed to zero; and (b) the efficacy to performance path was freed across groups. As a consequence, Model 2.2 differed from Model 2.1 in that equality constraints were imposed on the structural parameters of the following paths: ability to efficacy, attributions to efficacy, attributions/efficacy to metacognitive knowledge, metacognitive knowledge to combined strategy use, and efficacy to combined strategy use.

Thus, the hypothesized model as depicted in Figure 5, was estimated with the abovementioned a posteriori respecifications. The chi-square measure of this respecified model's fit (Model 2.2) was 134.71 with 137 degrees of freedom, indicating a very good fit of the model to the data ($p = .95$). The difference in chi-square values between Model 2.2 and Model 2.1 yielded $\Delta\chi^2(6, N=395) = 10.47$, which was statistically nonsignificant ($p > .05$). Thus, Model 2.2 provided a significantly better fit to the data than the unconstrained model. This implies that the relationships among the variables are essentially the same across groups, except those between ability and attributions, ability and performance, efficacy and performance, and combined strategy use and performance.

Three other competing models were also tested. These models were elaborations of Model 2.2 resulting from the freeing or constraining of one or more causal paths across groups. Model 2.3 is Model 2.2 with the unconstraining of the metacognitive knowledge to combined strategy use path; Model 2.4 is Model 2.2 with the constraining of the combined strategy use to performance path; and Model 2.5 is Model 2.2 with the freeing of the attributional style to combined strategy use path.

Model 2.3 was tested to ascertain whether the relationship between metacognitive knowledge and combined strategy use varied between RD and NA readers (antithesis of hypothesis no. 4). It has been suggested that in contrast to NA readers who use their metacognitive knowledge in the form of strategies, RD readers possess the metacognitive knowledge, but do not expend the effort to use the strategies (Short & Ryan, 1984). For Model 2.3, $\chi^2(136, N=395) = 130.36$, $p = .62$, the goodness-of-fit indices, GFI and NFI, are similar to Model 2.2. Although the standardized RMR for this model indicated that it was a better fit, when compared to Model 2.2, the PNFI for this model indicated it was slightly less parsimonious than Model 2.2. Although there was a significant difference in the chi-square statistic between these models, the difference was considered to be of borderline significance, leading to the rejection of Model 2.3. These findings support the hypothesis that the magnitude of

the relationship between metacognitive knowledge and combined strategy use is similar across groups.

The relationship between combined strategic behaviour and reading performance is unclear in the research literature, particularly with regard to RD readers. Hence, Model 2.4 was tested to determine whether the relationship between strategic behaviour and reading performance was similar across groups (antithesis of hypothesis no. 5). For this model, $\chi^2(138, N= 395) = 140.09, p = .43$, and the difference in chi-square was significant, leading to its rejection. These findings suggest that although the relationship between metacognitive knowledge and performance is similar across groups as was indicated in Model 1, combined strategy use differs in its mediating effect across groups.

In Model 2.5, the path from attributions to combined strategy use was added, on the basis of the criteria of substantive interpretability and expected gain in model fit (MacCallum, 1986). The modification indices in Model 2.2 suggested that an even better fit could be obtained by adding a direct path from attributional style to combined strategy use. Moreover, Borkowski and Muthukrishna (1992) theorized a direct relationship between attributions and strategy use (more specifically metacognitive strategy use). Addition of this path resulted in an improved fit. Although this improvement in fit was significant and resulted in an extremely well fitting model, an anomalous negatively valenced coefficient for the direct path of attributions to combined strategy use resulted. This result was contrary to theory, logic, and the cumulative findings of empirical research pertaining to this relationship. Thus, although the data suggested a significant improvement in fit by including the path attributions to combined strategy use, freeing this path adversely affected the precision and interpretability of the path estimate. Model 2.5 was therefore rejected.

In conclusion, Model 2.2 provided the best fit to the data.

H.3 Model 3

This section presents a summary of the nested sequence of models tested for the structural part of hypothesized Model 3. This sequence of models, for the most part, provided a test of the hypotheses presented in section 6.2.3.1 of Chapter 6.

As was the case for Models 1 and 2, a series of hierarchically nested models was tested. Before testing hypothesized Model 3, the least restrictive model was estimated to test the causal relationships proposed in Figure 6 (see p. 91), but with the structural parameters allowed to vary across the groups. As before, the path ability to metacognitive knowledge was fixed to zero. The fit statistics listed in Table H4 suggested that Model 3.1 provided a good fit to the data. Chi-square was small in relation to the degrees of freedom, $\chi^2(169, N= 395) = 145.28, p = .91$, suggesting that this model provided a reasonably good fit to the data. Examination of the structural parameter estimates indicated that, in addition to the ability to attributions, ability to performance, and efficacy to performance paths, the cognitive strategy use to performance path was also dissimilar across groups.

