

The Effects of a 15-minute
Direct Instruction Intervention in the
Regular Mathematics Class on
Students' Mathematical Self-efficacy and
Achievement

Rhonda Maree Farkota

DipT (Melbourne State College)
BEd (Melbourne State College)
MEd Studies (Monash University)

Thesis submitted as part of the
requirements for the degree of
Doctor of Education

Faculty of Education
Monash University

September, 2003

Table of contents

| | |
|---|-------------|
| LIST OF FIGURES | v |
| LIST OF TABLES..... | vi |
| LIST OF APPENDICES | ix |
| ABSTRACT | x |
| DECLARATION | xii |
| ACKNOWLEDGEMENTS | xiii |
| | |
| CHAPTER 1: MATHEMATICS – THE CURRENT SITUATION..... | 1 |
| Introduction | 1 |
| Possible causes of student failure | 4 |
| Self-efficacy | 6 |
| Transition | 8 |
| Transition and mathematics..... | 9 |
| Mathematics, transition and at risk students | 12 |
| Summary | 14 |
| Purpose and significance of the study | 16 |
| Research questions | 17 |
| Structure of the thesis | 18 |
| | |
| CHAPTER 2: REVIEW OF RELATED LITERATURE | 22 |
| Overview of the chapter | 22 |
| Constructivist approaches to teaching: The current trend..... | 22 |
| Constructivist teaching in practice | 26 |
| Potential problems with constructivist approaches | 28 |
| Research on constructivist approaches..... | 33 |
| The progressive education model..... | 33 |
| The open education model..... | 35 |
| Direct Instruction..... | 37 |
| Direct Instruction in practice | 38 |
| Potential problems with Direct Instruction | 41 |
| Direct Instruction research | 43 |
| Merging views..... | 49 |
| | |
| CHAPTER 3: SELF-EFFICACY | 54 |
| Overview of the chapter | 54 |
| Self-efficacy | 54 |
| Self-efficacy and self-concept | 56 |
| Sources of self-efficacy | 58 |

| | |
|---|------------|
| Self-efficacy and engagement | 61 |
| Self-efficacy and motivation | 62 |
| Self-efficacy and self-regulation | 63 |
| Self-efficacy and modeling..... | 64 |
| Self-efficacy and mathematics | 66 |
| Conclusions | 71 |
| CHAPTER 4: METHODS AND INSTRUMENTS..... | 74 |
| Introduction | 74 |
| A personal perspective on researching mathematics..... | 74 |
| Method of study..... | 76 |
| Procedures | 77 |
| Like school groups | 78 |
| Socioeconomic status | 79 |
| Sampling..... | 82 |
| Data collection..... | 85 |
| Instruments | 86 |
| Student questionnaires..... | 87 |
| General attitude | 88 |
| Perceived control..... | 88 |
| Task specific confidence | 89 |
| Reaction to challenge | 89 |
| Engagement | 91 |
| Self-efficacy scale | 91 |
| Ratings in mathematics | 92 |
| Short answer items | 92 |
| Mathematics assessment..... | 94 |
| Experimental intervention | 95 |
| Teacher survey..... | 101 |
| Summary | 101 |
| CHAPTER 5: VALIDATION OF INSTRUMENTS | 102 |
| Introduction | 102 |
| Data screening | 103 |
| Missing data | 103 |
| Rasch model | 105 |
| Validation of student questionnaires | 106 |
| Perceived control subscale | 107 |
| Engagement subscale..... | 109 |
| Reaction to challenge subscale..... | 111 |
| Task specific confidence subscale..... | 113 |
| General attitude subscale..... | 116 |
| Self-efficacy scale | 119 |
| Validation of mathematics assessment..... | 121 |
| Summary | 125 |
| CHAPTER 6: SELF-EFFICACY CROSS-GROUP COMPARISONS | 126 |

| | |
|--|-----|
| Introduction | 126 |
| Research questions | 126 |
| Statistical methods..... | 127 |
| Analysis of questionnaires: Aims..... | 127 |
| Comparison of control and experimental students' on self-efficacy scale | 128 |
| Comparison of control and experimental students' on perceived control subscale..... | 132 |
| Comparison of control and experimental students' on engagement subscale | 134 |
| Comparison of control and experimental students' on reaction to challenge subscale | 136 |
| Comparison of control and experimental students' on task specific confidence subscale | 138 |
| Comparison of control and experimental students' on general attitude subscale..... | 141 |
| Comparison of control and experimental students' by gender on self-efficacy scale | 144 |
| Female v's male (control group) | 144 |
| Female v's male (experimental group)..... | 147 |
| Female control + female experimental v's male control + male experimental | 150 |
| Summary | 153 |
| Do students' self-efficacy beliefs in mathematics change as they move through the first year of secondary school? | 155 |
| Does students' perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics change as they move through the first year of secondary school? | 155 |
| Do the changes in the self-efficacy beliefs of the students in the control group differ from those in the experimental group? | 158 |
| Do the changes that occur in the self-efficacy beliefs of the female students differ from those of the male students? | 158 |
| Conclusion..... | 160 |

CHAPTER 7: MEASURE OF MATHEMATICS ACHIEVEMENT

| | |
|--|------------|
| CROSS-GROUP COMPARISONS..... | 161 |
| Introduction | 161 |
| Statistical methods..... | 162 |
| Analysis of mathematics achievement: Aims | 162 |
| Comparison of control and experimental students' mathematics achievement..... | 165 |
| Comparison of control and experimental students' mathematics achievement: Like-school group 9 | 169 |
| Comparison of control and experimental students' mathematics achievement: Like-school group 6 | 172 |
| Comparison of control and experimental students' mathematics achievement: Like school group 2 and group 5 | 176 |

| | |
|---|------------|
| Comparison of control and experimental students' on mathematics achievement: Like school group 6 and group 9 | 180 |
| Comparison of control and experimental students' mathematics achievement by gender..... | 183 |
| Comparisons of mathematics achievement: Male v's female | 183 |
| Like school group by gender | 186 |
| Summary | 189 |
| Conclusion..... | 196 |
| CHAPTER 8: RELATIONSHIP BETWEEN SELF-EFFICACY AND MATHEMATICS ACHIEVEMENT | 198 |
| Introduction | 198 |
| Research questions | 198 |
| Statistical methods..... | 199 |
| Relationships analysis: Aims | 200 |
| Developing a mathematics scale | 200 |
| Relationship: Self-efficacy and mathematics achievement..... | 202 |
| Relationship: Task specific confidence and mathematics achievement..... | 203 |
| Relationship: Self-efficacy, task specific confidence, and mathematics achievement for at risk students | 205 |
| Summary and discussion of results | 206 |
| Conclusion..... | 207 |
| CHAPTER 9: SELF-RATINGS IN MATHEMATICS AND PERCEPTIONS OF THE MATHEMATICS LESSON | 209 |
| Stage 1 Ratings in mathematics..... | 209 |
| Stage 1 short answer items | 212 |
| Stage 2 | 216 |
| Stage 2 Ratings in mathematics..... | 217 |
| The mathematics lesson..... | 219 |
| Summary | 228 |
| CHAPTER 10: SUMMARY AND DISCUSSION..... | 233 |
| Summary | 233 |
| Discussion | 265 |
| Implications for teaching..... | 277 |
| Implications for research..... | 284 |
| Concluding words..... | 287 |
| APPENDICES..... | 291 |
| REFERENCES | 313 |

List of figures

| | | |
|-----------|--|-----|
| Figure 1 | Math percentile scores across Follow Through models | 44 |
| Figure 2 | Like school group boundaries | 79 |
| Figure 3 | Intervention teacher presentation script lesson 34, number pattern strand, equations and inverse operations strand..... | 97 |
| Figure 4 | Intervention teacher presentation script lessons 21-23 and 29, number pattern strand | 98 |
| Figure 5 | Intervention teacher presentation script lesson 24 and lesson 146, problem solving strand | 99 |
| Figure 6 | Item thresholds for perceived control subscale | 108 |
| Figure 7 | Item fit for perceived control subscale | 109 |
| Figure 8 | Item thresholds for engagement subscale..... | 110 |
| Figure 9 | Item fit for engagement subscale..... | 111 |
| Figure 10 | Item thresholds for reaction to challenge subscale..... | 112 |
| Figure 11 | Item fit for reaction to challenge subscale..... | 113 |
| Figure 12 | Item thresholds for task specific confidence subscale..... | 114 |
| Figure 13 | Item fit for task specific confidence subscale..... | 116 |
| Figure 14 | Item thresholds for general attitude subscale | 117 |
| Figure 15 | Item fit for general attitude subscale | 118 |
| Figure 16 | Item thresholds for self-efficacy scale..... | 120 |
| Figure 17 | Item thresholds for PATMaths revised..... | 123 |
| Figure 18 | Item fit for PATMaths revised | 124 |
| Figure 19 | Like school group LOTE statistics | 193 |
| Figure 20 | Significant outcome comparisons of basic skills achievement across Follow Through models | 267 |
| Figure 21 | Significant outcome comparisons of self-esteem across Follow Through models | 272 |

List of tables

| | | |
|----------|--|-----|
| Table 1 | School ID, like school group, metropolitan region and number of students..... | 83 |
| Table 2 | Stage 1 male and female student numbers | 85 |
| Table 3 | Stage 1 short answer items | 93 |
| Table 4 | Stage 2 short answer items | 93 |
| Table 5 | Content structure PATMaths revised tests | 95 |
| Table 6 | Descriptive statistics for stage 1 and stage 2, self-efficacy scale | 129 |
| Table 7 | Descriptive statistics for stage 1 and stage 2, perceived control subscale..... | 132 |
| Table 8 | Descriptive statistics for stage 1 and stage 2, engagement subscale | 135 |
| Table 9 | Descriptive statistics for stage 1 and stage 2, reaction to challenge subscale | 137 |
| Table 10 | Descriptive statistics for stage 1 and stage 2, task specific confidence subscale | 139 |
| Table 11 | Descriptive statistics for stage 1 and stage 2, general attitude subscale..... | 141 |
| Table 12 | Descriptive statistics for stage 1 and stage 2, self-efficacy scale and subscales | 143 |
| Table 13 | Descriptive statistics for scale and subscale by gender, control group | 145 |
| Table 14 | Descriptive statistics for scale and subscale by gender, experimental group..... | 148 |
| Table 15 | Descriptive statistics for scale and subscale by gender, combined groups | 151 |
| Table 16 | Descriptive statistics for stage 1 and stage 2, mathematics achievement measure | 166 |
| Table 17 | Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 9..... | 170 |

| | | |
|----------|---|-----|
| Table 18 | Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 6..... | 173 |
| Table 19 | Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 2 (experimental) and group 5 (control)..... | 177 |
| Table 20 | Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 6 and group 9..... | 180 |
| Table 21 | Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender | 184 |
| Table 22 | Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender, control group..... | 184 |
| Table 23 | Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender, experimental group | 185 |
| Table 24 | Descriptive statistics for mathematics achievement measure by gender, like school groups..... | 186 |
| Table 25 | Summary statistics for stage 1 and stage 2, mathematics achievement measure | 192 |
| Table 26 | Summary statistics for like school group LOTE, EMA/Youth allowance, community employment..... | 194 |
| Table 27 | Percentage of students in each stanine score category | 201 |
| Table 28 | Descriptive statistics control group by category, self-efficacy variable | 202 |
| Table 29 | Descriptive statistics experimental group by category, self-efficacy variable | 203 |
| Table 30 | Descriptive statistics control group by category, task specific confidence variable..... | 204 |
| Table 31 | Descriptive statistics experimental group by category, task specific confidence variable..... | 205 |
| Table 32 | Descriptive statistics for students' self-rating, stage 1 current study and stage 2 Fullarton study | 210 |
| Table 33 | Descriptive statistics for students' self-rating, stage 2 current study and stage 3 Fullarton study | 218 |
| Table 34 | Frequency distribution % of the students' liking of mathematics stage 2 | 221 |

| | | |
|----------|--|-----|
| Table 35 | Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you like the most?..... | 222 |
| Table 36 | Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you think you have learned the most? | 223 |
| Table 37 | Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you think you have learned the least? | 226 |
| Table 38 | Frequency distribution % of the students' stage 2 responses to the item: Compared to primary school how do you think you are doing in maths?..... | 227 |

List of appendices

- Appendix 1 Approval: Standing Committee on Ethics in Research Involving Humans
- Appendix 2 Response to the Standing Committee
- Appendix 3 Response to Standing Committee Approval, 10 December
- Appendix 4 Permission from the Department of Education, Employment and Training (DEET) to conduct research in Government schools
- Appendix 5 Explanatory statements and permission slips
- Appendix 6 Optical mark reader answer sheet
- Appendix 7 Teacher survey
- Appendix 8 Stage 1 student questionnaire
- Appendix 9 Stage 2 student questionnaire
- Appendix 10 Intervention sample lesson

Abstract

While the Direct Instruction experimental intervention employed in this thesis was based on the Engelmann model it differs in fundamental aspects. Specifically designed to cater for the diverse academic levels present in any given classroom it aimed to accommodate and elevate every student's academic skill level. Satisfactory academic performance is composed of a balance of, on the one hand possessed skills and on the other a certain belief in self. Given that self-efficacy is well accepted as an accurate predictor of academic performance the study also examined the effects of the intervention on students' self-efficacy from the perspective of Bandura's (1986) social cognitive theory. Since the decline in student self-belief over the transition years has long been recognised as a problem, the study was conducted in 54 regular Year 7 mathematics classrooms comprising 967 students. With at risk students forming a subsidiary focus, socioeconomic status was a relevant consideration in the selection of schools.

Employing a pretest-posttest control group design, the experimental intervention was implemented in the first 15 minutes of the regular mathematics lesson. The data were collected immediately prior to administering the experimental intervention and shortly after the trial period concluded. Pretesting both groups enabled the scores to be used as a statistical control, to analyse gain scores.

Analysis of the questionnaire data showed no significant change in the self-efficacy beliefs of the control group, whereas in contrast, there was a significant gain in the self-efficacy beliefs of the experimental group.

Analysis of the mathematics assessment data showed a significant growth in mathematics achievement for both the control and experimental groups. Pretest comparison showed that the difference in the means for the two groups was statistically significant in favour of the control group, whereas the posttest difference was trivial and not statistically significant. This revealed a particularly significant achievement overall in favour of the experimental group. Significantly, the biggest growth in mathematical achievement in the entire study appeared in the experimental group regarded most at risk. Overall it was found the behaviourally based Direct Instruction intervention had a positive effect on mathematical achievement and self-efficacy.

The findings in this study add to the existing body of evidence attesting to the effectiveness of competently designed, properly implemented teacher-directed programs in the important though often neglected mathematical domain of laying the foundational skills. It is to be hoped they also contribute to the long-standing debate between teacher-directed and constructivist student-directed learning adherents by illustrating that some skills are better acquired through one approach and some through the other. The research and literature reviewed herein shows that in the employment and cultivation of higher order skills where reasoning and reflection are required, a constructivist approach would seem more appropriate. But when it comes to the acquisition of basic mathematical skills the findings in this study clearly show a competently designed, properly implemented teacher-directed approach is ideally suited.

Declaration

I affirm to the best of my knowledge that this thesis does not incorporate any material previously submitted for a degree or diploma in any university or other institution and does not contain any material previously published or written by another person except where due reference is made in the text of the thesis.

Rhonda Maree Farkota

Acknowledgements

My sincere thanks are due to the many people who have assisted this study.

I would like to thank my supervisor Dr Barbara Clarke for her conscientious guidance, her critical reading of the drafts, and continual encouragement and support. I would also like to thank Associate Professor David Harvey for the initial inspiration, and subsequent comments and criticism.

On the statistical aspects I am indebted to Dr John Barnard for his ongoing advice and suggestions for the presentation of results.

I wish to acknowledge my colleagues at the Australian Council for Educational Research for their generous help throughout the study and a special thanks to Gloria Locock, and to the library staff for their endless pursuit of my requests.

I would like to express my gratitude to the various Principals, mathematics coordinators, teaching staffs, and students who participated in the study, for their cooperation and assistance.

Finally, I am deeply indebted to my best friend and mentor, my husband, for his encouragement and unconditional support, without him by my side nothing would have been possible.

Chapter 1: Mathematics – The current situation

Introduction

Battista (1999) informs us 'mathematics anxiety is widespread' (p. 426), and while Townsend, Moore, Tuck and Wilton (1998) have described angst about mathematics as 'feelings of tension and *anxiety*' arising out of dealing with mathematics in 'a wide variety of ordinary and academic situations' (p. 41), it has also been characterized as apprehension and fear and sometimes dislike (McLeod, 1992).

More than twenty years ago a Governmental Committee of Enquiry described the widespread perception of mathematics amongst the adult population of the UK as a 'daunting subject' (Cockcroft, 1982, p. 6, para. 16) and the situation in that country is no different today. Lamenting the problem-solving skills in the US workplace Gordon (1997) claims, 'half of American adults are close to functionally illiterate' (p. 14). And Cossey (1999) is scathing, describing the understanding of mathematics in the US as 'a national joke' (p. 443). In our own part of the world the research perspective presented in *A National Statement on Mathematics for Australian Schools* reveals that a large body of 'anecdotal and research evidence' exists showing many people have a dislike and even a fear of occasions which expose them to mathematics (Australian Education Council and Curriculum Corporation, 1991, p. 7). In his monograph analysing standards in Australia, Ayres (2000) concludes there is a widely held perception amongst people generally that fundamental understanding

in mathematics is not what it used to be. After analysing the performance of Australian students in *The Third International Mathematics and Science Study* (TIMSS), Stacey (1997) advises 'we have certainly succeeded in NOT emphasising arithmetic computation the percentages of students correct are near or below the international average' (p. 45). National data from the US indicates similar findings and there are like concerns in Europe (Cumming, 2000).

It does seem beyond argument that today's students possess an embarrassing paucity of skills when it comes to mathematics, and every year literally thousands upon thousands of them are leaving school as young adults and taking up employment without being able to communicate mathematically. This worrisome truth presents alarming sociological implications for the nature, efficacy and future direction of the country's workforce. In the US, Gordon (1997) informs us that large numbers of business people see the poorly educated workforce as 'the principal threat to their future success' (p. 14). In Gordon's opinion the problem lies in the fact that the expectations of the US school system over the larger part of last century were low — it was expected to furnish the nation with a 'grunt labour force' suitable for 'assembly-line jobs' and little else (p. 14). Most students in public education were not expected to rise above fourth-to-sixth grade levels in Mathematics and English. Dietz (1998) looks at new studies in US education which examine whether the sort of preparation students are receiving will enable them to prosper 'in an increasingly technology-based workforce' and goes on to speculate that if the conclusions he draws from these studies are correct they indicate 'a breakdown in math and science education' (p. 40). Indeed, just such a breakdown has been highlighted in a recent UNICEF (2002) report

analysing the latest cross-national surveys into educational performance, which showed nearly 40 per cent of US grade 8 students were incapable of applying basic mathematical knowledge to 'straightforward situations' (Zammit, Routitsky, & Greenwood, 2002, p. 20). Further, it showed that 27 per cent of Australian grade 8 students were in a similar condition and countries such as the UK, New Zealand, Italy, Spain and Greece were in an even worse condition than the US all showing percentages of over 40.

Cassidy (2000) reports that in the UK the government at ministerial level is concerned that Britain may very well end up suffering financial damage due to the poor mathematical abilities of the labour force. Clearly schools have a crucial role to play here. In Australia, Lamb (1997) concludes the findings in *Research Report Number 4: Longitudinal Surveys of Australian Youth* (LSAY) highlight the important role of schools in aiding students to enter the workforce, and he notes explicitly that teenagers with poor literacy and numeracy skills are the ones most likely to be unemployed both in the short and long term. Lamb's comments are supported by Marks and Fleming (1998, p. 5) who note that a government report on youth unemployment advocates 'increasing the levels of literacy and numeracy as a means of improving the employment prospects for Australian youth' (House of Representatives Standing Committee on Employment, Education and Training, 1997).

In the light of all these findings and comment it is not surprising that the teaching of mathematics in the schools of the Western world is a major problem zone. It should, however, come as a bit of a shock to find that large numbers of students entering the teaching profession actually harbour negative feelings towards mathematics (Townsend et al., 1998;

Watson, 1987). If many of our prospective teachers are feeling this way about maths we have to wonder what sort of values are being passed on to the students in our schools when they are being taught maths. In arguing for more research focus on values in mathematics teaching, Bishop (1999) refers to 'the often-quoted negative views expressed by adults about their bad mathematics learning experiences' and goes on to 'speculate that the values transmitted to them were not necessarily the most desirable' (p. 1).

Possible causes of student failure

The obvious question is whether students' failure to learn relates to a problem with the curriculum, the teaching, or the student, or perhaps some combination of these (Carnine, 1997). Though there are many possible reasons why students are failing in mathematics it would seem that most of them are related to curriculum and methods of teaching rather than the students lack of capacity to learn (Carnine, 1991; Engelmann & Carnine, 1982; Jones, Wilson, & Bhojwani, 1997). Airasian and Walsh (1997) argue the teaching of mathematics in schools has not measured up to the needs of the vast majority of our students, and that not nearly enough instructional stress is put on the higher order skills. Placing teachers and curriculum at the centre of student failure to learn, Engelmann (1980) remarked over twenty years ago that, 'the theoretical approaches they use, and their attempts to translate these into concrete, specific instruction, do not work well' (p. 28). And the situation seems to be little different today: G. Reid Lyon, who runs the US federal government's research into reading and learning disabilities was quoted in the LA Times

as saying 'Learning disabilities have become a sociological sponge to wipe up the spills of general education' (O'Reilly & Poindexter, 1999, Dec. 12, p. 1). According to Lyon perfectly normal students end up in special education for no other reason than the fact they were not taught well in the first place.