Next, the hypothesized model as depicted in Figure 6, with the abovementioned a posteriori respecifications, was estimated. Thus, Model 3.2 differed from the previous model in that equality constraints were imposed on the structural parameters of the following paths: ability to efficacy, attributions to efficacy, attributions/efficacy to metacognitive knowledge, metacognitive knowledge to metacognitive/cognitive strategy use, metacognitive strategy use

Table H4. Fit indices for nested sequence of Theoretical Model 3

Competing models		χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR	$\Delta\chi^2$	Δdf	<i>p</i>	Accept?
3.1	Hypothesized Model 3 with no constraints on structural parameters across groups	145.28	169	.91	.95	.95	.67	.043	-	-	-	Yes
3.2	Hypothesized Model 3.1 with equality constraints placed on nine of the structural parameters across groups	159.01	178	.84	.95	.95	.70	.074	13.73	9	>.2	Yes
3.3	Model 3.2 with path from cognitive strategy use to performance constrained equal across groups	166.92	179	.73	.95	.94	.70	.072	7.91	1	<.01	No
3.4	Model 3.2 with path from metacognitive strategy use to performance unconstrained across groups	157.79	177	.85	.95	.95	.70	.074	1.22	1	>.2	No
3.5	Model 3.2 with path from metacognitive knowledge to metacognitive strategy use unconstrained across groups	157.28	177	.85	.95	.95	.70	.069	1.73	1	>.1	No
3.6	Model 3.2 with path from metacognitive knowledge to cognitive strategy use unconstrained across groups	156.52	177	.86	.95	.95	.70	.065	2.49	1	>.1	No
3.7	Model 3.2 with path from efficacy to metacognitive strategy use unconstrained across groups	157.38	177	.85	.95	.95	.70	.069	1.63	1	>.1	No
3.8	Model 3.2 with path from efficacy to cognitive strategy use unconstrained across groups	157.00	177	.86	.95	.95	.70	.066	2.01	1	>.1	No
3.9	Model 3.2 with path attributions to metacognitive strategy use set free but constrained equal across groups	157.48	177	.85	.95	.95	.70	.073	1.53	1	>.2	No
3.10*	Model 3.2 with path from metacognitive strategy use to cognitive strategy use set free but constrained equal across groups	160.13	179	.84	.95	.95	.71	.075	-	-	-	No, for theoretical reasons

Tests were carried out hierarchically; each model was tested against the last accepted model except for Model 3.10* which was not a nested model.

to performance, and efficacy to metacognitive/cognitive strategy use. Model 3.2 differed from the hypothesized model depicted in Figure 6 in the following ways: the path ability to metacognitive knowledge was fixed to zero; and the path efficacy to performance was allowed to vary across groups.

The chi-square measure of this respecified model's fit (Model 3.2) was 159.01 with 178 degrees of freedom, indicating an extremely good fit of the model to the data ($p = .95$). Three fit indices suggested an adequate fit (GFI = .95; NFI = .95, PNFI = .70), although the RMR of .074 suggested significant deterioration in fit. The difference in chi-square values between this model and Model 3.1 yielded $\Delta\chi^2(9, N= 395) = 13.73$, which was statistically nonsignificant ($p > .2$). Thus, the hypothesized model provided a significantly better fit to the data than the unconstrained model, implying that the relationships among the variables are essentially the same across groups except for those between ability and attributions, ability and performance, efficacy and performance, and cognitive strategy use and performance.

Although Model 3.2 provided a good fit to the data, an even more constrained model might also provide an adequate fit. In particular, a model was tested that specified no difference between the cognitive strategy to performance path across groups (antithesis of hypothesis no. 6). The only difference in paths between Model 3.3 and Model 3.2 is the constrained path of cognitive strategy use to performance. An equality constraint was added to test the hypothesis that RD readers did not differ in magnitude from NA readers in the relationship between cognitive strategy use and performance. Model 3.3 essentially tests the antithesis of the "inactive" and "inefficient" learner views. As both these views implicate strategic deficiencies in RD children, these views would predict that cognitive strategy use does not predict performance for RD children. The chi-square difference, $\Delta\chi^2(1, N= 395) = 7.91$, $p < .01$, indicated that the constraint imposed by Model 3.3 resulted in a significant deterioration of fit. Thus, it can be concluded that the relationship between cognitive strategy use and performance varies between the groups. The findings of this model also provides support for both the "inactive" and "inefficient" learner views of reading disability, which suggest that cognitive strategy use does not predict reading performance for RD readers.

Several other competing models were also tested. These were elaborations of Model 3.2 resulting from the freeing of one or more causal paths across groups. One of the hypotheses proposed for Model 3 was that the relationship between metacognitive strategy use and performance would be similar across groups (hypothesis no. 7). Model 3.4 tested the antithesis of hypothesis 7 by freeing the path between metacognitive strategy use and performance across groups. Given that there were no significant differences in the chi-square statistic, $\Delta\chi^2(1, N= 395) = 1.22$, $p > .2$, between these models, Model 3.4 was rejected in favour of the more parsimonious Model 3.2. The findings suggested that RD readers do not differ from NA readers in the strength of the relationship between metacognitive strategy use and performance.

Models 3.5 and 3.6 were tested to clarify the competing views of researchers on the relationship between metacognitive knowledge and metacognitive/cognitive strategy use for RD readers (see Chapter 6, section 6.1.2.2). Generally, for NA readers, the evidence for a positive correlation between metacognitive knowledge and metacognitive/cognitive strategy use outweighs the counterevidence. However, conflicting findings have been found for RD readers. Thus, Models 3.5 and 3.6 were tested to determine whether the relationships between

metacognitive knowledge and metacognitive/cognitive strategy use were similar across groups (antitheses of hypotheses no. 8 and 9). Model 3.5 allowed for the metacognitive knowledge to metacognitive strategy use path to be unconstrained. For Model 3.5, $\chi^2(177, N= 395) = 157.28, p = .85$, and the difference in chi-square (between Model 3.5 and Model 3.2) was statistically nonsignificant, leading to the acceptance of the more parsimonious Model 3.2. Model 3.6 tested whether the relationship between metacognitive knowledge and cognitive strategy use was similar across groups. This model tested the competing predictions of the "inactive" and "inefficient" learner views. There was no significant difference in the chi-square statistic between Models 3.6 and 3.2, and thus Model 3.6 was rejected in favour of the more parsimonious Model 3.2. These results provide some indirect support for the "inefficient" learner view which posits that metacognitive knowledge predicts cognitive strategy use for RD children.

The next two alternative models tested the antitheses of hypotheses 10 and 11, proposed for Model 3. The hypotheses tested in the following two models were that the strength of the effect of efficacy on metacognitive/cognitive strategy use would differ across groups. Model 3.7 allowed for the path between efficacy and metacognitive strategy use to be unconstrained across groups whilst Model 3.8 allowed for the path between efficacy and cognitive strategy use to be unconstrained across groups. For both these models, the differences in chi-square were statistically nonsignificant, leading to the acceptance of the more parsimonious model, Model 3.2. As hypothesized, these findings imply that the relationships between efficacy and metacognitive/cognitive strategy use are essentially the same across groups.

The next competing model, Model 3.9, was tested to investigate Borkowski and Muthukrishna's (1992) theoretical proposition that attributional style predicts metacognitive strategy use or executive processes. In Model 3.9, the path attributional style to metacognitive strategy use was freed but constrained equal across groups. The difference in chi-square was statistically nonsignificant, $\Delta\chi^2(1, N= 395) = 1.53, p > .1$, leading to its rejection and the acceptance of the more parsimonious model, Model 3.2. These findings do not support Borkowski and Muthukrishna's prediction that attributional style predicts metacognitive strategy use.

The last competing model was tested to investigate Borkowski et al.'s (1989, 1990, 1992) theoretical proposition that metacognitive strategy use predicts cognitive strategy use. To achieve this the correlation between the disturbances of the metacognitive and cognitive strategy use latent variables was replaced in Model 3.10 by a direct path from metacognitive to cognitive strategy use. This direct path was constrained to be equal across groups. Both models appear to fit the data equally well. However, releasing the path from metacognitive strategy use to cognitive strategy use, rendered some of the other theoretically predicted paths (viz., efficacy to cognitive strategy use, and metacognitive knowledge to cognitive strategy use) to be statistically nonsignificant. This model was considered unacceptable, given previous research findings supporting these relationships, and hence Model 3.10 was rejected.

In conclusion, Model 3.2 was accepted as providing the best fit to the data.

H.4 Model 4

This section presents a description of the competing models that were tested for hypothesized Model 4. A number of competing models in this section provided a test of the hypotheses presented in section 6.2.4.1 of Chapter 6.

As with previous models, the approach taken here was to fit an initial structural model (which had no equality constraints on its structural paths across groups), followed by several alternative models. The competing models tested were largely exploratory in nature. This was because no hypotheses were formulated regarding the differential effects of the attributional subcomponents, given the lack of theory and empirical findings in the achievement literature.

Before testing the hypothesized model, the least restrictive model was estimated to test the causal relationships proposed in Figure 7 (see p. 94). However, the structural parameters were allowed to vary across the groups. Again, the path ability to metacognitive knowledge was fixed to zero. The fit statistics (see Table H5) suggested that Model 4.1 provided a good fit to the data. Chi-square was small in relation to the degrees of freedom, $\chi^2(166, N=395) = 152.19, p = .77$, suggesting that this model provided a good fit to the data (GFI, NFI = .95, PNFI = .66, RMR = .044).

The analysis of Model 4.1 revealed a number of interesting results which were contrary to predictions. Success attributions significantly predicted efficacy for NA readers but not for RD readers. Conversely, failure attributions significantly predicted efficacy for RD readers but not for NA readers. Only failure attributions were a significant predictor of metacognitive knowledge in both groups. These findings suggest that the hypothesized model in Figure 7 may warrant further respecification.

Next, the a posteriori hypothesized model was estimated, with respecifications similar to Model 3. This model differed from Model 4.1 in that equality constraints were imposed on the structural parameters of the following paths: ability to efficacy, success/failure attributions to efficacy, success/failure attributions to metacognitive knowledge, efficacy to metacognitive knowledge, metacognitive knowledge to metacognitive/cognitive strategy use, metacognitive strategy use to performance, and efficacy to metacognitive/cognitive strategy use.

The chi-square value of the fit of Model 4.2 was 164.47 with 177 degrees of freedom, indicating a good fit of the model to the data ($p = .74$). The difference in chi-square values between Model 4.2 and Model 4.1 yielded $\Delta\chi^2(10, N=395) = 12.28$, which was statistically nonsignificant ($p > .2$). Thus, the hypothesized Model 4.2 provided a better fit to the data than the unconstrained structural model (Model 4.1).

The findings of Model 4.1 revealed that success attributions may not significantly predict metacognitive knowledge. The next model (Model 4.3) was tested to ascertain whether fixing to zero the path success attributions to metacognitive knowledge would result in a deterioration of fit. For this model, $\chi^2(178, N=395) = 165.09, p = .75$, and the difference in chi-square (between Model 4.3 and Model 4.2) was statistically nonsignificant, leading to the acceptance of the more parsimonious Model 4.3. These findings suggest that success attributions do not significantly predict metacognitive knowledge.