Students understandably get confused if something is not properly explained. This commonly occurs where teachers use complex terminology, or fail to express themselves clearly. The failure to use appropriate analogies, or provide meaningful examples, or overloading students with information all tend to create learning problems (Engelmann & Carnine, 1982; Jones et al., 1997). One would think the obvious way to handle learning failure would be to improve the quality of teaching (Darch, Carnine, & Gersten, 1984; Gersten, Woodard, & Darch, 1986). But what sort of teaching — *what practices makes a teacher effective and what should the teacher's role be?* Traditional teaching refers to teaching 'organised around the transfer of information from the knowledgeable teacher to the uninformed student' (Goldsmith & Shifter, 1997, p. 23). The research of Rosenshine and Stevens (1984) demonstrated that with traditional teaching students had higher academic gains when teachers adopted a consistent pattern of demonstration, guided practice, and feedback.

Battista (1999), on the other hand argues traditional teaching fails to promote students mathematical thought process 'resulting in stunted growth' (p. 430). Here the teacher acts as the sole mathematical authority 'telling things to students' (Goldsmith & Shifter, 1997, p. 27) while the students are simply the passive listeners accepting 'mathematical validity

as being established by the teacher' (Simon, 1997, p. 73). With traditional teaching students are also expected to acquire skills by observing and copying teacher behaviour. Boaler (1997) describes these students as 'submissive and rule-bound' (p. 152).

Though educators cannot agree on what aspects of curriculum or which teaching practices are to blame for the poor mathematical standard of Western society, it seems it is not the students who are at fault. In considering the increasingly high number of students who have problems in acquiring mathematics skills, Westwood (2000) concludes 'there is no convincing evidence (with a few exceptions) that their problems result from any lack of innate potential or from cognitive deficits' (p. 2).

Self-efficacy

No one doubts how important confidence is when it comes to things like sport but how does it affect something like academic performance? For quite some time now educators have been speculating on its importance under the nomenclature of *self-efficacy*. Bandura, who first introduced the construct in 1977, defines self-efficacy as 'belief in one's capability to organize and execute the courses of action required to manage prospective situations' (Bandura, 1997, p. 2). According to Bandura (1986) enactive attainments, in other words, actual experiences, 'provide the most influential source' of self-efficacy (p. 399). To relate this back to sport, if in the past Jack has beaten Jim easily every time they played tennis, then providing all things remain the same Jack will be confident of beating him in the future. Jack's confidence quite naturally stems from positive experiences. Similarly if Jill has consistently gained high marks in algebra

her confidence, or self-efficacy in algebra will be high. Again this feeling stems from positive experiences. Naturally enough the same rules apply to negative experiences. If Jack was always being beaten at tennis by Jim then his confidence in beating that friend in the future would be low. And if Jill was always getting her algebra wrong then her self-efficacy in algebra would be low. Thus it would seem in the academic world as in the world generally, *success breeds success, failure breeds failure*.

The fact that experiences of mastery are particularly influential sources of our information on self-efficacy (Bandura, 1986) has important implications for the self-enhancement model of academic achievement (Pajares, 1996). If we want to elevate student achievement we should concentrate on elevating their self-efficacy (Pajares, 1996).

Hanchon Graham (2000) captures the essence of Banduras' work (Bandura, 1986, 1997) when she comments, 'by providing numerous opportunities for small victories, teachers increase the chance that all students experience the performance attainments' (p. 12). In this respect teacher encouragement can also be an effective persuasion but only when students have experienced authentic mastery, as Erikson (1959) cautions 'children cannot be fooled by empty praise and condescending encouragement' (p. 95).

Hanchon Graham (2000) eloquently puts the case for making the classroom a place where 'mistakes are viewed as opportunities' and treated as 'an integral part of the learning process' (p. 12). She argues that success is success and the fact that it may take extra industry to acquire should not diminish its value. Her study *Self-efficacy, Motivation*

Constructs, and Mathematics Performance of Middle School Students supports Bandura's (1977; 1986; 1997) findings in establishing the important role self-efficacy plays in student motivation and achievement. What students think about their ability to solve mathematics tasks is generally an accurate predictor of their actual performance (Pajares & Miller, 1994). In other words they will normally have a good idea as to how well they will perform in specific academic situations.

Satisfactory academic performance of course is composed of a balance of on the one hand possessed skills and on the other a certain belief in self. These predictions by students about their ability are naturally enough based on experience and it is generally accepted students experience most of their problems in the transition years. When students enter this *transitional door*, when they move from primary to secondary school, their self-efficacy suffers. *The Middle Years Numeracy Research Project* recognised that a significant number of students in Years 5 – 9 have difficulty 'maintaining their levels of performance over the transition years' (Siemon, 2000, p. 21). Hanchon Graham (2000) calls for more research into this area and it is partly in response to the call that this author is including in her thesis an evaluation of self-efficacy and its effects on students in the transition year.

Transition

Transition has been described as 'a short term life change characterized by a sharp discontinuity with the past' (Brammer, 1992, p. 1). Although there has been remarkably little research done on it, the particular discontinuity that occurs between primary and secondary school has long

been recognised as a problem period. The situation has been aptly summed up by Clarke (1989) who notes it is commonly referred to as 'the transition problem' (p. 2). Indeed, it is generally accepted that students' grades fall off considerably in almost any school transition (Midgley, Feldauer, & Eccles, 1989).

In their analysis of curriculum differences in transition and their relation to student achievement and satisfaction, Power and Cotterell (1981) found 'Major curriculum discontinuities in mathematics' (p. 18). De Groot (2000) discussing the work of Rice (1997) sees the shifts that occur in school transition as being 'related to and found in two discontinuities', these being the 'environment' and the 'social structure' of the school (p. 4). She offers a third dimension in the transition stage, '*the learning environment*, and particularly the learning environment of mathematics' (de Groot, 2000, p. 5).

Transition and mathematics

As with research into transition generally the area of transition and mathematics has received scant attention from researchers. Fullarton (1998), in her study of student engagement during transition from primary to secondary school, makes the point that 'despite great changes to the mathematics curriculum over the last ten years, the area of transition has been largely neglected' (p. 14). Although there have been studies researching student attitudes toward mathematics (e.g., Bay, Beem, Reys, Papick, & Barnes, 1999), few have studied the subject from the students perspective of their transitions in learning mathematics (e.g., de Groot, 2000; Fullarton, 1998; Hanchon Graham, 2000).

In the transition from primary to secondary school mathematics there is a considerable shift in what is happening conceptually. Power and Cotterell (1981) state that, 'while the primary mathematics program aims to foster an intuitive grasp of basic mathematical ideas and processes, the secondary course sets out to introduce mathematics as a formal, logical system' (p. 18).

Whilst specifically discussing transition and mathematics Fullarton (1998) points to the shift that occurs in secondary schools towards a 'more formal and abstract curriculum' and notes that learning difficulties students may encounter at this stage 'can critically affect their self-confidence' (p. 4). Singling out the differences in presentation of the math programs between primary and secondary education she goes on to stress the potential detrimental impact learning problems can pose at this stage. Changes in terminology can also pose a transitional problem. Power and Cotterell (1981) found that much of the material covered in secondary maths 'turns out to be a repetition of work done in primary school but couched in unfamiliar, abstract language' (p. 18). This writer can think of no sensible reason for waiting till students enter secondary school before introducing them to specific mathematical terminology. If students are not conversant with mathematical language and cannot perform basic mathematical operations then the acquisition of the more complex maths skills required in secondary school is almost impossible (Jones et al., 1997).

Howsoever these transitional problems occur in the maths area, it seems they can be picked up early. Clarke (1989) claims the effect on students' mathematics performance can be discerned in the very year after primary school. Care should be taken in the categorisation of students and their

individual transitional problems, for students of ostensibly similar academic standard may not enjoy a similar academic experience over this period. Ellerton and Clements (1988) make the significant point that transition can affect students differently even though they are of comparable ability.

Moore's (1989) study: *The Transition Problem A Study of the Mathematics Curriculum in the Primary — Secondary Transition in West Gippsland*, involving six post-primary schools, found around half of the primary school students leaving Year 6 were not proficient at prescribed work for that year and were unfamiliar with most of the prescribed texts and resources. Moore's findings raise two important issues. Enquiry firstly must be made as to whether student lack of mastery to this degree applies across a wider spectrum in primary classrooms and secondly whether it should be categorised as purely a *transitional problem*. The blame that has been laid on the transition years just might be misplaced. In order to determine what a student's level of mastery is, that student's performance needs to be comprehensively assessed and until there is adequate assessment procedure in place within the primary system we must be careful not to dump all the blame on the transition per se.

As can be seen the area of transition and mathematics raises important issues that have for too long been neglected. This research aims to explore these issues with a subsidiary focus on *at risk* students who for the purposes of the research will be defined as 'learners who may be experiencing difficulty' (AREA, 1999, Preface).

Mathematics, transition and at risk students

It is a sad fact but at risk students provide the education system with considerably more problems than are given due consideration by the authorities. Secondary teachers in mathematics classes generally do not have the time to cater for individual differences (Meadowcroft, 2000), so students with limited maths skills make little or no progress (Jones et al., 1997). Less obviously the more able students also suffer if they are not challenged in that they tend to lose interest and their potential is not exploited (Miller, Mills, & Tangherlini, 1995). Further, the *streaming* of students according to ability is not acceptable to many teachers (Meadowcroft, 2000) because they feel the minority groups who get placed in the lower streams are being 'discriminated against' (Hallam & Toutounji, 1996, p. 4).

It has already been noted that students who cannot follow mathematical language or implement basic mathematical operations will find it nigh impossible to acquire the more complex maths skills required in secondary school (Jones et al., 1997). And this situation obviously can only worsen when dealing with at risk students. The low achievement of these students and their persistent failure will inevitably have a negative effect on their self-perception of mastery and their confidence, and have a similar effect on their progress (Jones et al., 1997; Miller & Mercer, 1997). The repeated failure of these students imbues them with a sense of helplessness and they become so accustomed to getting things wrong that they lack the incentive to try any more (Fulk, Brigham, & Lohman, 1998). Just where these students rank academically when tested alongside their peers will come as a shock to many. In examining the educational disparities

between countries and the treatment of at risk students UNICEF asked the question: 'How far behind are the weakest students allowed to fall?' (2002, p. 9). UNICEF found that in the worst affected countries, which were the US, Germany, New Zealand and Belgium, the low achievers lagged approximately five years behind their middle-achieving peers. Though Australia performed better in the study it still came an undistinguished fifteenth in the twenty-four countries included.

According to Jones, Wilson, and Bhojwani (1997) when at risk students move on to adolescence, they will already have experienced 'many years of failure and frustration' (p. 152) and will have achieved little by the time they finish school. They argue that we will only improve maths education in secondary schools by first improving maths education in *primary schools*.

Going on to demand maths education be based on empirical research, Jones and colleagues criticize the shift towards constructivist and student-directed learning as being 'appealing but unvalidated trends' that are 'logically antithetical' (1997, p. 152) to existing research. Hempenstall (1996) lends support lamenting the lack of attention given to empirical research into effective teaching practices.

If Jones and colleagues (1997) are correct in attributing our student's lack of basic mathematical skills to student-directed learning the question must be asked whether it is appropriate especially in the crucial upper primary years to rely solely on one style of learning. This proposition would seem to be even more crucial when dealing with at risk students. It hardly seems reasonable to expect at risk students to direct their own learning through the complicated multi-faceted world of mathematics and come out

understanding it all. Perhaps the learning of these skills would be more satisfactorily accomplished if a teacher-directed approach were adopted either by way of a properly scripted program or a competent mathematics teacher. Unfortunately there is a dearth of effective instructional aids available and as Jones and colleagues (1997) point out, not many teachers have the time or training to design effective educational mathematical material. A significant focus of this research will be on the effectiveness of the scripted teacher-directed intervention (the experimental intervention), which has been designed specifically to fill this void.

Summary

It would thus seem, in the Western world, at least, that the current state of mathematics teaching in schools is in an unhealthy condition and the adult population generally has the most fragile of grasps on the subject. And although we have high numbers of students with poor mathematics skills there is no compelling evidence to show this is the fault of the students. If we accept the veracity of the two preceding sentences, logic dictates we must have a long hard think about everything relating to how our students are being taught mathematics and empirical research into the situation is a matter of urgency.

Given that the importance of self-efficacy in the prediction of academic performance is undenied it would also seem we should be seeking out ways of elevating students' self-efficacy beliefs. If the old adage *success breeds success* holds true we should be furnishing our students with multiple opportunities to experience success even if only small. And if we acknowledge that successful performance is the most powerful source of

self-efficacy beliefs, then, we should also acknowledge just how much more powerful the incremental and flow-on effect of many small successes must be. While discussing the importance of success we should be careful not to frown too much or look down on failure. The ability to learn from mistakes can be a powerful tool in the ultimate attainment of success, and besides, it breeds that most admirable trait: resilience.

We have seen that while little research has been done in the transition area it has long been recognised as a major problem zone with mathematics being a major casualty. For at risk students, mathematics classes in the transition years must be especially disheartening experiences as secondary teachers in mathematics lack the time to deal with their particular problems. While there is no doubt genuine problems do arise in the transition years, it would seem that until there is adequate procedure in place within the primary system to have students comprehensively assessed we must hold off on laying all the blame on the transition per se.

To conclude, it is crucial we recognise that an Australian workforce lacking basic skills in mathematics in this increasingly hi-tech age is a worrisome possibility, and one that is perhaps even now a reality. In the face of such a scenario the prospects for this country as a whole would be grave indeed, and if the matter is not quickly and competently addressed we may well be saddled with the *grunt labour force* adverted to in the introduction.

Purpose and significance of the study

This study will focus on student learning in mathematics in the first year of secondary school to determine the effects of a daily 15-minute teacher-directed intervention. It involves gathering information about student self-efficacy and achievement in mathematics so as to measure the effect of the experimental intervention. The collected data will also be examined to determine precisely what sort of measurable relationship exists between achievement and self-efficacy.

The idea that students act on their perceived capability has important implications for classroom practice and given that self-efficacy is well accepted as an accurate predictor of academic performance it seems wise to examine teaching methods that might increase students' self-efficacy. By understanding how transition students come to estimate their ability at this critical juncture, educators will be better equipped to provide interventions aimed at improving students' self-perception so they can better utilise their talents to exploit their potential. It is hoped the measured outcomes of this study will also go some way towards striking a positive balance and help promote a better understanding of the respective circumstances in which both student-directed and teacher-directed approaches to learning are best employed. An ancillary focus of the study will centre on at risk students who for the purposes of the research will be defined as 'learners who may be experiencing difficulty' (AREA, 1999).

Research questions

The essential focus in this study will be on what effects, if any, a Direct Instruction intervention in the regular mathematics classroom has on students' mathematical self-efficacy and achievement? To comprehensively explore this question the following subsidiary questions will be addressed:

- Will the students taught with the Direct Instruction intervention develop a higher mathematical self-efficacy than the students taught without it?
- Will the at risk students taught with the Direct Instruction intervention develop a higher mathematical self-efficacy than the at risk students taught without it?
- Will the growth in knowledge and understanding in mathematics of the students taught with the Direct Instruction intervention exceed that of the students taught without it?
- Will the growth in knowledge and understanding in mathematics of the at risk students taught with the Direct Instruction intervention exceed that of the at risk students taught without it?
- What relationship, if any, will exist between students' self-efficacy and students' achievement?

Further research questions will be taken into account where appropriate.

Structure of the thesis

The following two chapters review the literature related to this thesis. Chapter 2 examines the student-directed versus teacher-directed learning debate through a review of the literature and research pertaining to the *Constructivist approaches* alongside that relating to *Direct Instruction*. The origin of these teaching practices and specifically their essential elements are described, together with their methods of implementation. Potential benefits and problems arising out of these approaches are analysed and the possibility is mooted of applying the selective use of a particular approach to a particular learning situation on the basis of a *fitness for purpose* appraisal.

Chapter 3 examines the role self-efficacy plays in the learning process and analyses the various constructs that come into play in predicting and mediating academic performance. Differences between self-efficacy beliefs and self-concept will be first explained then findings on the relationship between self-efficacy and engagement, motivation, self-regulation and modeling are summarised. The research and literature on mathematics self-efficacy is reviewed and conclusions are drawn from this and the preceding chapters.

Chapter 4 presents the researcher's personal perspective then generally describes the gathering of information about student self-efficacy and achievement in mathematics so as to measure the effect of the experimental intervention. Details concerning the procedures and the instruments used to implement the research are outlined, and descriptions of the schools, which participated in the study, are provided. Also

discussed in the chapter is the quantitative research method whereby data were collected from students participating in the study at two stages during their first year of secondary school.

Chapter 5 outlines the validation of the questionnaires and the mathematics assessment instruments. The self-efficacy scale of the student questionnaires consists of five subscales, which, along with the mathematics assessment items, are validated using Rasch (1960; 1980) measurement and the findings are recorded in this chapter.

Chapter 6 details and discusses the student self-efficacy cross-group comparisons data analysis from each stage of the study. As the questionnaires were designed to measure five interrelated subdomains of student self-efficacy over the first year of secondary school, data analyses were carried out to examine the effects on the scale and each subscale to measure any change between Stage 1 and Stage 2. Comparisons between the control and experimental group of students are presented as well as gender differences.

The following subsidiary research questions provided the focus for the chapter:

- Do students' self-efficacy beliefs in mathematics change as they move through the first year of secondary school?
- Does students' perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics change as they move through the first year of secondary school?

- Do the changes in the self-efficacy beliefs of the students in the control group differ from those in the experimental group?
- Do the changes that occur in the self-efficacy beliefs of the female students differ from those of the male students?

Chapter 7 details and discusses the student mathematics achievement cross-group comparisons data analysis from each stage of the study. As the mathematics assessment was employed primarily to determine whether there was any growth in student knowledge and understanding of mathematics, data analyses were carried out to examine any difference in performance between Stage 1 and Stage 2. Comparisons between the control and experimental group of students are presented as well as comparisons within *like school* groups. Gender differences in mathematics achievement are also investigated.

The focus for the chapter revolved around an analysis of the following matters:

- Identify the differences in mathematics achievement of the control group and the experimental group between Stage 1 and Stage 2 of the study.
- Examine the differences in mathematics achievement between the control group and the experimental group.
- Determine what effect, if any, socioeconomic factors had on mathematics achievement.

- Determine what effect, if any, gender differences had on mathematics achievement.

Chapter 8 investigates the relationship between student self-efficacy beliefs and mathematics achievement. More explicitly it explores the actual effects students' self-efficacy has had on their mathematics achievement over time. Data analyses were carried out to explore this relationship, and comparisons between the control and the experimental groups of students are examined as well as comparisons within each group. Low achieving student differences are also investigated.

Chapter 9 examines student responses to the self-ratings in mathematics and short answer items with a view to gaining further insight into their attitude towards learning with a particular focus on the mathematics lesson.

The major findings, conclusions and recommendations are presented in Chapter 10.

Chapter 2: Review of related literature

Overview of the chapter

This chapter examines the student-directed versus teacher-directed learning debate through a review of the literature and research pertaining to the constructivist approaches alongside that relating to Direct Instruction. The origin of these teaching practices and specifically their essential elements are described, together with their methods of implementation. Potential benefits and problems arising out of these approaches are analysed and the possibility is mooted of applying the selective use of a particular approach to a particular learning situation on the basis of a *fitness for purpose* appraisal.

Constructivist approaches to teaching: The current trend

The popular trend in education at the moment is called Constructivism. It is, according to Fosnot (1996) 'the most current psychology of learning' (p. 8). Student-directed learning as opposed to teacher-directed learning is the underlying philosophy behind constructivist approaches to teaching, and though within the constructivist camp there are divisions about how the child is best directed it is outside the scope of this thesis to deal with them individually. Fetherston (1997) argues that the term Constructivism is often used loosely, is nowhere clearly defined and appears to emanate from no established base of knowledge. Proponents of constructivist teaching

approaches (e.g., von Glasersfeld, 1995) believe that children learn best when they participate in activities relevant to them. These activities must be capable of maintaining student attention and also require that students make meaning of them for themselves (Simon, 1997).

Battista (1999) sums up the perspective thus: 'mathematical ideas must be personally constructed by students as they try to make sense of situations' (p. 71). Carpenter et al. (1999) further make the point that by 'constructing their own procedures for solving problems, students take responsibility for their own learning' (p. 59). According to Airasian and Walsh (1997) 'Constructivism is an epistemology, a philosophical explanation about the nature of knowledge' (p. 444). Instead of getting students to come up with the right answer von Glasersfeld (1995) declares it is important to get students to articulate how they arrived at their answer. Airasian and Walsh (1997) suggest that the way to achieve the problem-solving and critical thinking skills is by re-orienting instruction to non-rote outcomes where such skills as 'generalizing, analyzing, synthesizing, and evaluating are very important' (p. 446). It is argued this approach is more rooted in reality.

The basic assumption of a student-directed learning practice such as Constructivism is that a child learns best from instruction that is self-initiated and directed. Constructivist approaches to teaching assume that children's learning needs are best served by allowing them to pursue their individual interests. Each child's learning needs, which are seen to be unique, are revealed through their particular interests. Further, children should be encouraged to pursue their own particular individual interests (Grossen, 1993). According to the National Association for the Education of Young Children (NAEYC) 'Much of young children's learning takes place

when they direct their own play activities Such learning should not be inhibited by adult-established concepts of completion, achievement, and failure' (1987, p. 3).

In her comprehensive review of student-directed teaching methods Grossen (1993) reminds us that Constructivism is not a new concept, in fact student-directed learning has been around for over two thousand years (Matthews, 1992). Plato (1955) was of the opinion that knowledge acquired under compulsion 'never sticks in the mind' (para. 536), and with respect to early education 'let your children's lessons take the form of play' (para. 537). The notion that real learning could not occur under the control and direction of a teacher is commonly attributed to Piaget (1952) even though he never conducted research on student-directed learning (Grossen, 1993). Piaget (1970) did, however, state that: 'Each time one prematurely teaches a child something he could have discovered for himself, that child is kept from inventing it and consequently from understanding it completely' (p. 725). It is, of course, greatly debatable how much children are actually capable of discovering for themselves, and then the question must be asked: How long should it take them to discover? Further, how much truth is there in the proposition that a child can never understand completely that which he or she has not discovered for him or herself? Adams and Engelman (1996) make the point 'research on learning shows that students who initially mislearn require substantially more practice to relearn the concepts than they would have if they had learned it correctly' (p. 14).

Throughout the US and Canada student-directed learning practices are reappearing except now they are called *Developmentally appropriate*

practices and *Constructivism*. These practices have been defined and guidelines for their implementation have been developed by the NAEYC (1987).

In the review of literature on constructivist approaches to teaching that follows the researcher has gained valuable insight from Westwood's (2000) guide to teaching mathematics with its comprehensive reference list and straightforward discourse on Constructivism. Here Westwood states unequivocally his book is not 'about identifying students with learning difficulties' but about 'high-quality "first teaching" to prevent students failing in the initial acquisition of numeracy skills' (Preface). According to Westwood (1999), Australian teachers are being enthusiastically encouraged to adopt constructivist approaches in most curriculum areas, especially maths, language and science, and virtually all the teacher education programs in Australian universities are based on constructivist approaches. He quotes a small study in South Australia where 67 per cent of the teachers reported that in their methodology courses the only teaching approach they had been introduced to was Constructivism. Later he makes the point that most of today's teachers believe teacher-directed learning is outdated and that activity-based learning is the only way to go (Westwood, 2000). According to constructivists, teacher-directed learning methods are boring and repetitive and inappropriate for today's students (McCarthy & Schwandt, 2000). It is difficult to argue against the boring and repetitive aspects of teacher-directed learning methods, but as to whether or not they are entirely inappropriate is another matter and one which will be addressed later in this thesis.

Constructivist teaching in practice

Generally speaking constructivists see the teacher's role as mainly one of aiding student performance in the construction of knowledge, rather than providing explicit knowledge (von Glasterfeld, 1996). As far back as Plato (see page 24) there have been serious thinkers who believe it is neither necessary nor desirable to teach explicitly and that direct explanation may even be harmful. In eighteenth century Europe, Rousseau who believed mankind was benevolent by nature but could be, and in many cases was, corrupted by social interference, espoused constructivist beliefs. In his view children should be allowed to develop naturally, untrammelled by society's dictates. His exhortation to teachers was 'Give your pupil no lesson in words, he must learn from his experience' (Rousseau 1964 cited in Weir, 1990, p. 28). Commenting on current constructivist thinking Harris and Graham (1996) report that to some '*teaching* is a dirty word' (p. 27).

Confrey (1990) informs us that the constructivist classroom must be conducive to student needs giving them the time and space necessary to articulate their various points of view. McCarty and Schwandt (2000) describe the constructivist teacher as having two specific tasks, 'to establish a learning environment suited to providing perturbations for the student's mental constructive processes and to project a model of each individual student's stage in mental development and constructions' (p. 50).

According to Airasian and Walsh (1997), the constructivist classroom is one of activity, involvement and creativity conducive to the building of personal knowledge and understanding. Group discussion, which is seen as an important aspect of constructivist teaching, is largely implemented in

the form of cooperative problem solving, and can take place with or without an instructor (Lochhead, 1991). Fosnot (1996) views the constructivist classroom as 'a community of discourse engaged in activity, reflection, and conversation' (p. 29). Inagaki, Hatano and Morita (1998) are enthusiastic in supporting the principle of discourse between students. They see the fostering of student interaction in the classroom as an ideal way for students to construct mathematical knowledge, but they still see the teacher's intervention as essential.

Some educators, however, (Harris & Graham, 1996; Jones et al., 1997; Swanson & Hoskyn, 1998; Westwood, 2000) declare it's too much to expect children to actively discover their own path towards basic knowledge in literacy and numeracy. They see these tasks as being of immense difficulty. In support Yates (1988) asserts, 'exposure to good direct teaching will enable the child to develop a more substantial knowledge base that will bootstrap the child's thinking processes in subsequent situations both in and out of school' (p. 8).

It remains to be seen how well constructivist methods perform in the various curriculum areas and whether it is more favourable to some areas than others (Stodolsky, 1988). Some supporters are of the opinion Constructivism theory applies to every province of learning, others, however, are not so sure. A major finding of Stodolsky's (1988) research was that teachers generally have a diverse range of teaching styles and that subject matter was the major determinant of which particular style they adopted in any given situation. The way they taught mathematics might differ from the way they taught science, and the way they taught one particular aspect of mathematics might differ from their approach to some

other area of mathematics. The grade level being taught must also be considered because, 'generalizability may be muted and limited depending on grade level' (Stodolsky, 1988, p. 104).

The instructional design expert, Dick (1992) is concerned that 'other spokespeople for constructivism would make it appear that the theory applies to all domains of human learning' (p. 96). However, he questions whether or not it is a complete theory for learning everything. These are important matters, which shall now be dealt with in a little more depth.

Potential problems with constructivist approaches

That children can learn from the activities they initiate themselves is inarguable, and the idea that they should have ample opportunity to use their individual initiative is also sound. However, the idea that teachers should restrain themselves from teaching may just be going too far (Grossen, 1993). If knowledge is to be a matter of personal construction, and teachers are to be restrained from imparting knowledge, how can children be expected to come by knowledge of complex theories and concepts that have taken the best brains in the world centuries to put together (Matthews, 2000)?

Some advocates of student-directed learning (e.g., DuCharme, Earl, & Poplin, 1989; Poplin, 1988a, 1988b) reject the whole idea of teacher-directed instruction. They oppose the implementation of strategy instruction in education generally because it exemplifies explicit, non-constructivist, rote learning (Grossen, 1993).

Student-directed learning practices are not simple to implement successfully in the classroom (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989). The selection of tasks can be difficult (Romberg & Kaput, 1999). And the provision of concrete situations, in themselves, are no guarantee students will be able to relate them to their own worlds — they may not be relevant for all students and thus play no part in preparing them for a changing, dynamic world (Romberg & Kaput, 1999). To theorise about how children construct their own meanings is a worthwhile pursuit but it is a long way from creating ‘a clearly defined classroom teaching model’ (Westwood, 2000, p. 6). The construction of meaning is not an easy process (Carnine, 1991; Darch et al., 1984). Simply providing students with resources and a location for collaborative learning situations will not in itself facilitate successful learning (Westwood, 2000) as not all students can be expected to put everything together unaided (Darch et al., 1984). When dealing specifically with mathematics, Carpenter and Lehrer (1999) state that knowledge ‘must also be linked to knowledge of students’ thinking, so teachers have conceptions of typical trajectories of student learning and can use this knowledge to recognise landmarks of understanding in individuals’ (p. 31). Sowder and Philipp (1999) opine that teachers have to be given ‘opportunities to revisit and reconceptualize the mathematics of these grades and to come to understand the nature of mathematical knowledge and activity that are necessary for pedagogical effectiveness’ (p. 107).

Airasian and Walsh (1997) are well aware of the demands constructivist approaches place on teachers and caution, ‘listening and responding to student constructions will be difficult and time-consuming’ (p. 36). And in order for this to happen, teachers themselves need to thoroughly

understand the body of knowledge behind the discipline (Bransford, Brown, Cocking, Donovan, & Pellegrino, 2000). The implementation of a successful lesson based on constructivist principles requires both careful planning and considerable practical skills and it is a big ask to expect teachers to put in the required time to plan constructivist lessons with the appropriate degree of care and precision and then expect them to carry out classes effectively (Westwood, 2000). It does seem that the practical implementation of constructivist approaches in the classroom is a more daunting exercise than one might be led to believe by much of the literature put out by its advocates (Dick, 1992).

Although there are obvious potential benefits to be gained from properly implemented student-directed learning we must question whether or not this approach is appropriate for all (Westwood, 2000). Students will differ with respect to exactly what benefits, and how much meaning they can extract from a class where there is no explicit instruction (Darch et al., 1984). Indeed, to some the learning experience presented in the constructivist class may not appear at all warm or inviting, to some learners it may appear inhibiting and complicated, they may see themselves as being forced to make choices they feel ill-prepared to make — they may feel uncomfortable about making them (Perkins, 1992).

Some authorities argue that certain students learn better when they are explicitly taught and when the curriculum is structured (e.g., Becker, 1988; Darch et al., 1984; Engelmann, 1980; Tuovinen & Sweller, 1999). They state that students who face challenges in learning need to have more structured and explicit instruction, not less (Gersten & Carnine, 1984; Johnston, Proctor, & Corey, 1994). And certain students 'do not always relate the

knowledge they possess to new tasks, despite its potential relevance' (Bransford et al., 2000, p. 237). It is also argued that some students become confused and frustrated when exposed purely to discovery approaches to learning, and it is important to bear in mind that learners differ greatly in their need for teacher direction (Harris & Graham, 1996).

The traditional approaches to teaching concentrate on getting students to process information efficiently and effectively, whereas the constructivist approaches are directed towards students' awareness of, and reflection on the learning experience, and since learners will all have their own unique perspective of that experience the concept of global learner is not part of the constructivist perspective (Bednar, Cunningham, Duffy, & Perry, 1992).

When it comes to students with special needs Mastropiere, Scruggs, and Butcher (1997) assert 'special educators historically have been skeptical about the effectiveness of discovery, inquiry, or constructivist methods for students with disabilities' (p. 200). In fact a significant body of research indicates that many students with special needs require detailed and explicit instruction in order to acquire a variety of cognitive and metacognitive strategies (Carnine, 1997; Englert, 1984; Harris & Graham, 1996). Indeed, it would seem to Graham and Harris (1994) that for students with special needs to master the higher-order processes they may require 'more extensive, structured, and explicit instruction' (p. 284).

It should be noted that even some of the more capable students experience problems when exposed to a constructivist approach (Westwood, 2000). Various teachers have observed that even perfectly capable students can be frustrated by methods that require of them

discussion, analysis and reflection (Westwood, 2000). Some educators (e.g., Grossen, 1993; Hempenstall, 1997) are of the opinion that when dealing with students who have little knowledge in a subject area, it is better to provide them with more structure, but in classes where students have more knowledge 'the advantage of additional structure may disappear' (Tuovinen & Sweller, 1999, p. 340). Bransford et al. (2000) caution that knowledge students already have when they enter new situations can have the effect of misdirecting them.

Special consideration must be given to whether or not a constructivist approach is appropriate for every aspect of learning (Westwood, 2000). According to Gagne, C. W. Yekovich, and F. R. Yekovich (1993) 'Effective teachers appear to have many instructional strategies that are *conditioned* on student performance' (p. 466). It is well accepted that the learning processes involved in acquiring facts and concepts are different from the learning processes involved in the development of intellectual skills and strategies (Gagne et al., 1993). Whereas traditional practice was achieved by way of regular drills, revision and rote memorisation, constructivists assume students will achieve basic number knowledge and skills through problem solving, enquiry and discourse (Westwood, 2000).

Westwood (1999) argues because constructivist approaches cannot guarantee children will acquire fluency and automaticity with basic number and computation it is dangerous to exclude all manner of instruction. And Resnick cautions 'Although it is not new to include thinking, problem solving, and reasoning in *someone's* school curriculum, it is new to include it in *everyone's* curriculum' (1987, p. 7).

Research on constructivist approaches

The US Department of Education publication *Hard Work and High Expectations* (1992) concludes that western belief in students' inherent potential to learn constitutes the major difference between eastern and western cultures. In the Department's opinion this belief is responsible for America's failing academic standards. Eastern cultures remain firm in the belief that hard work is the major factor in the learning process (Stevenson & Stigler, 1992). 'Japanese schools not only teach the value of effort but teach children to make an effort' (United States Department of Education, 1992, p. 27). Significantly, Asian students spend more time studying and consistently achieve the highest scores on international assessments (Ma, 2000). The history behind these conclusions will become apparent in the following review.

The progressive education model

Student-directed education has been extensively evaluated both in England, where it was called *progressive education* and in America, where it was called *open education*. The English termed student-directed education, progressive learning, because they expected the practice to have a leveling effect on social class differences and provide the catalyst for social change. For over twenty years it was England's official educational policy (cited in Sharp, Green, & Lewis, 1975) and was officially adopted in the Plowden report (1967) remaining in effect till its ultimate rejection in 1992 (Grossen, 1993).

Progressive education, it was thought, would promote the democratic ideals of individual freedom and autonomy, however, qualitative evaluations of its effect on working class children showed that it actually had anti-progressive outcomes (Grossen, 1993). A comprehensive evaluation (Sharp et al., 1975) on a school that was judged to be a model of progressive education in England, found that the teachers gave successful students far greater attention, interacted with them more frequently and generally paid more attention to their activities. The higher performing children who were from a higher social class received more attention than the lower performing children who were from the lower working class families, thus perpetuating the inequalities of the very system progressive education was attempting to radically reform (Grossen, 1993). Sharp, Green, and Lewis (1975) suggested that, 'modern child-centred education is an aspect of romantic radical conservatism' (p. 227). They concluded that student-directed methods of learning actually had the effect of reinforcing the existing social class structure rather than leveling it, as was the intention (Grossen, 1993). They formed the opinion that progressive educators were 'unwilling victims of a structure that undermines the moral concerns they profess' (Grossen, 1993, p. 227).

The philosophy behind progressive education in England was against the implementation of standardised tests, thus the learning outcomes of the model were only evaluated with the relatively recent advent of international competitiveness in education. In 1992 there was an outcry when the English Department of Education and Science (DES) reported on an international comparison that more than 60 per cent of the schools in the English sample scored below the lowest scoring Japanese school (para. 49). The official report (Department of Education and Science, 1992)

unequivocally blamed the poor achievement levels of English students on the Government endorsed progressive learning model and in 1992 English educational policy officially endorsed teacher-directed instruction (Grossen, 1993). The conclusion was reached that, 'Whatever else they do primary schools must get their policies and practices right for teaching the basic skills of literacy and numeracy' (Department of Education and Science, 1992, para 50). In what was the lengthiest and most comprehensive implementation of student-directed learning practices on record the English admitted their experiment with progressive education had failed (Grossen, 1993).

The open education model

The US version of the British model of progressive education was called *open education* and it was thoroughly evaluated, along with other student-directed and teacher-directed learning models, in Project Follow Through (for the original Follow Through report see Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977), the largest, most expensive research in the history of education (Bereiter & Kurland, 1981-1982). The project began in 1967 with Lyndon Johnson's war on poverty and was government-funded right up until 1995 (Grossen, 1995). This was a gigantic government initiative aimed at breaking the cycle of poverty by providing the underprivileged of the US with a better education. Over a period of almost 30 years and at cost to the taxpayer of more than a billion dollars Project Follow Through included over 70,000 children in over 180 schools (Hempenstall, 1997).

Abt Associates (1977), an independent assessment agency, gathered and evaluated the data from the open education model in all eight of its sites. Their results showed that open education had more negative outcomes, that is, significantly lower scores than those achieved in traditional education, than positive ones on measures of basic skills, cognitive development, and affect on self-esteem (Adams & Engelmann, 1996). Across multiple implementations and settings, open education was shown to be inferior to traditional education (Adams & Engelmann, 1996). The Abt report stated that the children in the open education model performed below children learning from traditional instruction and concluded that in most sites these children 'perform below expectations on a number of the basic skills and cognitive conceptual skills tests' (Abt Associates, 1977, p. 121).

An analysis of the comparison data in Engelmann, Becker, Carnine, and Gersten (1988) also shows that of all the learning models evaluated in Project Follow Through the lowest achievements were consistently obtained by the student-directed models. And the other student-directed models that differed in some way from the British progressive education model fared no better. The test evaluations by Abt Associates (1977) of all four student-directed models were consistently below those of comparable disadvantaged children learning in traditional classrooms (Bereiter & Kurland, 1981-1982).

The student-directed learning adherents have been accused (Grossen, 1993; Hempenstall, 1996) of failing to reassess their ideas in the light of the empirical data and make the appropriate instructional adjustments. According to some academics they have simply repackaged the same old themes (Grossen, 1993; Matthews, 1992), renamed the model and pushed

it as hard as they could. Indeed, Matthews (1992) has described Constructivism as *old wine in new bottles*. Whether this criticism is entirely justified is of course open to debate, but what is undeniable is that their push has been successful. The education model now being promoted for elementary school reform in the US is called *developmentally appropriate practice* (DAP) and it is so remarkably similar to the open education model that fared so badly in Project Follow Through that it is impossible to discern any significant difference (Grossen, 1993). The DAP model contains the same student-directed learning themes that were the basis of progressive education and open education. Grossen (1993) argues that the underlying philosophy of DAP is also identical as can be seen from the NAEYC position statement where the theories of Piaget (1952), Erikson (1963) and Vygotsky (1978) are specifically acknowledged as the guiding theories (National Association for the Education of Young Children, 1987).

Direct Instruction

Direct Instruction is a comprehensive system of education involving all aspects of instruction from the actual organisation and management of the classroom to the quality of teacher and student interaction, and design of curriculum materials (Gersten & Carnine, 1986; Gersten, Carnine, & Woodward, 1987). It should be noted that the term Direct Instruction as it appears in the research literature means different things to different people (Stein, Carnine, & Dixon, 1998), which often leads to confusion and non-constructive discourse about Direct Instruction in the educational community (for a basic analysis of these terms see Cotton & Savard, 1982). Direct Instruction has been referred to as teacher-directed

instruction, which is the opposite of the child-centred approach where the teacher is regarded as simply a facilitator for the students. In contrast, Direct Instruction is a highly structured system of teacher-student interactions (Stein et al., 1998).

The University of Oregon Model (Distar) of Direct Instruction, which originated at the University of Illinois in 1964 (for more complete descriptions of the curriculum and the philosophy of instruction see Kinder & Carnine, 1991) was extensively evaluated in Project Follow Through and came out in the most positive light (for an in-depth review see Adams & Engelmann, 1996).

Primarily the Direct Instruction program was intended for use with disadvantaged children to compensate for the learning deficits prevalent in socioeconomically disadvantaged areas (Becker, 1978). The Distar Direct Instruction model, originally termed *direct-verbal instruction*, evolved out of the experiences of Engelmann, Becker, Carnine and their colleagues' work in the field of compensatory education (Gersten, 1985).

Direct Instruction in practice

The Direct Instruction model provides reading, language and arithmetic instruction (Gersten, 1985) using a model that concentrates on small student-teacher ratios, rapid instruction, positive reinforcement and immediate corrective feedback (Becker, 1978). Heavy emphasis is placed on thorough teacher training and the efficient monitoring of student progress. The program involves explicitly detailed teacher instructions termed *scripts*, a signal system for cueing student response, and the

application of procedures to stimulate motivation. Other features of the Direct Instruction model include the teaching of general case problem solving strategies; use of oral instruction as opposed to written worksheets; and systematic correction procedures that transform errors into constructive learning experiences (Gersten & Maggs, 1982).

Direct Instruction is based on the assumptions that all children can be taught, but in order to *catch up*, low-performing students must be taught more, not less, and that the teaching of more involves the efficient use of technology and time (Cotton & Savard, 1982). It is a teaching model that sets out to control all the variables that affect the academic performance of students. A primary element of Direct Instruction is faultless communication (Engelmann & Carnine, 1982), and for faultless communication to take place instructional materials and teacher delivery must be clear and unambiguous (Kinder & Carnine, 1991). To maintain control over the instructional environment all Direct Instruction lessons are scripted for the teacher with precise instructional wording. This guarantees that teachers introduce the concepts in an orderly and efficient manner. The scripts give teachers specific teaching examples, sequenced teaching tasks, and a variety of test examples. This not only minimises the opportunity for errors and confusion to creep in, it allows teachers to more efficiently utilise instructional time. Engelmann (1980) is adamant that instructional material must have 'sequences or routines that are consistent with a single interpretation. A sequence or routine must pass the test of logical inspection to assure that the appropriate generalizations are described and no other generalizations are described' (p. 35). On the other hand, however, Gersten, Woodward and Darch (1986) point out that 'no curriculum is teacher-proof' (p. 23).

In the introductory stages of Direct Instruction the overall strategy becomes apparent. Every step in the application of rules and in problem solving must be explicitly taught; teachers must demonstrate each step involved in the application of a rule or strategy with detailed explanation (Vail & Huntington, 1993). Firstly, suitable examples are given to portray the intended meaning of the concept in an appropriate order of presentation. Students are provided with sequenced series of examples and are taken through the steps involved in application of the rule or strategy (Gersten et al., 1986). To ensure the learner focuses on the examples and their various features, examples are used that share the greatest possible number of features. Subsequently examples are selected that have only one attribute in common with a variety of irrelevant attributes to make the common attribute stand out (Carnine, 1991). To illustrate differences, minimally different examples are juxtaposed and treated differently. To illustrate sameness, greatly different examples are juxtaposed and treated the same. Sequences include instances when the rule is applied and instances when it is not, that is, when similar examples except for the critical quality are used to aid in discrimination (Vail & Huntington, 1993). In successive instruction the overt steps are faded and the number of questions reduced resulting in student strategy becoming increasingly covert (Fielding, Kameenui, & Gersten, 1983). Finally to comprehensively test student understanding, examples are juxtaposed that bear no predictable relationship to each other. Ultimately students will apply the strategies alone and silently (Kinder & Carnine, 1991).