Table H5. Fit indices for nested sequence of Theoretical Model 4

Competing models	χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR	$\Delta\chi^2$	Δdf	<i>p</i>	Accept?
4.1 Hypothesized Model 4 with no constraints on structural parameters across groups	152.19	166	.77	.95	.95	.66	.044	-	-	-	Yes
4.2 Hypothesized Model 4.1 with equality constraints placed on eleven of the structural parameters across groups	164.47	177	.74	.94	.95	.70	.050	12.28	10	>.2	Yes
4.3 Model 4.2 with path success attributions to metacognitive knowledge fixed to zero for both groups	165.09	178	.75	.94	.95	.70	.050	0.62	1	>.3	Yes
4.4 Model 4.3 with path failure attributions to metacognitive knowledge unconstrained across groups	162.96	177	.77	.94	.95	.70	.053	2.13	1	>.05	No
4.5 Model 4.3 with paths success and failure attributions to efficacy unconstrained across groups	159.67	176	.81	.95	.95	.69	.050	5.42	2	>.05	Yes, borderline

Tests were carried out hierarchically; each model was tested against the last accepted model.

Two other competing models were also tested. These were elaborations of Model 4.3 resulting from the freeing or constraining of one or more causal paths across groups. Model 4.4 allowed for the path between failure attributions and metacognitive knowledge to be unconstrained across groups. This constraint was removed to test the hypothesis that the relationship between failure attribution style and metacognitive knowledge would vary across groups (antithesis of hypothesis no. 4). Given that there was no significant difference in the chi-square statistic, $\Delta\chi^2(1, N= 395) = 2.13, p > .05$, between these models, Model 4.4 was rejected in favour of the more parsimonious Model 4.3. Hence, it can be concluded that the relationship between failure attributions and metacognitive knowledge is similar across groups.

The next competing model (Model 4.5) tested the hypotheses that the paths success attributional style to efficacy and failure attributional style to efficacy would differ in magnitude across groups (antitheses of hypotheses no. 5 and 6). The difference in chi-square was considered to be of borderline significance, leading to its acceptance. This model was accepted in light of the findings of Model 4.1 which suggested that success and failure attributions have differential effects on efficacy, and that these effects vary in magnitude across groups.

The above findings between attributional style and efficacy seem contradictory to the findings of Models 1, 2 and 3 at first sight. Models 1 through 3 suggested that the magnitude of the effect of general attributional style on efficacy is equal across groups. This inconsistency may be explained by suggesting that the joint or common influence of success and failure attributions on efficacy is equal, but the individual effects of the subcomponents vary in magnitude across groups.

In conclusion, Model 4.5 provided the best fit to the data.

APPENDIX I

ERROR VARIANCE-COVARIANCE MATRICES FOR STRUCTURAL MODELS 1 THROUGH 4

Table II. Common metric completely standardized solution of error variances and covariances for NA and RD readers for revised theoretical Model 1

	1	2	3	4	5	6	7	8	9	10	11	12
<i>NA readers</i>												
1. contros	0.31											
2. stabs		0.28										
3. stabf			0.49									
4. controf			0.16	0.48								
5. inters					0.27							
6. effica			0.10			0.56						
7. efficb							0.50					
8. msi			0.12					0.55				
9. ira			-0.10						0.41			
10. pat							0.25			0.28		
11. class							0.15				0.42	
12. iq												0.44*
<i>RD readers</i>												
1. contros	0.48											
2. stabs	0.13	0.44										
3. stabf			0.26									
4. controf				0.38								
5. inters				-0.11	0.76							
6. effica						0.36						
7. efficb							0.40					
8. msi								0.44				
9. ira									0.77			
10. pat										0.03*		
11. class							-0.10				0.65	
12. iq												0.45*

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

All residuals are significant at $p \leq .05$ (two-tailed).

* Unstandardized values of these parameters were fixed.

Table 12. Common metric completely standardized solution of error variances and covariances for NA and RD readers for revised theoretical Model 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>NA readers</i>														
1. contros	0.28													
2. stabs		0.28												
3. stabf			0.48											
4. controf			0.15	0.48										
5. inters					0.28									
6. effica			0.10			0.59								
7. efficb							0.47							
8. msi			0.11					0.51						
9. ira		-0.09							0.43					
10. strat1										0.44				
11. strat2		-0.08							-0.07		0.28			
12. pat							0.21			0.11		0.21		
13. class							0.15						0.41	
14. iq														0.45*
<i>RD readers</i>														
1. contros	0.44													
2. stabs	0.13	0.45												
3. stabf			0.27											
4. controf				0.38										
5. inters				-0.12	0.73									
6. effica						0.35								
7. efficb							0.37							
8. msi								0.40						
9. ira									0.74					
10. strat1										0.24				
11. strat2											0.45			
12. pat												0.03*		
13. class							-0.10						0.64	
14. iq														0.44*

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; strat1 = self-report measure of combined strategy use; strat2 = behavioural measure of combined strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

All residuals are significant at $p \leq .05$ (two-tailed).

* Unstandardized values of these parameters were fixed.