Potential problems with Direct Instruction

Much of the critical literature on Direct Instruction hinges around philosophical issues such as the proper role of the teacher and the real nature of the learning process (Hempenstall, 1996) and while these are indeed matters that demand constant academic attention, they are outside the ambit of this thesis.

A serious drawback with Direct Instruction is that it requires considerably more work from the teachers in respect of both the acquisition of new skills and in the actual teaching process itself. The role of the teacher can seem a complex one and where there is a perception by teachers that the innovation is highly complex, the level of implementation is low (Paul, 1977).

Another major criticism with Direct Instruction is that there is little room for the teacher to move — the models are highly restrictive and some teachers view themselves as little more than automatic delivery systems blindly spitting out instructions to automatic receivers (Barnes, 1985). The most notable feature of Direct Instruction is the high degree of specificity in terms of teaching behaviours (Hands, 1993). Doyle and Ponder (1977) regard specificity as an asset, though Fullan and Pomfret (1977) quite justifiably talk of the *dilemma of explicitness*. Highly specific innovations can be unsuitable in a variety of situations — some people may be overwhelmed by them (Fullan, 1982). In this writer's opinion innovations should not be so specific as to not allow teachers to adapt them to suit their own classrooms. There is no doubting the specificity of Direct Instruction, the lessons are entirely scripted, and the behaviour of the teacher is carefully defined. Generally speaking, people don't like being

restricted and teachers are no exception. According to Hands (1993) this probably constitutes the major problem with Direct Instruction.

Direct Instruction is seen by some teachers as a threat to their decision-making ability and an impingement upon their creative talents and capacity to innovate (Fields, 1986). Gersten and Guskey (1985) reported that at first teachers felt too constricted by a method that took all their decision-making capacity away from them. Researchers (Becker, 1977; Cole & Chan, 1990) noted that some teachers regarded Direct Instruction programs as too highly disciplined and prescribed.

As mentioned earlier there is some merit in the argument that teaching methods relying on rote memorisation of basic number facts and algorithmic procedures are boring and don't require students to reflect and think much (see Wakefield, 1997), but much of the criticism Direct Instruction has received is vague and unsubstantiated (Hempenstall, 1997). For reasons difficult to discern Boomer (1988) was of the opinion that Direct Instruction's 'side effects may be lethal' (p. 12), Reetz and Hoover (1992) that students might find it unacceptable, and Schweinhart and Weikart (1986) went further claiming Direct Instruction was directly responsible for leading children to delinquency. However, according to Mills, Cole, and Dale (2002) gender difference provides a more likely explanation than the earlier Schweinhart and Weikart findings, specifically stating that Direct Instruction may be implemented 'without fear that the method will result in later antisocial behavior' (p. 93).

Some Direct Instruction critics thought the tightly controlled instruction might discourage children from expressing themselves freely and

consequently have a detrimental effect on their self-esteem (Stebbins et al., 1977). Then, against the weight of considerable evidence favouring Direct Instruction with special education students, Kuder (1991) specifically finds it may be inappropriate for those in special education.

In the quest for an ideal approach to learning Bay, Staver, Bryan and Hale (1992), promote discovery learning as the way to go, while Weaver (1991), relying on questionable support, is of the opinion Whole Language is more effective. The debate will probably never be settled to the satisfaction of all but it seems to this writer important to focus more on what the research is telling us, and less on eloquent concepts that appeal to our sense of the aesthetic.

Direct Instruction research

The aforementioned Project Follow Through was instigated to identify teaching models that could elevate the academic performance of America's underprivileged schools from the 20th to the 50th percentile. The approaches taken by the different models came under three theoretical orientations. The first of these was a *behavioristic* orientation founded on the notion that all behaviours are learned; this was termed the Basic Skills Model. Here the teacher directs the learning process reinforcing the behavioural objectives. The next orientation was *cognitive development* founded on the developmental sequence characterising normal cognitive growth; this was termed the Cognitive/Conceptual Skills Model. Here the teacher provides cognitive materials and experiences appropriate to the age of the students. The third orientation was *psychodynamic* founded on the notion that the development of the 'whole child' is essential for educational

improvement; this was termed the Affective Skills Model (Adams & Engelmann, 1996, p. 68). Here the teacher provides an environment in which children decide what is best for their personal growth.

In the final analysis it was shown that the students being taught under Direct Instruction, a behavioristic orientation, scored close to the 50th percentile in every subject, the targeted level of performance for all subjects (Stebbins et al., 1977). Under the other models students consistently scored lower than the 20th percentile (Adams & Engelmann, 1996). As shown in Figure 1 with the exception of students in the Direct Instruction and Behavior Analysis models the math results were consistently poor. These were the only models to achieve above the 20th percentile. 'Direct Instruction students scored 20 percentiles ahead of the second place group (Behavior Analysis) and 37 percentiles higher than the last place group (Cognitive Curriculum)' (Adams & Engelmann, 1996, p. 82).

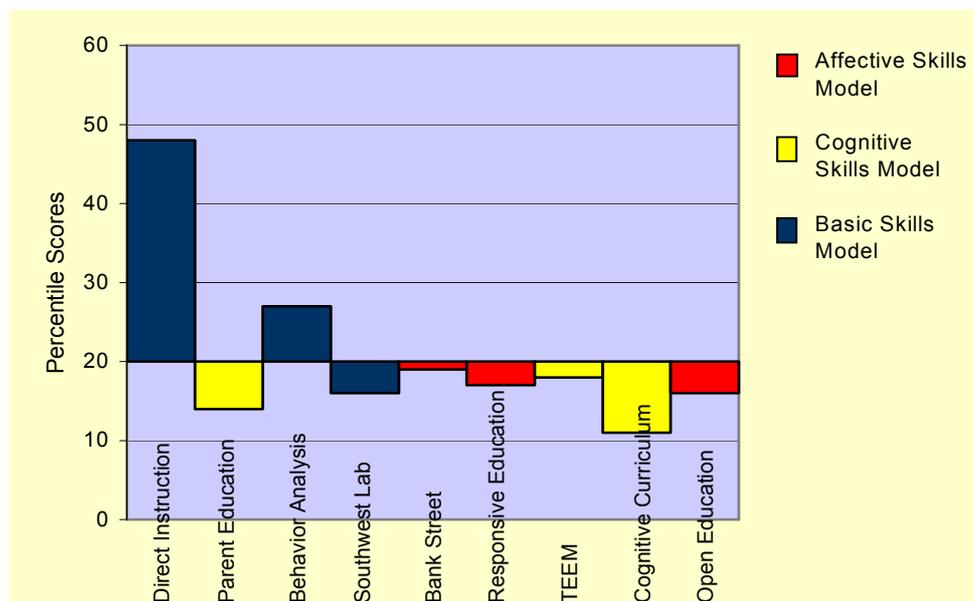


Figure 1 Math percentile scores across Follow Through models

Note. Adapted from Adams & Engelmann, 1996, p. 83.

Years later Gersten and Carnine (1986) reanalysed the data to look for implications that might relate to special education and mathematics and found that 'Direct Instruction Follow Through students achieved at a much higher level than is typical for students with similar demographic characteristics ... in fact, their mean performance was at a level comparable to their middle-income peers' (p. 402). However, Gersten, Woodward and Darch (1986) argue that 'Mean scores can be deceptive since the mean is heavily influenced by some extremely high scores' (p. 25).

The analysts of the Follow Through evaluation data (Bereiter & Kurland, 1981-1982; Lindsley, 1992; Stebbins et al., 1977) all agreed that structured, teacher-directed instruction resulted in stronger academic outcomes than the popular child-centred models.

The two high-scoring models according to our analysis are Direct Instruction and Behavior Analysis; the two low-scoring are EDC Open Education and Responsive Education sponsors of both the Direct Instruction and Behavior Analysis models call their approaches "behavioral" and "structured" and both give a high priority to the three Rs. EDC and Responsive Education, on the other hand, are avowedly "child-centered". Although most other Follow Through models could also claim to be child-centered, these two are perhaps most militantly so and most opposed to what Direct Instruction and Behavior Analysis stand for. (Bereiter & Kurland, 1981-1982, p. 16-17)

The core of Direct Instruction is its almost unrelenting focus on 'skill mastery by *all* students' (Gersten et al., 1986, p. 28). While initially the bulk of research into Direct Instruction was carried out in the provinces of those with learning difficulties and the disadvantaged or underprivileged its teaching principles have proved to be of value over a much wider range of learners. Lockery and Maggs (1982) listed a broad array of Direct Instruction research

findings showing success with a diverse range of students including average children, children with mild, moderate or severe skill deficits, children in withdrawal classes and special classes in regular schools, disadvantaged students, and children whose first language was not English.

Gersten (1985) reviewed studies of students with a range of disabilities and concluded overall that Direct Instruction resulted in higher academic gains than traditional approaches. He opined the mastery criterion (in excess of 90 per cent) was particularly important for special education students, suggesting a more formative evaluation where only one instructional variable was manipulated. He also called for more research into instructional dimensions in order to highlight those variables associated with academic gains.

White (1988) analysed the effect of Direct Instruction on special education students through a meta-analysis of 25 studies, 21 of which included students with mild disabilities (Stebbins et al., 1977). The focus was on those studies employing equivalent experimental and comparison groups. It is concluded that none of the studies significantly favoured the comparison group. In fact, more than half the measures significantly favoured the Direct Instruction group. The data clearly showed that Direct Instruction was effective over a broad range of grades from elementary through to secondary (Bender, 1993). White concluded that 'instruction grounded in Direct Instruction theory is efficacious for both mildly and moderately/severely handicapped learners, and in all skill areas on which research has been conducted' (p. 372).

Kavale (1990) comes out strongly in support for the Direct Instruction model. He concluded in his summary of research into Direct Instruction and effective teaching that the Direct Instruction model was five to ten times more effective for learning disabled students than other models. Direct Instruction and Precision Teaching are referred to by Binder and Watkins (1990) as the approaches best supported by research in the English-speaking world to address the problems of teaching.

It is worth noting here that while there has been a substantial amount of research into the University of Oregon Direct Instruction model since Follow Through, precious little of it has been into mathematics. The findings in the following relatively diverse range of studies where Direct Instruction was employed in the teaching of mathematics are significant.

Darch (1984) conducted a study to examine the effectiveness of a method that teaches fourth graders to translate word story problems into mathematical equation form in a step-by-step explicit manner that closely paralleled the Direct Instruction Follow Through programs. This method was compared to a method developed from a composite of four basal arithmetic texts adopted for use in the state of Oregon. Posttest results indicated a significant positive effect for the explicit model.

Kelly, Carnine, Gersten, and Grossen (1986) conducted a study to examine the effectiveness of a 10-day intervention in teaching fractions to learning-disabled and remedial high school students who pre-tested at less than a 50 per cent accuracy level. The study compared the *Mastering Fractions*, Direct Instruction videodisc program to a basal program. The videodisc program provides a system for mastering addition, subtraction,

and multiplication of fractions, common and mixed numbers, fractions on a number line and simplifying. 'The DI group showed significant differences at the posttest (95.1% vs. 79.1%). After 2 weeks, a maintenance check showed high retention for the DI group (93.8% vs. 70.2%)' (Adams & Engelmann, 1996, p. 57).

Moore and Carnine (1989) conducted research on problem solving requiring the use of ratios and proportions, and found high school students with disabilities who received Direct Instruction scored as well as their non-disabled peers receiving traditional math instruction.

Grossen and Ewing (1994) conducted research on the application of fractions, decimals, and percentages, with Year 5 and Year 6 low ability students who received Direct Instruction and found they scored significantly higher than high ability students in a constructivist treatment.

Tarver and Jung (1995) conducted a study to compare the effects of a Direct Instruction mathematics curriculum, Connecting Mathematics Concepts (CMC), and a discovery learning mathematics curriculum, Math Their Way combined with Cognitively Guided Instruction (MTW/CGI), with students in Year 1 and Year 2. At the end of Year 2, CMC students had significantly higher scores than MTW/CGI students on both the computation and the concepts/applications components of mathematics achievement. In addition, the CMC students had significantly higher scores on a survey of student attitudes towards mathematics. Comparisons of grade equivalents suggest that Direct Instruction CMC curriculum benefited high ability as well as low ability students.

It is also worth noting research has shown that when students with learning difficulties are removed from Direct Instruction programs though they continue to perform better than their socioeconomic peers who did not have the benefit of instruction under the Direct Instruction program, they soon begin to lag behind the national norms (Cotton & Savard, 1982). Though in the first place Direct Instruction research and program development was concerned solely with instructional design for basic skills its scope has been broadened to take in higher order skills (Kinder & Carnine, 1991).

On examination of the research and reviews of Direct Instruction one has to conclude the results are impressive. More than twenty years ago in a major review of research literature into school effectiveness Cotton and Savard (1982) concluded that much well designed research had gone into the concept of Direct Instruction and they found that students taught with this method 'consistently outperformed comparison students in all basic skill areas'.

Merging views

Various scholars (W. Bishop, 1999; Casazza, 1993; Drecktrah & Chiang, 1997; Harris & Graham, 1996; Westwood, 2000) are now seriously arguing that there are definite advantages in striking a balance between teacher-directed learning and student-directed learning (for an in-depth examination of teacher attitudes in this regard see Grant, 1998).

According to Westwood (2000), certain constructivists (e.g., Noddings, 1990) openly acknowledge that students need basic skills when it comes to solving certain problems or understanding certain concepts, and that to properly establish these skills they will need teacher-directed learning and

periodic practice. Airasian and Walsh (1997) warn, 'it is a misunderstanding to consider teaching methods such as memorization and rote learning useless' (p. 447). Von Glasersfeld, a committed constructivist (cited in Airasian & Walsh, 1997) acknowledges there are 'matters that can and perhaps must be learned in a purely mechanical way' (p. 447). And while Battista (1999) wants to see teaching that encourages reasoning and reflection rather than rote learning, he is mindful of the importance for students to become fluent with the basic number facts which are necessary for computation and problem-solving. The aforementioned Inagaki, Hatano and Morita (1998) while encouraging discourse amongst students, also see the teacher's intervention as essential.

It seems fairly well accepted that problem-solving skills normally operate from a knowledge base that has been acquired through practice (Dougherty & Johnston, 1996). And it is actually when the knowledge in a discipline is being acquired that the foundations for effective problem solving are being laid (Resnick, 1987). This knowledge base can be instantly tapped into without any great mental effort to aid in the application of higher-order processes. Automaticity, which originates from practice (Dougherty & Johnston, 1996), gives students the opportunity to maximise their mental powers on more complex tasks (Bloom, 1986). If the essential knowledge required for automaticity is to be stored in the student's long term memory it needs to be explicitly taught (Engelmann, 1980) and practiced repeatedly (Dougherty & Johnston, 1996). Genuine competence in both problem-solving and basic skills only comes with constant practice (Engelmann, Carnine, & Steely, 1991).

Beck, Perfetti, and McKeown (1982) demonstrated that a student of average ability required 16–22 presentations (a degree of instruction rarely encountered in today’s educational environment) of a new concept before learning and remembering it. Clearly, if a student of average ability needs this amount of presentations to learn a new concept, students with learning difficulties will almost certainly require more (Vail & Huntington, 1993). This places considerable demands on the teacher (not to mention the student), both in terms of actual teaching and degree of intensity of teaching (Vail & Huntington, 1993). The teacher would be expected to develop the curriculum incorporating ‘daily review, weekly review, monthly review, and quarterly review of materials presented previously, as well as the presentation of any new concepts or skill materials at least 20 times each’ (Vail & Huntington, 1993, p. 164).

In a move appealing to common sense, the legal concept of *fitness for purpose*, which requires vendors to warrant the goods they are selling are reasonably fit for the purpose for which they are being sold, has been brought into the education equation. Galton, Hargreaves, Comber, Wall, and Pell (1999) argue that the method of instruction best suited to the type of learning required by the lesson should be adopted and in deciding this matters such as age and ability of students need to be taken into account. It is their assertion that teaching methods should be assessed for ‘fitness for purpose’ (p. 184).

It would seem, at least when it comes to the acquisition of basic skills, that an instructional approach is more suitable. While explicit teaching is not necessary in order for children to acquire basic mathematical skills, it is generally accepted that most mathematical understanding, both basic and complex can be successfully acquired through its implementation (Bjorklund,

1995). Resnick (1987) sees mathematics as different from other disciplines in that here the 'particular knowledge structures must be learned' (p. 38).

Grant (1998) in his study of teacher attitude towards systemic reform found that while teachers may embrace student-directed learning in one subject area they may ignore it in another. The combined approach is seen as especially relevant when teaching children with special needs. For these children Harris and Graham (1996, p. 29) are of the opinion that:

No one intervention or viewpoint can address the complex nature of school failure or success, or, for that matter, of social inequalities and inequities. We, like other advocates of constructivism or whole language believe that an integration of knowledge and successful practices is critical in today's schools.

Certain educators feel that expecting our teachers to take on major teaching reforms 'with only their own meagre resources at hand seems naïve' (Grant, 1998, p. 208). The teachers of today have neither the training nor the time to sit down and design math curricula and then go through a comprehensive evaluation process. Jones et al. (1997) state unequivocally 'teachers must have at hand effective instructional procedures, materials, and other resources. At the present time they must do much of the work of improving mathematics education themselves' (p. 161).

Alexander (1995) warns, 'While ideology dictates a teacher role of facilitator and encourager, common sense (not to mention recent classroom research) indicates the benefits for children of purposeful intervention by the teacher, especially of a kind which generates cognitive challenge' (p. 31). Whatever reforms are made in education, however, they

'will not carry the day unless they tap the effort as well as the ability of our children' (United States Department of Education, 1992, p. 19).

Chapter 3: Self-efficacy

Overview of the chapter

This chapter examines the role *self-efficacy* plays in the learning process and analyses the various constructs that come into play in predicting and mediating academic performance. Differences between self-efficacy beliefs and self-concept will be first explained then findings on the relationship between self-efficacy and engagement, motivation, self-regulation and modeling are summarised. The research and literature on mathematics self-efficacy is reviewed and conclusions are drawn from this and the preceding chapters.

Self-efficacy

Bandura has taken the abstract concept of *confidence in learning* and placed it firmly in the academic realm. Under the nomenclature of self-efficacy he has pioneered the analysis and refinement of the role confidence plays in the learning process in great depth. It is not without reason he has consistently distinguished self-efficacy from confidence (Bandura, 2002) for without its own academic title and defined domain, it would have been a difficult task to bring to the general area of confidence in learning the serious level of academic discourse the subject has attracted.

To perform a task competently one requires not only the requisite skills but also the self-belief in one's ability to implement performance. General operative efficacy requires constant improvisation of a range of skills and

sub-skills in order to handle circumstances, which often change in unpredictable and ambiguous ways. Indeed, activities that we would regard as routine are seldom executed in precisely the same way; there is always something slightly different about the way they are effected. According to Bandura's (1986) *Social Cognitive Theory* what we do in any given circumstance is governed in large part by what we think we can do. In other words self-efficacy is not so much concerned with one's skills but what one thinks one can do with those skills. Bandura portrays an image of human behaviour and motivation where peoples' self-beliefs are fundamental components. He defines an individual's perception of self-efficacy as 'a judgement of one's capability to accomplish a certain level of performance' (Bandura, 1986, p. 391). It is his contention that self-efficacy beliefs powerfully influence the choices people make, the amount of effort they expend, the length of time they will persevere in adverse circumstances, and the measure of anxiety or confidence they will bring to a given situation. He terms the judgement of the likely consequence of that performance an 'outcome expectation' (Bandura, 1986, p. 391).

Self-efficacy beliefs are personal judgements and the satisfaction individuals gain from their activities is largely determined by their own standards of self-evaluation. Thus when we talk about Jack or Jill's self-efficacy we are actually discussing their perceived self-efficacy, how well they think they can perform a given task. Those with high personal expectations will naturally enough set themselves high standards and attribute any failure to personal lack of effort whereas someone with comparable skill but with low self-expectation will attribute failure to personal inability. Similarly those with a strong sense of efficacy are undeterred by obstacles, in fact they serve to make them work harder —

they seek to influence and produce their future rather than foretell it (Bandura, 1986). It is also important to realise that if the student does not have the required knowledge and skill to perform a particular task then high self-efficacy in itself will be insufficient to successfully complete that task (Schunk & Zimmerman, 1997).

Self-efficacy and self-concept

One must be careful not to confuse self-efficacy with self-concept for there exists between the two a conceptual distinction, which has often been blurred (Marsh, 1990; Pajares, 1996). Hanchon Graham (2000) attributes this in part to the fact that 'self-concept has itself had a long history of ambiguity with regard to its own definition, measurement and interpretation' (p. 57). Some researchers into academic motivation have even gone as far as to employ the terms interchangeably (see Chapman, 1988; Meece, Wigfield, & Eccles, 1990; Wilhite, 1990). Generally speaking though, research into self-concept has employed more generalised indexes and relied more on correlational data than experimental (Bong & Clark, 1999).