Table 13. Common metric completely standardized solution of error variances and covariances for NA and RD readers for revised theoretical Model 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>NA readers</i>																
1. contros	0.28															
2. stabs		0.28														
3. stabf			0.48													
4. controf			0.11	0.48												
5. inters					0.28											
6. effica			0.10			0.57										
7. efficb					0.07		0.47									
8. msi			0.11					0.51								
9. ira	-0.10								0.43							
10. meta						-0.10				0.57						
11. coga					0.09					0.13	0.39					
12. metb									-0.12			0.41				
13. cogb	-0.09												0.26			
14. pat							0.21			0.16				0.19		
15. class							0.15								0.46	
16. iq																0.45 ^a
<i>RD readers</i>																
1. contros	0.44															
2. stabs	0.12	0.44														
3. stabf			0.38													
4. controf				0.29												
5. inters				-0.15	0.72											
6. effica						0.35										
7. efficb			0.05 ^b				0.38									
8. msi								0.37								
9. ira									0.75							
10. meta								-0.09		0.23						
11. coga				-0.10							0.34					
12. metb												0.56				
13. cogb												0.12	0.65			
14. pat														0.03 ^a		
15. class							-0.10					0.10			0.65	
16. iq																0.44 ^a

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; meta = self-report measure of metacognitive strategy use; coga = self-report measure of cognitive strategy use; metb = behavioural measure of metacognitive strategy use; cogb = behavioural measure of cognitive strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

All residuals are significant at $p \leq .05$ (two-tailed).

^a Unstandardized values of these parameters were fixed.

Table I4. Common metric completely standardized solution of error variances and covariances for NA and RD readers for revised theoretical Model 4

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>NA readers</i>																
1. contros	0.27															
2. stabs		0.28														
3. stabf			0.34													
4. controf				0.32												
5. inters					0.28											
6. effica				0.11		0.56										
7. efficb					0.06		0.48									
8. msi			-0.12					0.49								
9. ira		-0.09	-0.07						0.42							
10. meta						-0.10				0.58						
11. coga					0.09					0.13	0.39					
12. metb								-0.12				0.40				
13. cogb		-0.08											0.26			
14. pat							0.21			0.16				0.20		
15. class							0.14								0.46	
16. iq																0.45 ^a
<i>RD readers</i>																
1. contros	0.34															
2. stabs		0.33														
3. stabf			0.38													
4. controf				0.25												
5. inters					0.72											
6. effica						0.35										
7. efficb							0.37									
8. msi								0.38								
9. ira									0.73							
10. meta								-0.09		0.24						
11. coga											0.34					
12. metb												0.56				
13. cogb													0.12	0.64		
14. pat															0.03 ^a	
15. class							-0.09					0.10			0.64	
16. iq																0.44 ^a

Key of variables names: contros = controllability dimension of success attributions; stabs = stability dimension of success attributions; stabf = stability dimension of failure attributions; controf = controllability dimension of failure attributions; inters = internality (locus) dimension of success attributions; effica = general measure of self-efficacy; efficb = task-specific measure of self-efficacy; msi = MSI measure of metacognitive knowledge; ira = IRA measure of metacognitive knowledge; meta = self-report measure of metacognitive strategy use; coga = self-report measure of cognitive strategy use; metb = behavioural measure of metacognitive strategy use; cogb = behavioural measure of cognitive strategy use; pat = standardized measure of reading comprehension; class = classroom measure of reading performance; iq = intellectual ability as measured by WISC-R.

All residuals are significant at $p \leq .05$ (two-tailed).

^a Unstandardized values of these parameters were fixed.

APPENDIX J

SUPPLEMENTARY ANALYSES: THE IMPACT OF ETHNICITY AND GENDER ON THE FOUR HYPOTHESIZED MODELS

Many researchers have failed to report the ethnic and gender memberships of their subjects, despite the finding that ethnicity and gender differences occur between NA and RD readers. Furthermore, empirical findings indicate that these demographic variables should be considered, particularly when research includes an examination of motivational beliefs and strategy use (Eccles et al., 1983; S. Graham & Long, 1986; Wigfield & Asher, 1984).

Empirical research on the effect of ethnicity on motivational beliefs is scarce; however, there is some evidence which suggests that ethnicity influences academic performance. In contrast, research on gender differences is prolific. Although findings are inconsistent across studies, girls generally report lower estimates of their academic abilities and lower achievement expectations, even if they perform as well as or better than boys (Eccles et al., 1983). These expectancy related perceptions have the strongest influence among junior high school students (Berndt & Miller, 1990; Meece et al., 1990), and may have a direct or indirect impact on students' use of reading strategies. However, recent research suggests that the gender differences in students' achievement related perceptions are quite small in magnitude (Hyde, Fennema, Ryan, Frost, & Hopp, 1990) and represent situational effects rather than enduring dispositions (Deaux & Major, 1987).

As the aforementioned relationships involving demographic variables of ethnicity and gender are of less interest to this study, these variables are included as control variables. It is important to note that the results of the present study revealed overall mean gender and ethnicity differences on some of the variables under study. Hence, one secondary goal of this study is to determine whether the inclusion of these control variables appreciably changes the pattern of relationships proposed in the four hypothesized models (see Figures 4 - 7, Chapter 6).

J.1 Hypothesis

It is expected that:

1. The inclusion of gender and ethnicity variables will not change the basic hypothesized relationships amongst the variables in Models 1 through 4.

J.2 Results: LISREL analyses of hypothesized models which included ethnicity and gender variables

Causality was further built into the four hypothesized models (Figures 4 - 7, Chapter 6) by the inclusion of ethnicity and gender to determine the extent to which the structure of the models was affected by the inclusion of these demographic variables. Ethnicity and gender were each specified by a single indicator. Because these measures were relatively objective measures, they were assumed to be measured without error. The observed variables were

therefore made isomorphic with the underlying theoretical constructs (i.e., the observed variables were given a loading of 1.0 and a residual of 0). Simultaneous multisample analyses were conducted for each of the four models following the specifications of the best fitting models in the main analyses (see Appendix G). However, there was a small difference in the specification of the measurement models which contained the demographic variables: the IRA measure of metacognitive knowledge was assumed to be influenced by ethnicity.