Shavelson, Hubner and Stanton (1976) in broad terms define self-concept as being a 'person's perception of himself' (p. 411). Bandura (1986) refers to it as 'a composite view of oneself that is formed through direct experience and evaluations adopted by significant others' (p. 409). Pajares (1996) explains that self-concept embraces judgements about personal competence, feelings of self-worth and evaluative reactions. More detailed definitions have been attempted (see Bong & Clark, 1999) but they inevitably intrude into the more task specific area of self-efficacy and thus tend not to be helpful (Hanchon Graham, 2000). In her discussion of the

differences between self-efficacy and self-concept Hanchon Graham (2000) adverts to an important distinction existing in the manner in which they are measured. The former involving judgements of capacity to perform specific tasks, the latter involving judgements of self-worth with respect to performance. Bong and Clark (1999) see self-concept as being a more generalised view of one's competence and as such of limited use as a predictor of effective task implementation. Though self-concept judgements are never task specific they may be domain specific and while the two are strongly related (Marsh, 1990) this is not necessarily so. A student with high self-efficacy in mathematics may feel no great self-worth in high mathematical achievement because the student places little value on this skill (Pajares, 1997).

Some academics (see Marsh, Walker, & Debus, 1991) view the difference between self-concept and self-efficacy as one that flows from their respective sources. When assessing one's self-worth inevitably one has to compare oneself with one's peers whereas when assessing one's capacity to accomplish a specific task the first thing one goes to is one's previous experience in this area. This way of looking at self-efficacy, however, has limits for if one has no prior experience at a task then social comparison becomes critical (Bandura, 1997).

It must be understood that self-efficacy beliefs are of a context specific nature (Schunk, 1991) and Bandura (1977; 1986; 1997) has gone to great lengths to stress they be measured at 'the optimal level of specificity' (Hanchon Graham, 2000, p. 30). In other words whatever self-efficacy belief is being assessed it should not only be related to the subject's prior performance of the same or similar exercises it should be administered as close in time as

possible to the prior performance (Bandura, 1997). Though herein lies the major difference from self-concept it should be noted that a hierarchical model of self-concept using domain specific indices is now widely seen as being an appropriate model for research purposes (Marsh & Shavelson, 1985). Introduced by Shavelson (in 1976) this model distinguishes between general, social, physical, emotional and academic self-beliefs further subdividing the latter into Maths, Science, English and History.

Although in the past definitions and measurement have been equivocal (Hanchon Graham, 2000) the findings clearly show there is a strong relationship between self-concept and self-efficacy (see Byrne & Shavelson, 1987; Marsh, 1989; Shavelson & Bolus, 1982). Equally clearly, the more precise the measurement of self-concept becomes the closer the boundary will be moved towards self-efficacy and the more difficult it will be to determine whether in fact a boundary exists.

Sources of self-efficacy

The four major sources of self-efficacy are: performance attainments, vicarious experiences, verbal persuasion, and physiological states (Bandura, 1986). Although they are referred to as sources of self-efficacy they do not translate directly into self-efficacy beliefs (Pajares, 1997); what actually happens is people base their judgements on their own individual interpretations of their actions (Pajares, 1997). Thus, two people sitting an exam may achieve the same result yet each forms a totally different performance judgement. For Jack the mark is higher than he is used to getting and he can see improvement, whereas for Jill the mark is lower and she can only see deterioration. Jack's self-efficacy is raised and Jill's is lowered.

By far the most influential of these sources is personal performance attainment (Bandura, 1977, 1986) and this is because it is based on personal experience of mastery (Pajares, 1997). While students engage in learning they automatically monitor their progress towards academic goals (Pajares, 1997). This behaviour has the effect of modifying self-efficacy beliefs in that as goals are attained students know they are capable of performing certain tasks and their confidence with respect to future learning is enhanced (Schunk, 2001). Their performances provide dependable data on which to base their self-efficacy (Schunk, 1998). Just as repeated success raises self-efficacy beliefs repeated failure lowers them (Bandura, 1977, 1986). Repeated success establishes a strong sense of self-efficacy where the odd failure here and there is not likely to impact upon one's perception of one's abilities (Bandura, 1977).

Where such personal experience is lacking the natural thing to do, where possible, is to look to one's peers (Pajares, 1997) who have had experience and use them as a model to estimate one's own capability (Bandura, 1986). How the vicarious data will affect the self-efficacy appraisal depends on the criteria employed to evaluate the ability (Suls & Miller, 1997), but it is generally accepted that learning vicariously speeds up the learning process and can have the effect of shielding the learner from negative experiences (Schunk, 2001). Although some activities, like running and jumping, provide clear factual bases for evaluating individual capabilities one must generally turn to the performance of others to gauge one's ability in objective terms (Bandura, 1993). Almost every performance we complete is evaluated in terms of social criteria: whether it be the size of the fish you caught yesterday, the exam you sit today, or the wage you will take home at the end of the week, social comparison data plays a

major role in self-efficacy appraisals (Bandura, 1993). Bandura (1986) maintains 'even the self-assured will raise their perceived self-efficacy if models teach them better ways of doing things' (p. 400). The point has also been made that we do not always turn to real people for models when self-evaluating. 'Quite often, we can and do compare our experience with an internalized standard' but these standards are 'themselves based on personal experience' (Gruder, 1977, p. 38).

Verbal persuasion is often used as a means of inducing people to think they are capable of doing certain things, but as a means in itself of engendering self-efficacy it is of limited potency (Pajares, 1997). Verbal persuasions are at their most effective when the person subject to persuasion has valid cause to believe he/she is capable of performing the task (Chambliss & Murray, 1979). While there is no doubt that thoughtfully employed teacher encouragement can elevate self-efficacy (Schunk, 2001) lifting beliefs of personal competence to unrealistic levels not only invokes failure it invalidates the efforts of the persuader and serves to further lower the subjects perceived self-efficacy (Bandura, 1986). It should be noted that just as positive social persuasion is capable of contributing to successful performance, derogatory comments about a person's individual competence in a particular province can have the most detrimental effect (Pajares & Miller, 1994) if that person is already lacking confidence in that area (Bandura, 1986). Verbal persuasion as a means of engendering self-efficacy should thus be viewed as a delicate instrument, to be treated with respect and applied with care.

Whilst physiological indices such as perspiring palms, hollow feelings in the stomach and elevated heart rates may be important indicators of degrees of

confidence at an individual level, they vary greatly between people. Stress or tension in some people may spur them to greater heights and success while in others it may induce negative reactions and failure. Bandura (1986) notes that differing interpretations of arousal have differing effects on self-efficacy perceptions: what makes one person frightened may well see another fired up. He points out that whatever the implications of arousal on an individual's self-efficacy beliefs they derive from past experience. Generally speaking though 'fear reactions generate further fear through anticipatory self-arousal' (Bandura, 1986, p. 401) and physiological states of high arousal have a debilitating effect on performance.

Self-efficacy and engagement

According to Bandura (1986) students decide on the basis of their perceived academic capacity how much time and effort they will put into solving an academic problem. Students with high self-efficacy beliefs will persevere longer and harder even in particularly difficult circumstances (Bandura, 1986) and it is this effort and perseverance, which they invest in an academic task that has been termed *engagement* (Skinner, Wellborn, & Connell, 1990). Conversely students with low self-efficacy in a particular skill area are reluctant to engage in tasks where those skills are required and if they do they are more likely to quit when encountering difficulty (Bandura, 1986; Pajares, 1997).

Fullarton (1998) in her study, which focuses on motivation and engagement, found that throughout the entire study 'effort was seen as the most important strategy for success' (p. 204). She makes the critical point that 'to foster motivation and enhance perceived control, it is important for

children to be made aware of the connections between their efforts and outcomes' (Fullarton, 1998, p. 214). According to Fullarton there is merit in encouraging students to view their successes as the result of their 'high effort and ability' (1998, p. 214). The fact that students are aware that effort can be an effective strategy though is in itself insufficient; students need to be aware that they personally can produce the effort required to achieve the desired outcome (Fullarton, 1998).

Self-efficacy and motivation

'Without aspirations and evaluative involvement in activities, people remain unmotivated, bored, and underdeveloped in their capabilities' (Bandura, 1997, p. 17). And even individuals who are highly efficacious and skilled may not act in accord with their beliefs and abilities if they don't have the incentive — for example, where the stakes are low. The role that motivation plays in the learning process has received increased attention in recent years (Pajares, 1996) and in America the lack of it has been blamed as a major reason for students low mathematical achievement (Schunk, 1998). Motivation with respect to self-efficacy has been defined as 'a process whereby goal-directed effort is initiated and sustained' (Schraw & Brooks, 2001, p. 2). A major element in motivation is a belief in one's personal competence and this is best displayed in perseverance (Schunk, 1990). If students are not motivated to engage, knowledge and understanding of cognitive strategies will be insufficient to sustain them in the learning process (Pintrich & de Groot, 1990). Students feelings of self-efficacy are enhanced as they see their learning progress and this in turn sees their motivational constructs enhanced, however, if they think they lack the capacity to achieve their goals their motivational constructs will

diminish (Schunk, 2001). Naturally enough, if students see their peers successfully negotiate a task they are inclined to think they will be capable of similar performance. Thus, the simple knowledge in itself that their peers have performed a task successfully can enhance student self-efficacy (Schunk, 2001). Once students with high self-efficacy have pursued and attained a goal they will aim for even higher goals, which incorporate new challenges and require mastery of different motivational constructs (Bandura, 1993).

Generally speaking for goals to enhance motivational constructs they should be either context specific or incorporate specific performance standards so progress towards them can be easily gauged. General goals fail to enhance motivation whereas specific goals do; similarly goals set way in the future enhance motivation less than do immediate goals (Schunk, 2001). However, Resnick (1987) cautions, 'Motivation for learning will be empty if substantive cognitive abilities are not developed, and the cognitive abilities will remain unused if the disposition to thinking is not developed' (p. 50).

Self-efficacy and self-regulation

Though in the normal course of events modeling precedes self-regulation (Schunk & Zimmerman, 1997) and indeed in many instances provides the template for it, for convenience sake in this thesis self-regulation will be dealt with first.

Fullarton (1998) asserts that self-efficacy serves to regulate behaviour, an issue which Schunk (1998) sees as taking on more importance amongst

educators. Self-regulated learning according to Schunk (2001) is the learning that comes from 'students' self-generated thoughts and behaviors that are systematically oriented toward the attainment of their learning goals' (p. 125). It includes such things as organising an effective working environment, listening to and analysing instructions, correlating old knowledge with new, effectively utilising resources, and self-belief in one's ability to learn (Schunk, 1989). While some of these behaviours may operate generally, 'learners must understand how to adapt processes to specific domains and must feel efficacious about doing so' (Schunk, 2001, p. 125). Students with effective self-regulatory skills not only utilise their time and accomplish a particular task more efficiently, they save themselves unnecessary stress (Bandura, 1997). On the other hand students lacking these skills tend to put off doing tasks (Pajares, 1996) or do them badly, if at all (Pajares, 1997).

'The ultimate development of students' academic self-regulatory skill depends on the growing synergy between their use of self-regulated learning processes and derived forms of self-motivation, such as perceived self-efficacy' (Schunk & Zimmerman, 1997, p. 206).

Self-efficacy and modeling

By way of definition '*Modeling* refers to cognitive, affective, and behavioral changes that derive from observing models' (Schunk, 2001, p. 128). Generally speaking models are persons or characters upon whose behaviour observers will base their own individual behaviour (Schunk, 1987). Peers and authority figures such as parents, adults and teachers are generally the most powerful models for students (Schunk, 1987) and it is important not to underestimate the effect they can have. Although they

may be unaware of it at the time of learning, students may be acquiring certain knowledge from the simple observation of a model (Schunk, 1987, 2001). Watching teachers rehearse learning material leads students to think that they may be able to do likewise and this belief creates in those students 'a sense of self-efficacy for self-regulation and motivates students to engage in rehearsal' (Schunk, 1998, p. 143). It is by watching models behave and seeing what happens as a result of that behaviour that students work out what they can expect to happen if they behave similarly (Bandura, 1977). Indeed, the expectation of what is likely to happen in the event of specific behaviour serves as a powerful motivational effect (Bandura, 1986) as well as providing an informational function (Schunk & Zimmerman, 1997).

By watching models successfully perform tasks, observers not only acquire particular behaviours and skills (Schunk, 1998) but they pick up important information with respect to correctly sequencing actions (Schunk, 2001). By observing models students can see what requirements are necessary and in what order they must be undertaken to successfully negotiate a task (Schunk, 1998).

It is even claimed that students feelings of self-efficacy towards a task, which they have been unable to successfully negotiate in the past can be enhanced by their observance of a model successfully perform the same operation (Schunk & Hanson, 1985). Through the observation of others they may pick up on 'mental processes that might otherwise have remained entirely implicit' (Resnick, 1987, p. 40).

Although by observing models students can pick up valuable self-regulatory skills and build their self-efficacy so as to personally employ these skills effectively (Schunk & Zimmerman, 1997) the manner in which the consequences of the modeled behaviour affect the observer are partially dependent on the observer's self-efficacy (Schunk & Zimmerman, 1997).

The four elements involved in learning by observation of models are attention, retention, production, and motivation (Bandura, 1986). And while in the first place students often require repeated modeling, correcting and practice to successfully complete a task (Schunk, 1998), over time students will be able to negotiate specific tasks with less and less support (Schunk & Zimmerman, 1997).

Self-efficacy and mathematics

For a variety of reasons most of the self-efficacy research in education has been in mathematics. One obvious reason is that the results of mathematical performance are easier to quantify than other areas of education making the subject more research-friendly (Hanchon Graham, 2000). Mathematics is also a major curriculum item in both primary and secondary schools and provides an important assessment tool for a broad range of educational purposes. Two other factors accounting for the popularity of self-efficacy research into mathematics are firstly: the degree of variation in students' self-efficacy beliefs is higher in the mathematics domain (Miller, Greene, Montavalo, Ravindran, & Nichols, 1996). And secondly, the correlation between academic performance and self-efficacy beliefs is also higher in mathematics than any other academic area (Williams, 1994).

There are numerous studies substantiating the value of self-efficacy beliefs as a predictor of students' mathematics performance (for reviews see Bandura, 1986; Pajares, 1996, 1997). And it is well accepted that self-efficacy is a more accurate predictor when it comes to mathematics performance (Hanchon Graham, 2000) than other beliefs such as self-efficacy for self-regulatory practices (Zimmerman, Bandura, & Martinez-Pons, 1990), mathematics self-concept or mathematics anxiety (Pajares & Miller, 1994).

It is significant that in the mathematics domain, self-efficacy has been reported as having a stronger influence on performance than general mental ability (Pajares & Kranzler, 1995) long regarded as the most powerful predictor of academic outcomes (Hunter, 1986). Further, across all levels of ability, students with stronger self-efficacy, mathematically compute more accurately and with greater persistence (Collins, 1982).

Generally speaking social cognitive theorists accept that various other self-beliefs and motivation constructs play an important role in the prediction of mathematics achievement (Hanchon Graham, 2000). For example, self-efficacy for self-regulation (Zimmerman et al., 1990), mathematics anxiety (Cooper & Robinson, 1991), mathematics self-concept (Marsh, 1990), and value of mathematics (Meece et al., 1990) have all been found to predict mathematics-related outcomes.

There is widespread evidence supporting Bandura's contention that self-efficacy beliefs mediate the effect of existing skills on subsequent performance by influencing effort, persistence and perseverance (Schunk, 1991).

Using path analysis, Schunk (1981) demonstrated that modeling treatments elevated student performance and persistence when dealing with division problems by raising self-efficacy and enhancing skills.

Schunk (1982a) showed that effort attributional feedback of earlier performance elevated anticipatory self-efficacy of students and this in turn was partially responsible for their elevated performance in dealing with subtraction problems.

Schunk (1998), in a review of studies investigating the effect of modeling on student performance (e.g., Schunk, 1981, 1982b; Schunk & Gunn, 1986; Schunk & Hanson, 1985), reported that, irrespective of the modeling treatment condition, the confidence students showed in their judgements related in a positive manner to the skills they later demonstrated. From these findings it is apparent that 'students' self-efficacy perceptions are predictive of their subsequent performance' (Hanchon Graham, 2000, p. 44) and students with higher self-efficacy accurately solve more problems.

In a study embracing students of all mathematics ability levels Collins (1982) found that irrespective of student level of ability those with high self-efficacy worked more effectively and persistently when it came to solving problems.

Using path analyses Pajares and Miller (1994) showed that math self-efficacy was a more accurate predictor of students' problem solving ability than math self-concept, perceived usefulness, prior experience or gender. Importantly, these findings demonstrate that self-efficacy beliefs are

reflective of more than the simple sum of one's past experiences (Hanchon Graham, 2000).

Pajares and Kranzler (1995) found that by including a measure of general mental ability in the path model a better control for ability was provided, strengthening the mediational role of self-efficacy. Their findings also support Bandura's (1982, 1990) contention that self-efficacy beliefs play an important role in what one actually does with the ability one has. Probably the most significant finding to be gleaned from this study (Pajares & Kranzler, 1995) is that the direct effect of self-efficacy on performance is as powerful as the effect of general mental ability (Hanchon Graham, 2000).

Hanchon Graham (2000) investigated the influence of various mathematics self-beliefs on mathematics performance to determine if there are changes in these self-beliefs throughout the middle school years. Results from the study support the long-held contention that self-referent thought is a critical component of motivation and achievement.

Some of the earlier studies into this general area have reported little or no relationship between self-efficacy and performance. According to Pajares (1997), however, these studies 'often suffer from problems either in domain specificity or correspondence' (p. 28).

Norwich (1987) found in a path analysis of 9 and 10 year old students where the influence of mathematics self-concept and prior performance was controlled, self-efficacy made no independent contribution in predicting mathematics performance. In referring to this study Hanchon Graham

(2000) points out that 'Norwich entered the variables according to their assumed causal influences from a self-concept perspective, with self-concept entered first, followed by prior performance and by self-efficacy' (p. 14). She then goes on to comment that 'A more impartial test of the influence of self-concept and self-efficacy' would most likely have produced different results.

Benson (1989) reported that the path from mathematics self-efficacy to performance was not significant. Referring to Benson's study Hanchon Graham (2000) commented that it wasn't a genuine assessment of self-efficacy. She argued that using three global items to measure self-efficacy reflected more a performance prediction than a capability judgement.

Wilhite's (1990) finding that students self-assessment of memory ability was the most powerful predictor of their grade point average, followed by locus of control, is flawed according to Hanchon Graham (2000) and Pajares (1997), in that a global self-concept measure not corresponding with the outcome was employed to assess the self-efficacy beliefs.

Cooper and Robinson (1991), using a regression model with mathematics anxiety, the quantitative score on the American College Test (ACT-Q), and prior mathematics experience, revealed that 'self-efficacy did not account for a significant portion of the variance in math performance' (Hanchon Graham, 2000, p. 50). However, as with the aforementioned study by Wilhite, assessment was not consistent with the specifications prescribed by social cognitive theory (Hanchon Graham, 2000). Bandura (1986) has consistently stressed 'ill-defined global measures of perceived self-efficacy or defective assessments of performance will yield discordances' (p. 397).

While general mental ability has long been touted as the most powerful factor in predicting academic performance it can be seen from the literature reviewed herein that a considerable amount of studies have found self-efficacy beliefs to be equally powerful (Pajares & Kranzler, 1995). Towards the end of his chapter *Current Directions in Self-efficacy Research*, Pajares (1997) comments that research findings over the last twenty years have substantiated Bandura's proclamations about the crucial function of self-efficacy in the field of human behaviour. Then, in concluding, Pajares (1997) goes on to urge educators to focus on the self-beliefs students harbour in respect of their academic aptitude because of the important role these beliefs have to play in their 'motivation, self-regulation, and academic achievement' (p. 26).

Conclusions

It is apparent from the preceding chapters that the general province of mathematics teaching in the primary schools of the western world is in a state of crisis and confusion. And while it is obvious we have an enormous problem, it is equally clear we are nowhere near reaching consensus on how it should be addressed. The one point on which agreement can be said to exist is that the sorry state mathematics teaching is currently in, is not due to any lack of requisite intelligence on the part of our students or their innate inability to learn. The vast bulk of the problems associated with student learning it seems can be directly related back to the nature of the curriculum or the method of teaching. It would further seem that these problems are nowhere more apparent than in the transition years.

The research and literature relating to student-directed approaches to learning was examined alongside that relating to Direct Instruction and it was concluded that the empirical data heavily favoured the latter in being more effective. Despite these findings student-directed learning is, and has been for many years, the most popular method employed in the schools of Australia and for that matter, the rest of the western world. Almost every teacher education program in the Australian universities is based on this approach and throughout the US and Canada the same student-directed learning practices are currently being employed under new names such as 'Developmentally appropriate practices' and 'Constructivism'.

The notion that students act on their perceived capability has important implications for classroom practice and this is especially so when endeavouring to understand the problems faced by transition students. If we can determine how students come to estimate their ability at this critical juncture, we will be better equipped to provide interventions aimed at improving their self-efficacy beliefs so they can better exploit their talents and potential.

In this chapter the literature and research relating to the role self-efficacy plays in the learning process was reviewed and the various constructs that come into play in predicting and mediating academic performance were examined with a particular emphasis on mathematics. Analysis here showed that self-efficacy beliefs indeed play a powerful role in the learning process with some studies placing it on the same level as general mental ability, long recognised as the most powerful factor in the prediction of academic performance. Given this important role it seems wise to examine teaching methods that might increase students' self-efficacy and it is

hoped the measured outcomes of this study can make some contribution towards this end. The gathered data from the study will also be analysed in an attempt to gauge the true extent to which these self-efficacy beliefs actually do determine student achievement.

It is a fair assumption that if children are still lacking in basic mathematical skills by the time they reach the transition years then howsoever they have been taught hasn't worked. It is to be hoped that the Direct Instruction instrument utilised in this study will go some way towards helping decide how well suited this approach is for this crucial mathematical domain.

Chapter 4: Methods and instruments

Introduction

The purpose of this research study was to examine student learning in mathematics in the first year of secondary school to determine the effects of a 15-minute teaching intervention. This chapter outlines the researcher's personal perspective then generally describes the gathering of information about student self-efficacy and achievement in mathematics so as to measure the effect of the experimental intervention. Details concerning the procedures and the instruments used to implement the research are outlined, and descriptions of the schools, which participated in the study, are provided. Also discussed in the chapter is the quantitative research method whereby data were collected from students participating in the study at two stages during their first year of secondary school.

A personal perspective on researching mathematics

The researcher's focus on mathematics during students first year of secondary school was influenced by several interrelated factors:

1. A perennial problem teachers face at the start of each school year is the diverse and generally inadequate academic standard of their new class. In an ideal world, teachers could safely assume that the students entering their classroom on the first day of the school year would be capable of performing at an academic level appropriate to

their particular grade and the foundations necessary for them to progress satisfactorily on their academic paths would be firmly in place. In the subject of mathematics at least, all too often this is not the case.

2. Recognition that class sizes were large and there was limited time to prepare Year 6 students for secondary school.
3. A personal belief in the need to establish firm foundations in mathematics in all students, including those at risk, within the regular classroom.
4. Concern about the predominantly negative attitude of upper primary students towards mathematics and their generally low self-efficacy in the subject.

It was whilst working as an upper primary classroom teacher that the researcher, recognising the need for an intervention that addressed the above issues, set about the task of developing one. Because of the diverse academic levels present in almost any given classroom an essential element in the design of the intervention was that it specifically accommodate every student elevating their academic level, no matter what level they started from. To do justice to such a broad-based commission any design blueprint would of necessity have to cast a particularly wide net. Introduced to the University of Oregon Direct Instruction model at Monash University the researcher had been initially taken aback at the conceptual audacity of completely scripted lessons, but was intrigued by the rationale underlying the design and impressed with its attention to

detail. After initial trial implementations of the Oregon model the researcher could actually see the students improving but although the results came more quickly than with other approaches trialed, the program was not without its problems. The long lesson period requiring the undivided attention of both students and teacher, was a major drawback as both found it mentally exhausting. It seemed to the researcher, however, that there was much that could be accomplished within a shorter time frame if the script was radically rewritten. After two years of drafting and trialing it still seemed to the researcher that the underlying principles of the Oregon model were the most appropriate to adopt in the development of the intervention. Another year or so of redrafting and retrialing resulted in a much leaner modified version that sacrificed nothing in terms of content yet could be implemented in less than 20 minutes. This shorter time frame was a big improvement. It gave the students no time to get bored and they could actually see themselves learning as they monitored their progress on a daily basis.

Method of study

Quantitative research has been described as 'a journey of the facts' (Smith, 1983, p. 10), which ideally would be objective and value-free (see Popkewitz, 1984). With the quantitative approach subjectivity should be kept separate from the research enabling the researcher to view the facts objectively (Smith, 1983) — here, 'the researcher's viewpoint is not considered in the explanation of the research' (Hara, 1995, p. 353). In other words the value of the research lies in it being uncontaminated by the researcher's perspective. Choosing the appropriate approach to

research in any given circumstance is a critical factor and because the general purpose of this study was to examine the effects of a teaching intervention actually developed by the researcher, it was considered that the quantitative method, and more specifically a pretest-posttest control group design, would be the best suited. The pretest-posttest control group design consists of two groups where one is given an experimental treatment, and the other is not (Wiersma, 1991). The collection of data on both the control and experimental subjects is carried out immediately prior to administering the experimental intervention and shortly after it has concluded (in this study — 7 months). Pretesting enables the scores to be used as a statistical control, which was important in this study so as to analyse the gain scores in the subjects studied. In acknowledging the practical benefits that may be gleaned from analysis of gain scores, 'especially in instructional areas', Wiersma cautions, 'their meaning and reliability in the specific situation should be considered carefully' (1991, p. 108). Whilst the researcher is in complete concurrence that the variables in any given circumstance must be taken into account it is also true that the 'differences between pretest and posttest measures produced by an intervening experimental treatment or period of instruction can be highly reliable' (Zimmerman & Williams, Table of Contents, 1982).

Procedures

The project submission was approved by the Standing Committee on Ethics in Research Involving Humans (Appendix 1) at a meeting A8/2001 on 4 December 2001 with the provision that the following matter was satisfactorily addressed: If a parent does not consent to their child's results

being used in the research, how will the researcher ensure that the results are not included? The response provided to the Standing Committee (Appendix 2) was approved 10 December (Appendix 3). Permission was also obtained from the Department of Education, Employment and Training (DEET) to conduct research in Government schools (Appendix 4). At the outset of the study the number of subjects to be included was set at 1100.

Like school groups

In 2001, Standards and Accountability (a division of the office of school education, DEET) published school performance benchmarks (Department of Education Employment and Training, 2001) to help schools improve their effectiveness. In Victoria performance levels across the state are identified and presented so that individual schools are able to compare their performance with the performances of peer schools (Department of Education Employment and Training, 2001). These school performance benchmarks are produced for categories of *like* schools across the entire state. Schools have been divided into 9 groups based on the demographic background of their students. The groups are identified by the proportion of students for whom the main language spoken at home is not English (LOTE), and the proportion of students who receive the Education Maintenance Allowance (EMA) or Commonwealth Youth Allowance (YA) (Department of Education Employment and Training, 2001). The like school group boundaries are shown in Figure 2.

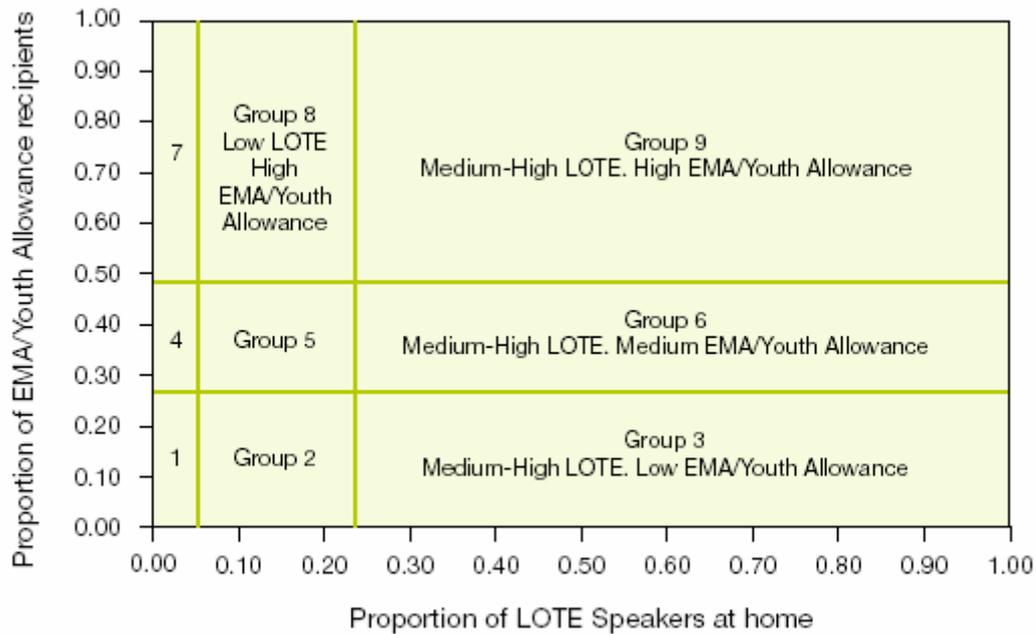


Figure 2 Like school group boundaries

Note. From *VCE Benchmarks 2000* (p. 82), Standards and Accountability Department, Department of Education, Employment and Training, Victoria. Copyright 2001 State of Victoria. Reprinted with permission.

Socioeconomic status

It doesn't seem to matter what country children live in, the wealthier, the better educated, the more upwardly mobile their parents are, the further their education will progress. Conversely children of poorer, less educated, less motivated parents will receive fewer educational opportunities. These are the findings of UNICEF's Innocenti Research Centre. In their November 2002 report they unequivocally conclude, 'educational disadvantage is born not at school but in the home' (UNICEF, 2002, p. 3).

Since at risk students form a subsidiary focus of this study, and because it is well accepted (Rothman, 1983) these students come from socially disadvantaged circumstances, the matter of socioeconomic status was a

relevant consideration in the selection of schools. This, however, proved to be no simple matter to consider for whilst accepting that 'socioeconomic status occupies a central place in educational research', particularly when it comes to 'educational provision for socioeconomically disadvantaged school students', there is 'no consensus regarding its definition and measurement' (Ainley, Graetz, Long, & Batten, 1995, p. 132).

Academic writers have long contended that greater educational benefits flow to the more privileged corners of Australia (Edgar, 1981; Rothman, 1983) and current federal government policy would seem to bear this out. Education Department calculations reported in an article entitled *Give to rich plan sparks outrage* show that 'between 2001-04, 2,600 private schools will receive \$14 billion in Federal grants and the 7,000 state schools just over half that sum' (Maslen, 2000, p. 14). Reporting on the role social status plays in education systems generally, Rothman (1983) comments that 'even though education systems are purported to be open to all students regardless of class or status, some groups repeatedly do better' (p. 76). Australia's indigenous people is one group that does not fare well in the education stakes. Indeed, in commenting on the Howard government's 1998 announcement that it was going to scrap the Abstudy program, Maslen complains that our Aborigines are 'among the most deprived in the world, especially when it comes to education' (2002, p. 10). In discussing the characteristics prevalent in Melbourne social groups that are disadvantaged Rothman points, inter alia, to 'high rates of unemployment' (1983, p. 86) and it has been generally accepted for some time in Australia that the *working class* is amongst the least privileged group (Wild, 1978). The Commission of Inquiry into Poverty (1976) specifically identified low status groups as being at risk in Australian schools.

Since member characteristics were important considerations in this study the researcher considered the purposeful sampling design the most appropriate method to employ. Purposeful sampling occurs where samples are 'selected in a nonrandom manner, based on member characteristics relevant to the research problem' (Wiersma, 1991, p. 428). The Australian Bureau of Statistics (1997) figures were considered in so far as they related to the labour force and unemployment, and schools were selected from suburbs in the Melbourne metropolitan area where the occupational status was primarily manual labour with unemployment rates predominantly between 8 – 19 per cent. The four like school groups (9, 6, 5, 2), representing the best available spread of the different Melbourne school populations, were those chosen by the researcher to participate in the study. Significantly, no schools were selected from groups 1, 4 and 7, where LOTE represents less than 4 per cent.

Two secondary schools were selected from a sampling frame comprised of schools within the four like school groups. These schools satisfied the following conditions:

1. They could implement the experimental intervention across all Year 7 classes in 2002.
2. They were situated in suburbs where the occupational status was predominantly manual labour.

Subsequent to receipt of DEET approval a third secondary school that met the above conditions and which had previously requested to be included in the study was officially approached.

Since the finding of suitable control schools proved difficult the experimental schools were each asked to suggest like schools for the researcher to approach. Fortunately these suggested schools responded positively. In December 2001 all schools participating in the study were officially approached by letter whereby the purpose of the study and the methodology to be employed was explained. All of the Year 7 students in attendance at the schools when Stage 1 data were collected were to be included in the questionnaire and mathematics assessment (with the exception of one experimental school where, because of the researchers financial constraints, only 50% of students were included). The parents of the participating students were all provided with explanatory statements and permission slips (Appendix 5), which were distributed and collected by the schools on behalf of the researcher.

The researcher offered to visit schools and further explain the purpose of the study if necessary. Two of the control schools accepted the offer and the researcher visited them in December 2001. Questionnaire Form 1, mathematics assessment instrument, answer response sheets, express post bags for return of questionnaire and student response forms for Stage 1 of the study were posted to reach each school by February 2002.

Sampling

So that variability across school systems could be better controlled the research design was confined to Government secondary schools. Thus, all schools participating in the study were selected from within the Education Department's Melbourne metropolitan regions. The school ID used to track

schools within the study, the like school group, the metropolitan region and the number of students from each participating school appear in Table 1.

Table 1 School ID, like school group, metropolitan region and number of students

| ID | LSG | Metropolitan region | Students | LOTE | EMA/YA |
|----|-----|---------------------|----------|-----------------|----------------|
| 1C | 5 | Eastern | 36 | 0.04 to <= 0.26 | <0.43 |
| 2C | 9 | Southern | 51 | >0.26 | >0.43 |
| 3E | 6 | Northern | 259 | >0.26 | >0.28 to <0.43 |
| 4C | 5 | Eastern | 108 | 0.04 to <= 0.26 | <0.43 |
| 5E | 9 | Western | 161 | >0.26 | >0.43 |
| 6E | 2 | Western | 99 | 0.04 to <= 0.26 | <0.43 |
| 7C | 9 | Western | 101 | >0.26 | >0.43 |
| 8C | 6 | Western | 152 | >0.26 | >0.28 to <0.43 |

Key: C = control school, E = experimental school, LSG = like school group

The three schools in like school group 9 included in this study are situated in distinctly working class suburban areas in the Southern and Western regions. Unemployment in these suburbs ranges from 15 – 19 per cent. One hundred and fifty-two students (Southern and Western region) are part of the control group and 161 students (Western region) are part of the experimental group. As shown previously in Figure 2, the students in this group are medium-high LOTE (> 0.26) and high EMA/YA (> 0.43).

The two schools in like school group 6 included in this study are situated in largely working class suburban areas in the Northern and Western regions. Unemployment in these suburbs ranges from 8 – 9 per cent. One hundred and fifty-two students (Western region) are part of the control group and 259 students (Northern region) are part of the experimental group. As shown previously in Figure 2, the students in this group are medium-high LOTE (> 0.26) and medium EMA/YA (> 0.28 to < 0.43).

The three remaining schools in this study are either like school group 2 or 5 and are situated in relatively middle-class suburban areas in the Eastern and Western regions. Unemployment in these suburbs ranges from 6 – 7 per cent. One hundred and forty-four students (Eastern region) are part of the control group and 99 students (Western region) are part of the experimental group. As shown previously in Figure 2, the students in this group are low LOTE (0.04 to ≤ 0.26) and low to medium EMA/YA (< 0.43).

The student questionnaire and the mathematics assessment were to be carried out by the mathematics teacher during the first two weeks in February, but because this period proved overly busy with school camp and other activities, the time was extended to include the entire month of February. All student responses to questionnaire and mathematics assessment and parent consent forms were to be returned to the researcher by the end of Term 1. However, due to time constraints it was decided to extend the deadline to allow consent forms to be reissued to those who had not yet returned them.

One control school had a large number of parents objecting to their children participating in the study so an additional school had to be found. Fortunately the Australian Council for Educational Research (ACER), who produces and provides the scoring service for the mathematics assessment instrument used in this study, found an ideal additional school almost immediately. The ACER database was searched for schools, which had already administered the mathematics assessment and were in the same like school group to the control school with the high attrition rate. Two schools were found but one was unsuitable as it was already implementing the experimental intervention. Personal contact was made

with the suitable school, which agreed to participate in the study and gave the researcher permission to use its mathematics assessment data. Consent slips and explanatory statements were given out and collected in second term and the questionnaire was administered.

A total of 8 schools comprising 54 classrooms took part in the study with 967 students having been given permission to participate. As can be seen in Table 2 a total of 445 females and 500 males were identified across the eight schools.

Table 2 Stage 1 male and female student numbers

| School ID | 1C | 2C | 3E | 4C | 5E | 6E | 7C | 8C | Total | Total C | Total E |
|--------------|----|----|-----|-----|-----|----|-----|-----|------------|---------|---------|
| Females | 21 | 30 | 104 | 60 | 84 | 46 | 35 | 65 | 445 | 211 | 234 |
| Males | 15 | 18 | 153 | 45 | 76 | 51 | 62 | 80 | 500 | 220 | 280 |
| Missing data | | 3 | 2 | 3 | 1 | 2 | 4 | 7 | 22 | 17 | 5 |
| Total | 36 | 51 | 259 | 108 | 161 | 99 | 101 | 152 | 967 | 448 | 519 |

Key: C = control school, E = experimental school

Data collection

The teachers implemented the two questionnaires and two mathematics achievement assessments administered during the research period. Students were informed that the basic purpose of the research was to seek their opinions about mathematics as a subject, how they thought they might perform in mathematics generally, and what degree of confidence they had in their mathematical ability. They were informed that their responses would be kept completely confidential. In order to avoid any problems the students might encounter with reading the items, the teacher was required to read the questionnaire aloud. Students were encouraged to inform the teacher if they

were experiencing difficulties in understanding any part of the questionnaire before being asked to record their responses on an optical mark reader (OMR) answer sheet (Appendix 6).

Within a week of completing the questionnaire the students were required to complete a 45-minute mathematics assessment, the Progressive Achievement Test in Mathematics (PATMaths) 2A Revised (Australian Council for Educational Research, 1997). The PATMaths assessment was administered in accordance with the guidelines laid down by the test developer ACER.

Procedures for Stage 2 were similar to Stage 1, the only difference being that the teachers in the experimental schools were invited to complete a small survey (Appendix 7) on the experimental intervention. Data collection for Stage 2 was conducted in October. While the students were completing their mathematics assessment, the teachers in the experimental schools were completing the survey.

Note that the Stage 2 data for one of the control schools (School ID 1) in the study were destroyed in a school fire , thus the variation in the number of control students between Stage 1 and Stage 2.

Instruments

Throughout the study several different instruments were employed: the Student Questionnaires, which will be discussed first, followed by the Short Answer Items then the PATMaths Achievement Test (Australian Council

for Educational Research, 1997) and finally the Experimental Intervention (Farkota, 2000).

Student questionnaires

RAPS-SE, is a student self-report measure — a component of the Research Assessment Package for Schools (Institute for Research and Reform in Education, 1998). This measure, in its earlier version (Wellborn & Connell, 1987), was first used in Australia in a study by Fullarton (1998), where the items were reworded to refer specifically to mathematics. Fullarton piloted the items where alternative wordings were formulated for those items that students or teachers flagged as difficult to understand. Fullarton commented ‘While care was taken to ensure that the original meaning of the question was retained, some questions were reworded to reflect Australian rather than American student vernacular’ (p. 83). The alternative wordings formulated by Fullarton were adopted in this study.

The student questionnaires implemented at Stage 1 and Stage 2 of the study are included in Appendix 8 and Appendix 9. These instruments comprised items concerning attitude towards mathematics generally, beliefs about mathematics, relatedness, perceived control, ongoing engagement, and reactions to challenge. Items concerning how students perceived other people's beliefs about the student's ability were also included. Each of these matters is described in the ensuing sections.

General attitude

The questionnaires begin with six items taken directly from the Fullarton (1998) study. They concern student attitude towards mathematics, perceived importance of mathematics, persistence at and understanding of mathematics, and are ranked on a four point Likert-type scale from very true to not at all true. Examples of such items are: 'I give up working on maths problems when I can't understand them' and 'I usually understand the work we do in maths'.

The next part of the questionnaires is made up of various detailed self-efficacy subscales. As previously discussed in the literature review, self-efficacy has been defined as the belief one has in one's ability to implement the course of action required to achieve a specific outcome (Bandura, 1997).

Perceived control

The Perceived control subdomain, is a component of the Beliefs about self domain of the RAP-SE (Institute for Research and Reform in Education, 1998). On the same four-point Likert scale, students were required to judge statements about the extent to which they thought outcomes in mathematics were due to effort, for example, 'Trying hard is the best way for me to do well in maths', or unknown strategies, for example, 'I don't know how to keep myself from doing badly in maths', and the extent to which they felt they were able to influence these factors.

Task specific confidence

As previously noted in chapter 3, self-efficacy beliefs are of a context specific nature (Schunk, 1991) and Bandura (1977; 1986) has gone to great lengths to stress they be measured at 'the optimal level of specificity' (Hanchon Graham, 2000, p. 30). In other words whatever self-efficacy belief is being assessed it should not only be related to the subject's prior performance of the same or similar exercises, it should be administered as close in time as possible to the prior performance (Bandura, 1997).

To address the context specific nature of the study the researcher developed 17 mathematics items with a view to discerning how confident students felt about responding correctly to these items in a mathematics assessment. The items were similar to those that they would subsequently be presented with in the mathematics assessment. Students were asked not to attempt to solve the problems but simply to provide confidence judgments as to how successful they thought they would be at solving each problem on a four point Likert-type scale that ranged from very confident to not at all confident.

Reaction to challenge

According to the Institute for Research and Reform in Education (1998) in reaction to challenge, 'Students who perceive a situation as challenging, actively persist in the face of failure through the use of effort, strategizing, problem-solving, information seeking, and experimentation'. On the other hand those students who feel threatened 'tend to feel incompetent and full of self-doubt' (p. 3). This is in direct line with Bandura's (1986) contention

that self-efficacy beliefs powerfully influence the choices people make, the amount of effort they expend, the length of time they will persevere in adverse circumstances, and the measure of anxiety or confidence they will bring to a given situation.

Six items measuring how students reacted to challenge were derived from the RAPS-SE (Institute for Research and Reform in Education, 1998) Reaction to challenge subdomain, a separate but interrelated part of the Engagement domain. The scale indicates incidents of academic failure, for example, performing badly in a test or failure to follow teacher explanations, and requires students to rank items on a four point Likert-type scale ranging from very true to not at all true.