Table J1. Fit indices for Measurement and Structural Models with demographic variables

Hypothesized models	χ^2	<i>df</i>	<i>p</i>	GFI	NFI	PNFI	RMR
<i>Measurement models only</i>							
Model 1	101.81	114	.79	.96	.96	.60	.037
Model 2	140.41	152	.74	.95	.95	.60	.040
Model 3	158.21	191	.96	.95	.95	.59	.040
Model 4	149.31	178	.94	.95	.95	.55	.038
<i>Full measurement and structural models</i>							
Model 1	111.89	123	.75	.96	.95	.64	.043
Model 2	158.86	168	.68	.95	.94	.66	.059
Model 3	184.98	213	.92	.95	.94	.66	.063
Model 4	185.47	208	.87	.95	.94	.64	.046

Preliminary analyses of these models indicated that there was a confounding effect of ethnicity on the IRA measure of metacognitive knowledge. Inspection of the means indicated that non-Europeans tended to have lower metacognitive knowledge scores on this measure suggesting that Europeans and non-Europeans tended to respond differently to this measure. This methodological problem was controlled for by allowing ethnicity to directly influence the IRA measure of metacognitive knowledge. This resulted in a significant decrease in chi-square, suggesting improved model fit. No substantial differences in the factor loadings of the latent constructs or the structural parameters occurred as a result of this modification; that is, it improved fit but did not alter measurement properties. Table J1 presents the fit indices of the measurement and structural systems of Model 1 through 4. The goodness-of-fit indices indicate that all four models fit extremely well to the data.

A comparison of the parameter estimates and the squared multiple correlations of the factors for the structural model with and without demographic variables is presented for each model and for each group (see Tables J2 through J5). In Tables J2 through J5, ability, ethnicity and gender represent the latent independent or exogenous variables whilst attributions, efficacy, metacognitive knowledge, strategy use, and performance represent the latent endogenous or predictor variables. In each pair of rows in the table, the non-bracketed numbers show the regression (Beta) coefficients and associated *t* values (italicized) for the predictors of that latent construct in that group. For comparative purposes, the corresponding Beta values and *t* values (italicized) for the model excluding gender and ethnicity are shown in brackets.

Table J2. Comparison of path weights and square multiple correlations for NA and RD readers in Model I with and without demographic variables

Dependent variable	Independent variables							R^2
	attributions	efficacy	metacognitive knowledge	performance	ability	ethnicity	gender	
<i>NA readers</i>								
attributions					.53 (.42) <i>5.11 (4.47)</i>	.11 <i>1.57</i>	.36 <i>4.51</i>	.30 (.18)
efficacy	.56 (.63) <i>8.70 (9.89)</i>				.31 (.24) <i>4.25 (3.85)</i>	.11 <i>1.82</i>	.23 <i>3.23</i>	.63 (.58)
metacognitive knowledge	.51 (.54) <i>4.78 (5.03)</i>	.41 (.36) <i>3.41 (3.15)</i>				-.01 <i>-0.07</i>	-.04 <i>-0.06</i>	.71 (.71)
performance		.22 (.24) <i>1.44 (1.97)</i>	.23 (.24) <i>2.52 (2.80)</i>		.41 (.36) <i>3.17 (3.60)</i>	.02 <i>0.35</i>	.10 <i>1.17</i>	.51 (.49)
<i>RD readers</i>								
attributions					.26 (.15) <i>2.17 (1.45)</i>	-.10 <i>-1.28</i>	.23 <i>2.69</i>	.08 (.02)
efficacy	.64 (.70) <i>8.70 (9.89)</i>				.43 (.32) <i>4.25 (3.85)</i>	-.02 <i>-0.30</i>	.17 <i>1.96</i>	.69 (.66)
metacognitive knowledge	.40 (.43) <i>4.78 (5.03)</i>	.28 (.26) <i>3.41 (3.15)</i>				-.03 <i>-0.39</i>	.20 <i>2.56</i>	.49 (.41)
performance		.64 (.60) <i>4.66 (5.07)</i>	.26 (.27) <i>2.52 (2.80)</i>		-.08 (-.03) <i>-0.63 (-0.32)</i>	-.12 <i>-1.75</i>	-.12 <i>-1.50</i>	.62 (.60)

Path weights for model without demographic variables are provided in brackets.

Associated t values are italicized; within group completely standardized solution is presented.

Table J3. Comparison of path weights and square multiple correlations for NA and RD readers in Model 2 with and without demographic variables

Dependent Variable	Independent variables								R ²
	attributions	efficacy	metacognitive knowledge	combined strategy use	perf.	ability	ethnicity	gender	
<i>NA readers</i>									
attributions					.53 (.43) <i>5.15 (4.50)</i>	.13 <i>1.87</i>	.38 <i>4.68</i>		.31 (.18)
efficacy	.54 (.62) <i>8.67 (10.09)</i>				.33 (.25) <i>4.56 (4.12)</i>	.13 <i>2.07</i>	.23 <i>3.33</i>		.64 (.58)
metacognitive knowledge	.42 (.44) <i>4.10 (4.26)</i>	.49 (.45) <i>4.08 (3.87)</i>				-0.02 <i>-0.26</i>	-0.04 <i>-0.67</i>		.70 (.69)
combined strategy use		.42 (.40) <i>3.51 (3.61)</i>	.48 (.49) <i>4.43 (4.73)</i>			.01 <i>0.24</i>	-.11 <i>-1.86</i>		.71 (.70)
performance		-.11 (-0.01) <i>-0.53 (-0.06)</i>		.65 (.62) <i>4.25 (4.42)</i>	.39 (.30) <i>3.22 (3.35)</i>	.03 <i>0.51</i>	.19 <i>2.01</i>		.65 (.62)
<i>RD readers</i>									
attributions					.26 (.15) <i>2.18 (1.40)</i>	-.10 <i>-1.34</i>	.23 <i>2.69</i>		.08 (.02)
efficacy	.62 (.70) <i>8.67 (10.09)</i>				.45 (.34) <i>4.56 (4.12)</i>	.00 <i>0.05</i>	.20 <i>2.30</i>		.70 (.67)
metacognitive knowledge	.32 (.34) <i>4.10 (4.26)</i>	.32 (.30) <i>4.08 (3.87)</i>				-0.05 <i>-0.56</i>	.20 <i>2.52</i>		.44 (.36)
combined strategy use		.29 (.29) <i>3.51 (3.61)</i>	.51 (.52) <i>4.43 (4.73)</i>			-0.07 <i>-0.98</i>	.04 <i>0.50</i>		.53 (.52)
performance		.64 (.58) <i>5.23 (5.50)</i>		.33 (.33) <i>3.72 (3.86)</i>	-.14 (-0.05) <i>-1.14 (-0.60)</i>	-.10 <i>-1.65</i>	-.15 <i>-1.94</i>		.66 (.64)

Path weights for model without demographic variables are provided in brackets.
Associated *t* values are italicized; within group completely standardized solution is presented.

Table J4. Comparison of path weights and square multiple correlations for NA and RD readers in Model 3 with and without demographic variables

Dependent variable	Independent variables									R ²
	attributions	efficacy	metacognitive knowledge	metacognitive strategy use	cognitive strategy use	performance	ability	ethnicity	gender	
<i>NA readers</i>										
attributions							.55 (.44) 5.30 (4.64)	.13 1.77	.38 4.74	.32 (.19)
efficacy	.52 (.60) 8.41 (9.76)						.32 (.25) 4.48 (4.09)	.13 2.14	.23 3.29	.61 (.55)
metacognitive knowledge	.46 (.48) 4.81 (4.97)	.45 (.41) 4.04 (3.82)						-.01 -0.14	-.04 -0.62	.69 (.68)
metacognitive strategy use		.43 (.40) 3.70 (3.68)	.45 (.46) 4.19 (4.50)					-.03 -0.47	-.08 -1.21	.66 (.66)
cognitive strategy use		.35 (.34) 2.85 (3.00)	.43 (.43) 3.89 (4.02)					-.05 0.78	-.15 -2.22	.55 (.52)
performance		-.05 (.05) -0.31 (0.35)		.24 (.22) 2.47 (2.39)	.43 (.42) 3.99 (4.12)		.37 (.29) 3.23 (3.31)	.03 0.46	.18 2.10	.66 (.63)
<i>RD readers</i>										
attributions							.26 (.15) 2.17 (1.37)	-.11 -1.35	.23 2.65	.08 (.02)
efficacy	.62 (.69) 8.41 (9.76)						.45 (.33) 4.48 (4.09)	.01 0.10	.19 2.20	.69 (.66)
metacognitive knowledge	.35 (.37) 4.81 (4.97)	.29 (.27) 4.04 (3.82)						-.05 -0.65	.20 2.54	.43 (.36)
metacognitive strategy use		.34 (.33) 3.70 (3.68)	.53 (.56) 4.19 (4.50)					-.02 -0.31	.09 1.17	.65 (.62)
cognitive strategy use		.22 (.22) 2.85 (3.00)	.42 (.41) 3.89 (4.02)					-.11 -1.39	-.04 -0.53	.33 (.32)
performance		.60 (.55) 4.85 (5.14)		.26 (.23) 2.47 (2.39)	.12 (.15) 1.28 (1.63)		-.11 (-.04) -0.97 (-.45)	-.11 -1.77	-.15 -1.86	.65 (.63)

Path weights for model without demographic variables are provided in brackets. Associated *t* values are italicized; within group completely standardized solution is presented.

Table J5. Comparison of path weights and square multiple correlations for NA and RD readers in Model 4 with and without demographic variables