Students' reactions to challenge are measured on three dimensions: positive, for example, 'I tell myself that I'll do better next time'; denial, for example, 'I say I didn't care about it anyway'; and anxiety, for example, 'I worry that the other students will think I'm dumb'. According to Bandura (1986) if a student's self-efficacy in mathematics is low this will be coupled with a high degree of math anxiety.

As the RAP-SE (Institute for Research and Reform in Education, 1998) included only one anxiety amplification item, three items indicating student mathematics anxiety were taken from the Fennema-Sherman Mathematics Attitudes Scales (Fennema & Sherman, 1976), which measures feelings of anxiety and associated symptoms with mathematics performance. Items in the scale range from feeling at ease, for example, 'Maths doesn't scare me at all'; to feeling distinct anxiety, for example, 'When I'm in maths classes I

usually feel uncomfortable and nervous', ranked on a four point Likert-type scale ranging from very true to not at all true.

Engagement

As previously discussed in chapter 3, students with high self-efficacy beliefs will persevere longer and harder in difficult circumstances and it is this effort and perseverance, which is called engagement. Conversely, students with low self-efficacy in a specific skill area are loath to engage in tasks where those skills are required and if they do they are more likely to quit when encountering difficulty.

One item indicating a student disinclination to engage, 'I dread having to do maths', was taken from the Hanchon Graham (2000) study.

The items derived from the (Institute for Research and Reform in Education, 1998) Ongoing engagement subdomain, measured the extent of student participation or involvement, for example, 'When I'm in maths classes I usually just act as though I'm working'. Again items were ranked on a four point Likert-type scale ranging from not at all true to very true.

Self-efficacy scale

In the construction of any scale the items need to be calibrated according to their degree of difficulty. The Rasch (see Rasch, 1960, 1980; Wright & Stone, 1979) calibration of a set of items involves the collection of responses from a group of persons and the estimation of the item difficulty

parameters and person ability parameters (J. Barnard, personal communication, 13 September, 2002).

The subscales, General attitude, Perceived control, Reaction to challenge, Task specific confidence and Engagement were subject to Rasch calibration and the scores for the subscales were combined to form a total self-efficacy score, which is described in the following chapter.

Ratings in mathematics

In Stage 1 of the study students were required to rank on a scale of 1 = excellent, to 5 = weak, how good they thought they were at mathematics, how good they would like to be, where their teacher, parents and classmates would place them on the scale, and how good they and their parents would like them to be at mathematics. In Stage 2 of the study students were again asked to rank on a scale of 1 = excellent, to 5 = weak, how good they thought they were in mathematics and where their classmates would place them on the scale. Students were asked to rank how much they liked maths on a scale 1 = Yes very much, to 5 = No. These rating items were taken directly from Fullarton (1998).

Short answer items

At Stage 1 of the study some short answer items were included to extract more detail on students' enjoyment of mathematics, their worries and expectations concerning transition to secondary school, and whether they thought maths classes would be different in secondary school than they

are in primary school. Students were given specific space to answer these items. The short answer items for Stage 1 of the study, which appear in Table 3, were also taken from Fullarton (1998).

Table 3 Stage 1 short answer items

Do you like maths? Explain why or why not.

What do you think will be the best things about going to secondary school?

Is there anything about secondary school that you aren't looking forward to?

What will you miss about primary school?

Do you think that maths classes will be different in secondary school than they are in primary school? If so, how do you think things will be different?

At Stage 2 of the study some short answer items were designed to extract more detail on student feelings about the maths lesson: the part they liked the most; the part from which they thought they learned the most; and the part from which they thought they learned the least. Students were given specific space to answer these items. These short answer items are shown in Table 4.

Table 4 Stage 2 short answer items

Which part of the maths lesson do you like the most?
If you chose OTHER (E) please describe it.

From which part of the maths lesson do you think you have learned the most?
Why do you think this is? If you chose OTHER (E) please describe it.

From which part of the maths lesson do you think you have learned the least?
Why do you think this is? If you chose OTHER (E) please describe it.

Compared to primary school how do you think you are doing in maths?
Why do you think this is?

An additional short answer item, 'Compared to primary school how do you think you are doing in maths? Why do you think this is?' was taken from Fullarton (1998).

Mathematics assessment

The Progressive Achievement Tests in Mathematics Revised (PATMaths Revised) (Australian Council for Educational Research, 1997) was the assessment instrument used in this study to measure student achievement in mathematics over approximately an eight-month period. It was prepared by the Measurement Division of ACER and constitutes an update and revision of the precursor first published by ACER and New Zealand Council for Educational Research in 1984.

There are six separate tests comprising a pair of parallel tests at each of three levels. 'Both forms of the test at each of the three levels of difficulty provide valid and reliable measuring instruments to estimate students' achievement in mathematics' (Australian Council for Educational Research, 1997, Teacher Manual, p. 2). The items all appear in multiple-choice format with each test including a variety of general mathematics topics. The tests specify a 45-minute testing period with additional administration time provided.

In this study Test 2A was used at Stage 1 and Test 2B at Stage 2, which is in line with the year level suitability recommended by ACER. The test items are arranged in content groups using the Mathematics — A Curriculum Profile For Australian Schools, Curriculum Corporation, (1994), also known as the National Profiles. The topic content of the tests used in this study is listed in Table 5.

Table 5 Content structure PATMaths revised tests

| PATMaths Revised Test National Profiles Strand | Question numbers |
|--|------------------------|
| 2A or 2B | |
| Number | 1 to 12 – 12 questions |
| Space | 13 to 19 – 7 questions |
| Measurement | 20 to 26 – 7 questions |
| Chance and Data | 27 to 33 – 7 questions |
| Number (no calculator) | 34 to 39 – 6 questions |

Table 5 illustrates that whilst the emphasis is on the Number strand, some questions are taken from other content strands. Though the Working Mathematically strand is not specifically identified the answers to almost all the questions will require skills from this strand.

A standard answer sheet was provided for all students and the tests were administered by the teachers. The researcher coded all answer sheets before having them machine scanned by the ACER OMR Scoring Service. Data files of the student responses were provided to the researcher by the ACER OMR Scoring Service for data cleaning and scoring. The researcher's data were checked against the ACER OMR Scoring Service results for any discrepancy. Details of the test validity and reliability are given in chapter 5.

Experimental intervention

- Experimental intervention: Elementary Math Mastery (Farkota, 2000)
- Format of intervention: Scripted teacher presentations designed for presentation to entire class
- Number of lessons: Maximum 160

- Scheduled time: 15–20 minutes at the beginning of the regular mathematics lesson
- Weekly schedule: Minimum 4 times per week
- Teacher’s material: Presentation book
- Student material: Exercise book

The experimental intervention is a Direct Instruction mental maths program specifically designed around Mathematics — A Curriculum Profile For Australian Schools (Curriculum Corporation, 1994). Comprising 20 different strands it is a daily program for the entire class requiring 15 minutes to implement, plus 5 minutes for feedback diagnosis and correction procedures.

Each of the experimental intervention's 20 strands (addition; subtraction; multiplication; division; number patterns; equations and inverse operations; whole number properties; fractions; decimals; measurement; space; geometry; average, percentage, ratio, chance; math language; money; time; algebra; visual perception; data analysis; problem solving) starts at base level and moves through its particular field interrelating with the 19 other strands that are being run concurrently (see for example Figure 3).

A number is divisible by 9 if the sum of the digits is divisible by 9.

Question 5. Is 225 divisible by 9?

Number pattern strand

$$\begin{array}{l} 36 \div 9 = 4 \\ 36 \div 4 = 9 \\ 90 \div 9 = 10 \end{array}$$

Equations and inverse operations strand

THIRTY-SIX divided by NINE equals FOUR.
The turn-around fact for this division is THIRTY-SIX divided by FOUR equals NINE.

Question 6. Write the turn-around fact for NINETY divided by NINE equals TEN.

Figure 3 Intervention teacher presentation script lesson 34, number pattern strand, equations and inverse operations strand

NOTE. What the teacher writes on the chalkboard appears in the text box; What the teacher says appears outside the text box; What the teacher says and simultaneously points to on the chalkboard appears in CAPS; What the teacher repeats is underlined.

The daily incremental portions learned by the students in each strand are small, and because they are reinforced and built upon in subsequent lessons (see for example Figure 4), they should be more easily retained (Engelmann, 1980). The teacher models each scripted lesson in the prescribed format with chalkboard presentations being an integral component.

| |
|-----|
| 9 |
| 1 8 |
| 2 7 |
| 3 6 |
| 4 5 |

Lesson 21

When counting in 9s the first five numbers are NINE, EIGHTEEN, TWENTY-SEVEN, THIRTY-SIX, FORTY-FIVE.

Look at the ONES DIGIT of each number. NINE, EIGHT, SEVEN, SIX, FIVE.

Question 5. When counting in 9s what number does the ones digit decrease by?

| |
|-----|
| 9 |
| 1 8 |
| 2 7 |
| 3 6 |
| 4 5 |

Lesson 22

When counting in 9s the first five numbers are NINE, EIGHTEEN, TWENTY-SEVEN, THIRTY-SIX, FORTY-FIVE.

Look at the TENS DIGIT of each number. ONE, TWO, THREE, FOUR

Question 5. When counting in 9s what number does the tens digit increase by?

| |
|-----|
| 9 |
| 1 8 |
| 2 7 |
| 3 6 |
| 4 5 |

Lesson 23

When counting in 9s the first five numbers are NINE, EIGHTEEN, TWENTY-SEVEN, THIRTY-SIX, FORTY-FIVE.

Question 5. When counting in 9s the ones digit decreases by 1 and the tens digit increases by 1. When counting in 9s what number comes just after 45?

Lesson 29

Question 5. When counting in 9s what number comes just before 45 and what number comes just after 45?

Figure 4 Intervention teacher presentation script lessons 21-23 and 29, number pattern strand

NOTE. What the teacher writes on the chalkboard appears in the text box; What the teacher says appears outside the text box; What the teacher says and simultaneously points to on the chalkboard appears in CAPS; What the teacher repeats is underlined.

At the outset the experimental intervention assumes nothing in terms of student academic level. Its aim is to accommodate every student, and to elevate every student's academic level no matter what level they start from (provided, of course, the students do not have significant learning disabilities).

In the early stages the questions are basic. Students with developed skills will find them simple, though in the researcher's experience this does not last long. These basic questions quickly build to questions designed to challenge every student (see for example Figure 5). Of course, the

challenge to some students will be greater than to others, but every student will have been taught the skills necessary to meet that challenge. The experimental intervention has been so designed that students are able to discover for themselves the formulae necessary to solve relatively complex problems automatically and speedily.

Once foundations to the core areas have been laid and tested, they are built on with small precise portions (see for example Figure 4). None of this incremental information is neglected in subsequent lessons. Students move on to questions that gradually increase in complexity (see for example Figure 5), all the while relying on the skills they will have acquired along the way. Questions shift from abstract numbers to real life situations so students will see the relative worth of mathematics in situations that arise every day in the world outside the classroom.

Question 20. A cork and bottle costs \$6.20.
The bottle costs \$5.10 more than the cork.
Find the cost of the cork.

Lesson 24

Question 20. The goal scorers for the
winning soccer team were Barry and Tanielle.
Barry scored half the team 's total plus one
more. Tanielle scored half the remaining total
plus one more. How many goals were scored
altogether?

Lesson 146

Figure 5 Intervention teacher presentation script lesson 24 and lesson 146, problem solving strand

NOTE. What the teacher says appears in black text; What the teacher repeats is underlined.

Everything students are taught will be revisited, developed further and gradually integrated into the whole mathematical scheme. This gradual and consistent development of skills is one of the key elements of the

Direct Instruction model upon which the experimental intervention has been based.

The traditional practice of teaching mathematics in single topics creates many problems for students. Presenting them with a welter of new information at once, expecting them to master it, then move onto another, often unrelated topic, master that too, move on, and so on, is a lot to ask of any student (Engelmann et al., 1991). The problem is compounded when students are not reacquainted with the topics throughout the year. The experimental intervention aims to circumvent this problem by running 20 concurrent strands over the entire school year so that students get familiar with the many connections existing between the various maths disciplines and become fluent in applying them.

Maintaining student focus is crucial to any academic program and this experimental intervention emphatically directs student attention on the learning process. This is achieved by the teacher orally introducing the concepts and questions at a pace easily enough accommodated by students who are concentrating, but which gives them no time to get bored. According to Kinder and Carnine (1991, p. 196) 'rapid pacing keeps students interested' and in addition to the pacing it is the researcher's experience with this experimental intervention that students stay alert because they know a question immediately follows the introduction of a concept. See Appendix 10 for a sample lesson.

Teacher survey

While the students in the experimental schools were completing their maths assessment (Stage 2), the teachers were invited to complete a small survey on the experimental intervention (Appendix 7). This experimental intervention survey questioned teachers as to how they felt about the program generally. For example: What have you found positive about the EMM program? What are your concerns about the EMM program? Would you like to see the program continue next year?

Summary

This chapter outlined the researcher's personal perspective and generally described the quantitative, pretest-posttest design employed in the research. It went on to explain how the study was implemented, and described the instruments and methods utilised throughout. The following chapter will outline the validation and reliability of the instruments.

Chapter 5: Validation of instruments

Introduction

The instruments employed in this study have been described in the previous chapter. In this chapter details concerning the validation of the questionnaires and mathematics assessment instruments will be provided. The self-efficacy scale of the student questionnaires consists of five subscales. As stated in chapter 4, with the exception of the task specific confidence subscale most items making up the self-efficacy scale were extracted from the Fullarton (1998) study. Fullarton who conducted principal components analysis considered this statistical technique particularly important in validating the scales 'because they had not been used with Australian students before' (p. 92). In this study the questionnaire subscales and mathematics assessment items were validated using Rasch measurement (see Rasch, 1960; Rasch, 1980; Wright & Stone, 1979) the results of which are recorded in this chapter.

An important feature of the present study is the Direct Instruction experimental intervention. A search of the literature failed to reveal any other study examining the effects of a Direct Instruction intervention on student self-efficacy in mathematics during the first year of secondary school. A further search supported Fullarton's (1998) claim that no other Australian study had used the RAPS-SE self-report measure. The insights gained from this study will thus be drawn from largely untrodden ground and should prove theoretically and practically useful in the field of mathematics both from a teaching and learning perspective.

Data screening

During each of the two stages of the study the student response sheets for the questionnaire and mathematics assessment were scanned by an optical mark reader (OMR). The scanned data entries provided by the OMR were checked for accuracy. A random sample of 10 per cent of students was chosen and their responses were entered by hand and checked against the scanned data. At each stage, the number of unmatched entries was found to be no more than 1 in 1000 (0.1%). Frequency analysis was performed for each variable and all the variables were within the range of possible scores. The data were then screened for outliers and none were identified.

Missing data

With respect to students' personal details an examination of the data set showed that 9 per cent of students failed to identify their gender on the response sheet section labeled sex. In most cases the researcher was able to identify the gender from the students' name and in cases where there was doubt the school was contacted. When verification was complete student identification details were destroyed in order to preserve anonymity.

It is well accepted that if the missing data were randomly spread throughout the data matrix the likelihood of any serious problems arising is minimal. Although missing data does present problems with data analysis Tabachnick and Fidell (1996) stressed that, 'the pattern of missing data is more important than the amount missing' (p. 60). It is also accepted that

Rasch models for measurement can handle missing data more appropriately than most other approaches. In discussing strategies to handle missing data, Ludlow and O'Leary (2000) claim 'it is a testimony to the fundamental strength of Rasch that the precise choice usually has little impact on the measurement framework' (p. 751). The student questionnaires in this study had very few items of data missing (less than 2%), and they were found to be randomly distributed.

With respect to the mathematics assessment if an item was skipped it was scored as incorrect because it was assumed the student saw its content and decided not to respond. The frequency of missing data was higher for Stage 1 (PATMaths 2A) with 1518 missing responses. This compared to only 494 in Stage 2 (PATMaths 2B). The improvement in the second stage was to be expected given the fact that students' mathematical knowledge can be taken to have grown as the year progressed. Further, in the Stage 1 assessment, 14 students who answered less than 13 of the 39 items were removed from the data set as this provided insufficient information to estimate their abilities. Twelve of these 14 students were from the control group. In the Stage 2 assessment, 4 students were removed. Again these students were from the control group. One student who answered B to every question was also removed.

Note that the Stage 2 data for one of the control schools in the study were destroyed in a school fire, thus the variation in the number of control students between Stage 1 and Stage 2.

Rasch model

Rasch measures in this study are expressed as *student abilities* and *item difficulties* on an interval scale as opposed to raw scores on an ordinal scale. A more comprehensive treatment of the Rasch model is unnecessary for the purposes of this thesis, however, a detailed account is laid out in Rasch (1960; 1980), Wright (1979), Andrich and Masters (1988). The Rasch model (Rasch, 1960, 1980) employed herein proposes a relationship between student ability and item difficulty expressing the relationship as the probability of a certain response. The more able the students the more chance they have to answer the item correctly. With this model the chances of a correct response are a function of the difference between the person's ability and the difficulty of the item (J. Barnard, personal communication, November 29, 2002).

A *logit* is 'the unit of measurement that results when the Rasch model is used to transform raw scores obtained from ordinal data to log odds ratios on a common interval scale. The value of 0.00 logits is routinely allocated to the mean of the item difficulty estimates' (Bond & Fox, 2001, p. 231) and typically estimated values vary between -3 and 3 logits where negative values indicate estimates below the mean and positive values indicate estimates above the mean. An ability or difficulty measure is obtained by converting a raw score percentage into odds of success. For example, a raw score of 30 per cent correct converts to -0.85 logits and a raw score of 80 per cent correct converts to 1.39 logits (J. Barnard, personal communication, November 29, 2002).

Validation of student questionnaires

For the purposes of this study the researcher undertook a Rasch rating scale analysis (see Andrich, 1997; Andrich & Masters, 1988) in order to determine the most appropriate selection of items for calibrating the subscales in the student questionnaires. The items were formed into five subscales: perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics. These five subscales made up the self-efficacy scale. To provide classical indices as well as Rasch based information about the properties of the items the QUEST computer program (Adams & Khoo, 1993) was employed. The results of the indices and the Rasch based information were examined to determine item difficulty and to see whether or not the items fitted the subscales.

In this study thresholds indicate the item difficulty for probability levels of 0.5. An item has average difficulty if the threshold value is 0 and all items are centred around this point. The higher the positive value, the more difficult the item and the higher the negative value, the easier the item. Thresholds will mostly be within the range -2 to 2 logits.

To determine whether or not all the items discriminated in a similar way between students, the infit mean square statistic (INFIT MNSQ) was considered. The acceptable range of the INFIT MNSQ for each item in this study was set from 0.77 to 1.30. Mean square values have an expected value of 1 and individual values above or below this show greater variation (values above 1) or less variation (values less than 1) than might normally be expected. Any value above 1.30 indicates that the item fails to properly discriminate while any item value below 0.77 indicates that the item overly discriminates. In

accord with Rasch theory all items outside the acceptable range were excluded from the scale (Smith & Kramer, 1992). Details of the item thresholds and INFIT MNSQ coefficients with their respective subscale follow.

Perceived control subscale

The 7 items of the perceived control subscale were subjected to Rasch rating scale analysis using QUEST (Adams & Khoo, 1993). The results of the item estimates indicated a mean item difficulty -0.23 (SD $.45$). The reliability of the estimates 0.99 is relatively high so the scores obtained from the perceived control subscale can be regarded as satisfactorily stable.

The item thresholds for the perceived control subscale (see Figure 6) shows that item 15 was the most difficult to endorse, that is, elicit the response *Very true*. This is shown as 15.3 (Note, the decimal shows the category e.g., $.2$ means the second category, that is, the responses *Sort of true* etc.). This is not to say that item 15 was the most difficult item to endorse as a whole in the subscale, but it does show that category 3 in this item was the most difficult to endorse. The easiest item to endorse was item 12. The step between $.3$, $.2$ and $.1$ of each item is reasonable showing there exists a clear distinction between the categories. The threshold values indicated that all the items included in the subscale had a satisfactory spread for the assessment of the different levels of perceived control.

In Figure 6 the x's represent the students' agreeability estimates whilst the numbers on the right refer to the item difficulties and their respective categories. Note that the estimates are shown on a common scale.