Dependent variable	Independent variables										R ²
	success attributions	failure attributions	efficacy	metacognitive knowledge	metacognitive strategy use	cognitive strategy use	performance	ability	ethnicity	gender	
<i>NA readers</i>											
success attributions								.56 (.43) <i>5.42 (4.53)</i>	.12 <i>1.69</i>	.39 <i>4.92</i>	.33 (.19)
failure attributions								.51 (.44) <i>4.60 (4.36)</i>	.12 <i>1.48</i>	.24 <i>2.70</i>	.26 (.20)
efficacy	.51 (.64) <i>3.25 (4.07)</i>	-.04 (-.06) <i>-0.23 (-0.36)</i>						.36 (.27) <i>4.59 (4.15)</i>	.15 <i>2.24</i>	.24 <i>3.18</i>	.61 (.56)
metacognitive knowledge		.38 (.42) <i>5.53 (5.95)</i>	.54 (.51) <i>5.12 (4.95)</i>						.01 <i>0.10</i>	.00 <i>0.02</i>	.68 (.69)
metacognitive strategy use			.43 (.39) <i>3.48 (3.38)</i>	.43 (.45) <i>4.36 (4.81)</i>					-.03 <i>-0.46</i>	-.08 <i>-1.08</i>	.63 (.63)
cognitive strategy use			.34 (.32) <i>2.68 (2.72)</i>	.42 (.42) <i>3.98 (4.23)</i>					.05 <i>0.77</i>	-.14 <i>-2.12</i>	.52 (.49)
performance			-.06 (.06) <i>-0.36 (0.48)</i>		.23 (.21) <i>2.41 (2.29)</i>	.42 (.41) <i>4.00 (4.11)</i>		.39 (.29) <i>3.24 (3.34)</i>	.03 <i>0.47</i>	.19 <i>2.12</i>	.65 (.62)
<i>RD readers</i>											
success attributions								.30 (.21) <i>2.27 (1.82)</i>	-.11 <i>-1.25</i>	.21 <i>2.18</i>	.09 (.04)
failure attributions								.21 (.11) <i>1.71 (1.05)</i>	-.11 <i>-1.33</i>	.23 <i>2.62</i>	.07 (.01)
efficacy	.09 (.11) <i>0.43 (0.53)</i>	.57 (.60) <i>2.79 (2.92)</i>						.47 (.36) <i>4.59 (4.15)</i>	.00 <i>0.04</i>	.20 <i>2.25</i>	.73 (.68)
metacognitive knowledge		.47 (.50) <i>5.53 (5.95)</i>	.31 (.28) <i>5.12 (4.95)</i>						-.03 <i>-0.43</i>	.14 <i>2.01</i>	.57 (.54)
metacognitive strategy use			.31 (.28) <i>3.48 (3.38)</i>	.56 (.60) <i>4.36 (4.81)</i>					-.02 <i>-0.35</i>	.09 <i>1.20</i>	.67 (.66)
cognitive strategy use			.20 (.19) <i>2.68 (2.72)</i>	.44 (.46) <i>3.98 (4.23)</i>					-.11 <i>-1.46</i>	-.04 <i>-0.45</i>	.36 (.36)
performance			.60 (.54) <i>4.62 (4.79)</i>		.26 (.23) <i>2.41 (2.29)</i>	.12 (.15) <i>1.26 (1.59)</i>		-.14 (-.06) <i>-1.17 (-0.65)</i>	-.10 <i>-1.62</i>	-.16 <i>-2.02</i>	.66 (.64)

Path weights for model without demographic variables are provided in brackets. Associated *t* values are italicized; within group completely standardized solution is presented.

As comparisons are made between the parameters within groups, the within group completely standardized solution is presented. Whilst these analyses revealed that some parameter estimates (notably the associations between ability and the other latent factors) varied systematically with the inclusion of ethnicity and gender, all models yielded similar patterns of significance of the paths among constructs for each group to that in the appropriate comparison model. These analyses suggested that the introduction of ethnicity and gender did not appreciably change the model estimates and the hypothesized effects. However, there was a slight increase in the parameter estimates of the following paths: ability to attributions, and ability to performance.

One important effect that is worthy of note is the relationship between ability and attributions for RD readers. When ethnicity and gender were included, ability significantly predicted success attributions for RD readers. Nevertheless, it is not of theoretical importance because the magnitude of this path still varied markedly across groups and remains an important distinguishing factor between NA and RD readers.

In general, ethnicity and gender had consistent effects across models. For NA readers, ethnicity significantly predicted efficacy, whilst gender significantly predicted attributions, efficacy, cognitive strategy use, and performance. For RD readers, ethnicity did not have a significant effect on any of the variables, but gender significantly predicted attributions (both success and failure), efficacy, and metacognitive knowledge. The inclusion of the demographic variables did not substantially improve the determinability of the model in explaining differences in reading comprehension performance between NA and RD readers.

J.3 Discussion and conclusion

Inferences about causality are often difficult to make in nonexperimental research because of many uncontrolled demographic/background sources of variance. However, two key extraneous variables in this study, students' ethnicity and gender, were controlled.

A secondary aim of the present study was the examination of the impact of ethnicity and gender on the basic structure of the hypothesized model. The introduction of ethnicity and gender did not markedly change the structure of the hypothesized model. Although there were small changes in the values of the parameter estimates, the hypothesized relationships were fairly robust. One finding that may be of interest is that the relationship between ability and success attributions was rendered significant with the inclusion of demographic variables. Even though this relationship is significant, this path still varies significantly across NA and RD groups.

More significant gender effects were found, in contrast to the small number of ethnicity effects. The results indicate that for both NA and RD readers, girls have more positive motivational beliefs in the domain of reading than boys. These findings are consistent with previous research on gender effects for NA (Hay, 1994; Marsh et al., 1985) and RD readers (Chapman & Wilkinson, 1988).

The lower scores for motivational beliefs in boys may be partly due to the fact that boys in general have fewer verbal skills than girls. The finding is also in line with Burns' (1982) suggestion that general gender differences may be related to the higher frequency with which

boys have problems during their primary schooling when compared to girls. Additionally, Burns suggested that the boys tend to view the classroom environment as being "female". This perception is reinforced by the predominance of women teachers at the primary and intermediate levels. Consequently, boys who have to cope with failures in a "female" environment may have greater difficulty in maintaining positive motivational beliefs in reading.

In conclusion, the above findings yielded some interesting results; however, they were considered exploratory in nature. Overall, the models stayed largely unchanged, with a few exceptions. Given these factors have not been investigated extensively, they may merit further investigation in future research. Their utility may be in specifying different interventions for boys and girls or possibly modifying the teaching environment. Consequently, it may be useful to check the effects of gender and ethnicity routinely in future studies and to look for any replication of the patterns obtained in the present study.