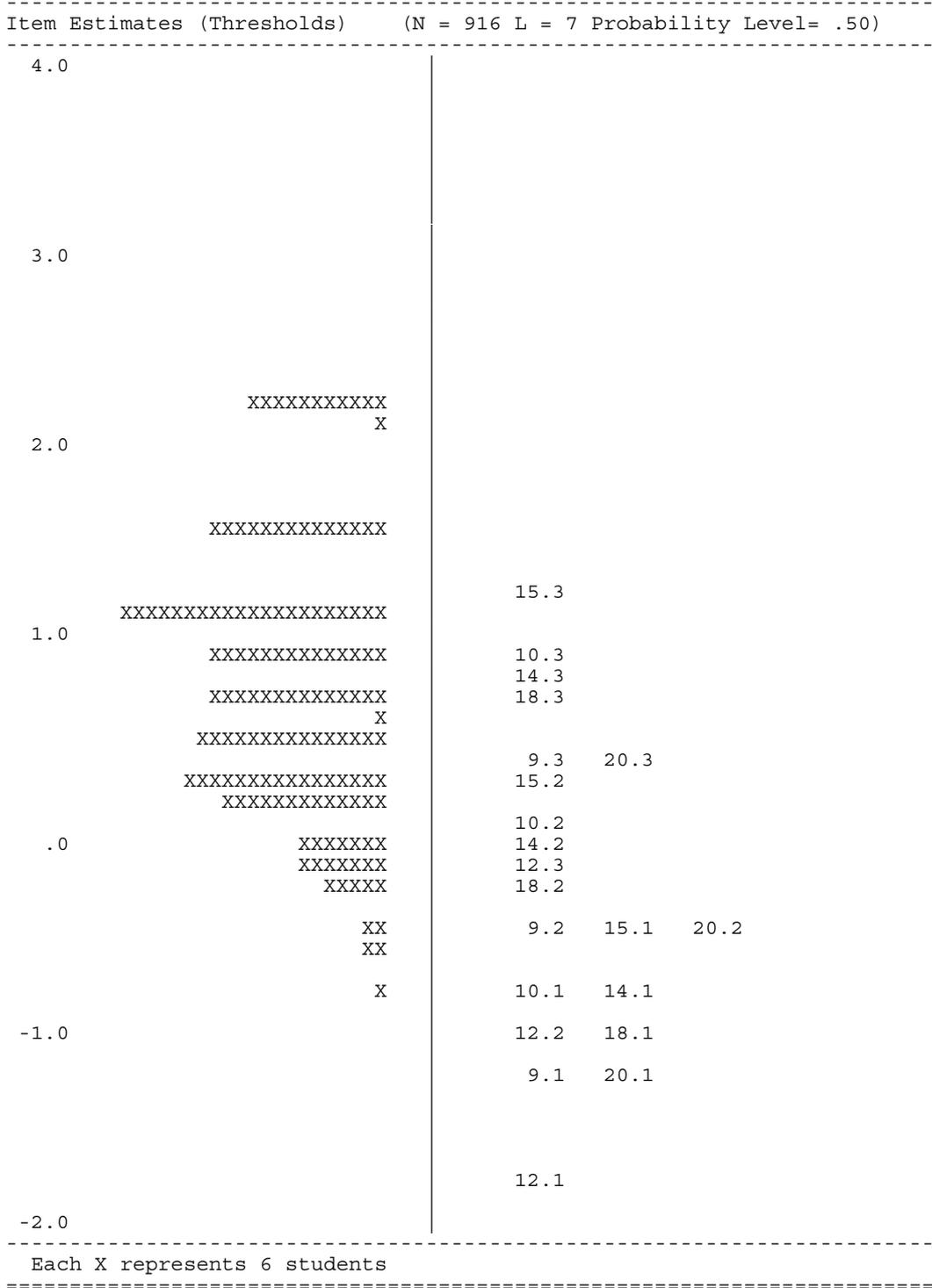


Figure 6 Item thresholds for perceived control subscale

The results of the item fit analysis (see Figure 7) for the perceived control subscale revealed that the 7 items comprising the subscale had INFIT MNSQ coefficients in the range 0.83 to 1.24 which is within the acceptable

range. The subscale mean INFIT MNSQ was 0.98 (SD .14). These results showed that the perceived control items fitted the Rasch model, thus they were all measuring the same trait. This being the case it was considered that the items combined to form an appropriate perceived control subscale. For the purposes of this study the sum of these item scores has been taken to present a total perceived control score.

```

-----
(N = 916 L = 7 Probability Level= .50)
-----
INFIT
MNSQ      .63      .71      .83      1.00      1.20      1.40
-----+-----+-----+-----+-----+-----+
 9 item 9      .          .          .          *          .          .
10 item 10     .          .          *          .          .          .
12 item 12     .          .          .          .          *          .
14 item 14     .          .          *          .          .          .
15 item 15     .          .          .          .          .          *
18 item 18     .          .          .          *          .          .
20 item 20     .          .          .          .          *          .
-----

```

Figure 7 Item fit for perceived control subscale Engagement subscale

The 9 items of the engagement subscale were subjected to Rasch rating scale analysis using QUEST (Adams & Khoo, 1993). The results of the final item estimates indicated a mean item difficulty -0.60 (SD .47). The reliability of the estimates 0.98 is relatively high so the scores obtained from the engagement subscale can be regarded as satisfactorily stable. Note that the mean item difficulty of -0.60 is less than the mean item difficulty of the perceived control subscale of -0.23 . This means that it was generally easier to agree with the statements in the engagement subscale than the statements in the perceived control subscale. This can also be seen in the distributions of the student abilities in Figure 6 and Figure 8. The item thresholds for the engagement subscale (see Figure 8) show there exists a clear distinction between the categories. The threshold values indicated that all the items included in the subscale had a satisfactory spread for the assessment of the different levels of engagement.

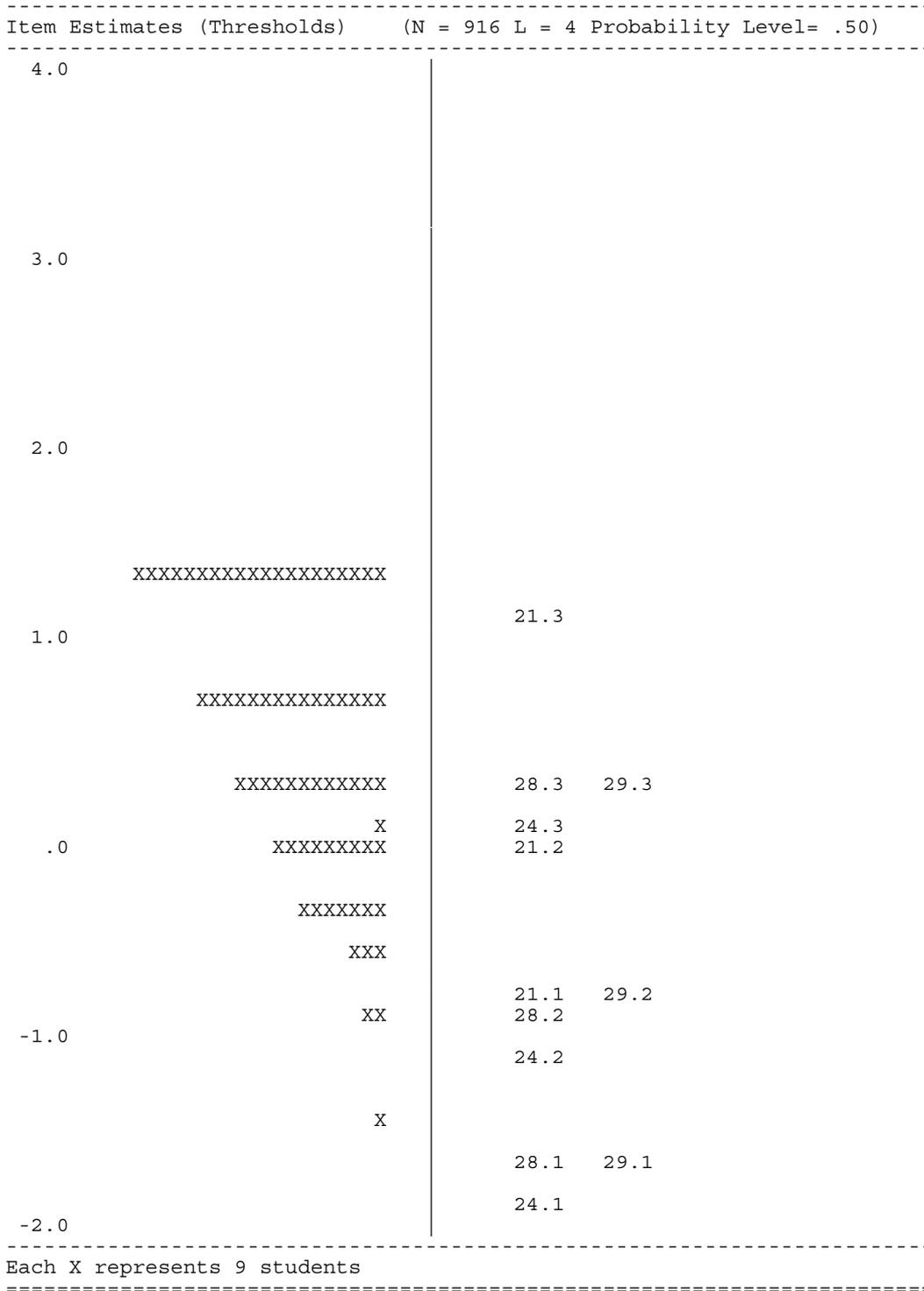


Figure 8 Item thresholds for engagement subscale

The results of the initial item fit analysis for the engagement subscale revealed that 4 of the 9 items had INFIT MNSQ coefficients outside the acceptable range and 1 item was on the lower cutting point 0.77. Further

investigation into these items showed that the poor fit indicated they did not form part of the set, which together defined a single measurement of engagement. Therefore, the 5 items were excluded from the final analysis.

The results of the final item fit analysis (see Figure 9) for the engagement subscale revealed that the 4 items comprising the subscale had INFIT MNSQ coefficients in the range 0.78 to 0.90 which is within the acceptable range. The subscale mean INFIT MNSQ was 0.84 (SD .05). These results showed that the items fitted the Rasch model, thus they were all measuring the same trait. This being the case it was considered the items combined to form an appropriate engagement subscale. For the purposes of this study the sum of these item scores has been taken to present a total engagement score.

```

-----
(N = 916 L = 4 Probability Level= .50)
-----
INFIT
MNSQ      .63      .71      .83      1.00      1.20      1.40
-----+-----+-----+-----+-----+-----+
21 item 21      .      *
24 item 24      . *
28 item 28      .      *
29 item 29      .      *
=====

```

Figure 9 Item fit for engagement subscale

Reaction to challenge subscale

The 10 items of the reaction to challenge subscale were subjected to Rasch rating scale analysis using QUEST (Adams & Khoo, 1993). The results of the final item estimates indicated a mean item difficulty -0.10 (SD .31). The reliability of the estimates 0.98 is relatively high so the scores obtained from the reaction to challenge subscale can be regarded as satisfactorily stable. The item thresholds for the reaction to challenge subscale (see Figure 10) show there exists a clear distinction between the categories. The threshold values indicated that all the items included in the subscale had a satisfactory spread for the assessment of the different levels of reaction to challenge.

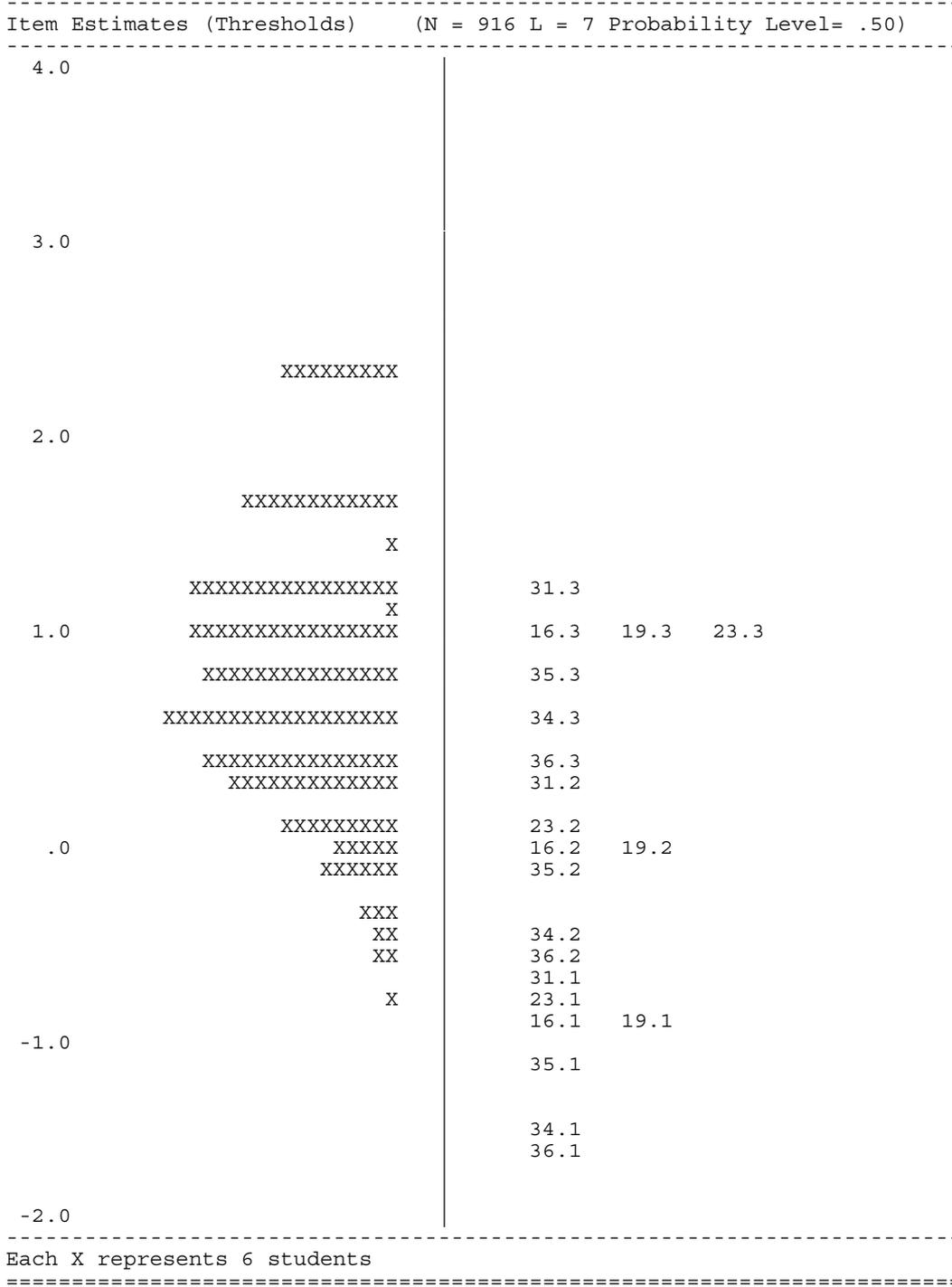


Figure 10 Item thresholds for reaction to challenge subscale

The results of the initial item fit analysis for the reaction to challenge subscale revealed that 2 of the 10 items had INFIT MNSQ coefficients outside the acceptable range. In 1 of these cases the poor fit of the item was below the acceptable range and in the other the poor fit was above the

acceptable range showing significant misfit. Further investigation into these items showed that their poor fits indicated they did not form part of the set, which defined a single measurement of reaction to challenge. One item was close to the upper cutting point 1.30 and demanded further inspection, which revealed the item did not discriminate well. Therefore, the 3 items were excluded from the final analysis.

The results of the final item fit analysis (see Figure 11) for the reaction to challenge subscale revealed that the 7 items comprising the subscale had INFIT MNSQ coefficients in the range 0.80 to 1.02 which is within the acceptable range. The subscale mean INFIT MNSQ was 0.92 (SD .09). These results showed that the items fitted the Rasch model, thus they were all measuring the same trait. This being the case it was considered the items combined to form an appropriate reaction to challenge subscale. For the purposes of this study the sum of these item scores has been taken to present a total reaction to challenge score.

```

-----
(N = 916 L = 7 Probability Level= .50)
-----
INFIT
MNSQ      .63      .71      .83      1.00      1.20      1.40
-----+-----+-----+-----+-----+-----+-----
16 item 16      .              .              .              *              .
19 item 19      .              .              .              *              .
23 item 23      .      *              .              .              .
31 item 31      .              .              .              *              .
34 item 34      .              .              *              .              .
35 item 35      .              .              *              .              .
36 item 36      .              *              .              .              .
=====

```

Figure 11 Item fit for reaction to challenge subscale

Task specific confidence subscale

The 17 items of the task specific confidence subscale were subjected to Rasch rating scale analysis using QUEST (Adams & Khoo, 1993). The results of the final item estimates indicated a mean item difficulty 0.13 (SD .33). The reliability of the estimates 0.99 is relatively high so the scores obtained from the task specific confidence subscale can be regarded as

satisfactorily stable. The item thresholds for task specific confidence subscale (see Figure 12) show there exists a clear distinction between the categories. The threshold values indicated that all the items included in the subscale had a satisfactory spread for the assessment of the different levels of task specific confidence.

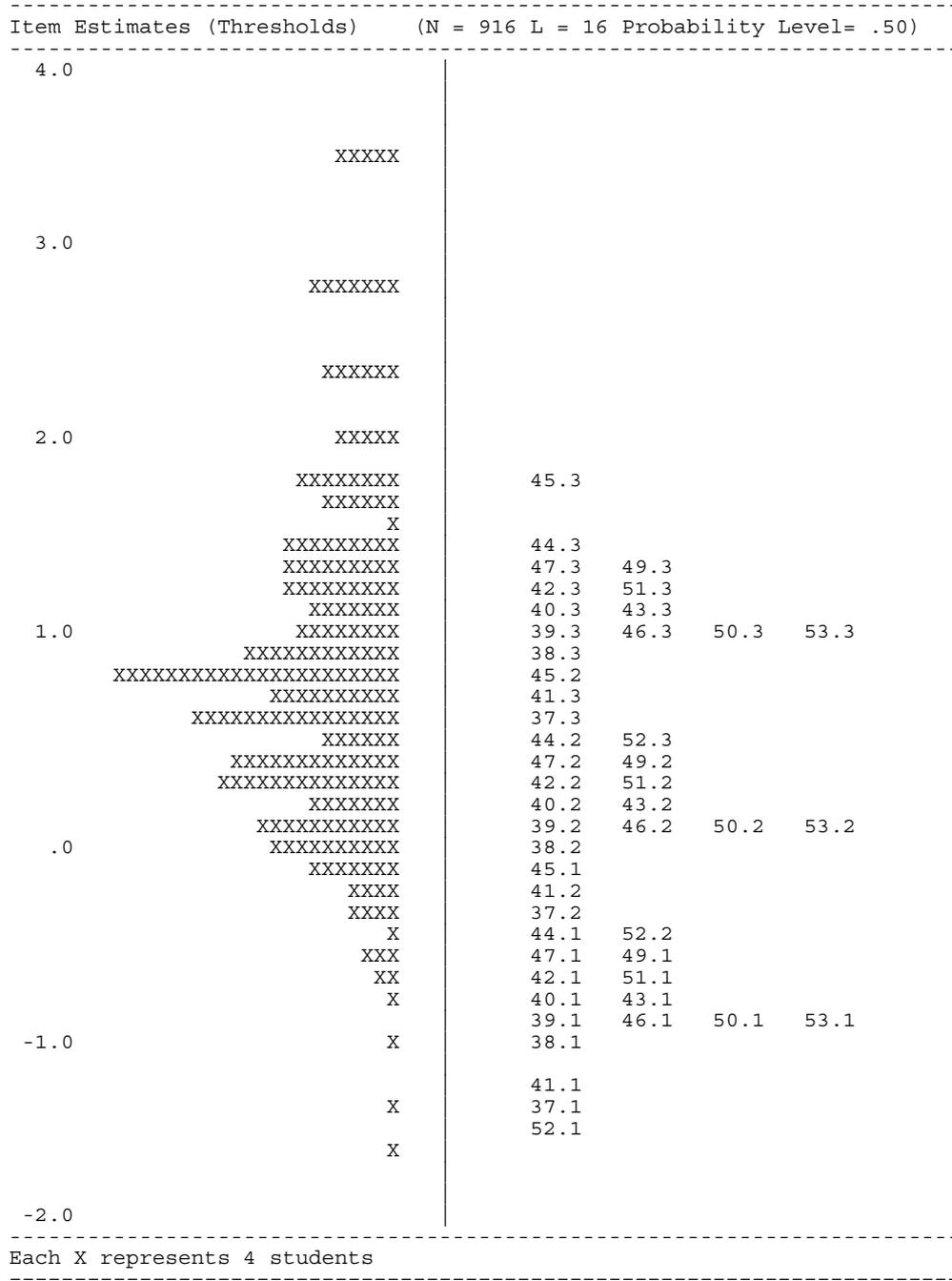


Figure 12 Item thresholds for task specific confidence subscale

The results of the initial item fit analysis for the task specific confidence subscale revealed that 1 of the 17 items had an INFIT MNSQ coefficient outside the acceptable range. In this case the poor fit of the item was just below the lower cutting point 0.77. Further investigation into the poor fit of this item indicated that it required the student to make a judgment about a task beyond their direct experience. Therefore, the item was excluded from the final analysis.

The results of the final item fit analysis (see Figure 13) for the task specific confidence subscale revealed that the 16 items comprising the subscale had INFIT MNSQ coefficients in the range 0.81 to 1.08, which is within the acceptable range. The subscale mean INFIT MNSQ was 0.93 (SD .08). These results showed that the items fitted the Rasch model, thus they were all measuring the same trait. This being the case it was considered the items combined to form an appropriate task specific confidence subscale. For the purposes of this study the sum of these item scores has been taken to present a total task specific confidence score.

```

-----
(N = 916 L = 16 Probability Level= .50)
-----
INFIT
MNSQ      .63      .71      .83      1.00      1.20      1.40
-----+-----+-----+-----+-----+-----+
37 item 37      .      .      *      .      .      .
38 item 38      .      *      .      .      .      .
39 item 39      .      .      *      .      .      .
40 item 40      .      .      .      *      .      .
41 item 41      .      .      .      *      .      .
42 item 42      .      .      *      *      .      .
43 item 43      .      .      .      *      .      .
44 item 44      .      *      .      .      .      .
45 item 45      .      .      .      .      *      .
46 item 46      .      .      .      *      .      .
47 item 47      .      .      *      .      .      .
49 item 49      .      .      .      *      .      .
50 item 50      .      .      .      *      .      .
51 item 51      .      *      .      .      .      .
52 item 52      .      .      .      .      *      .
53 item 53      .      .      .      *      .      .
-----+-----+-----+-----+-----+

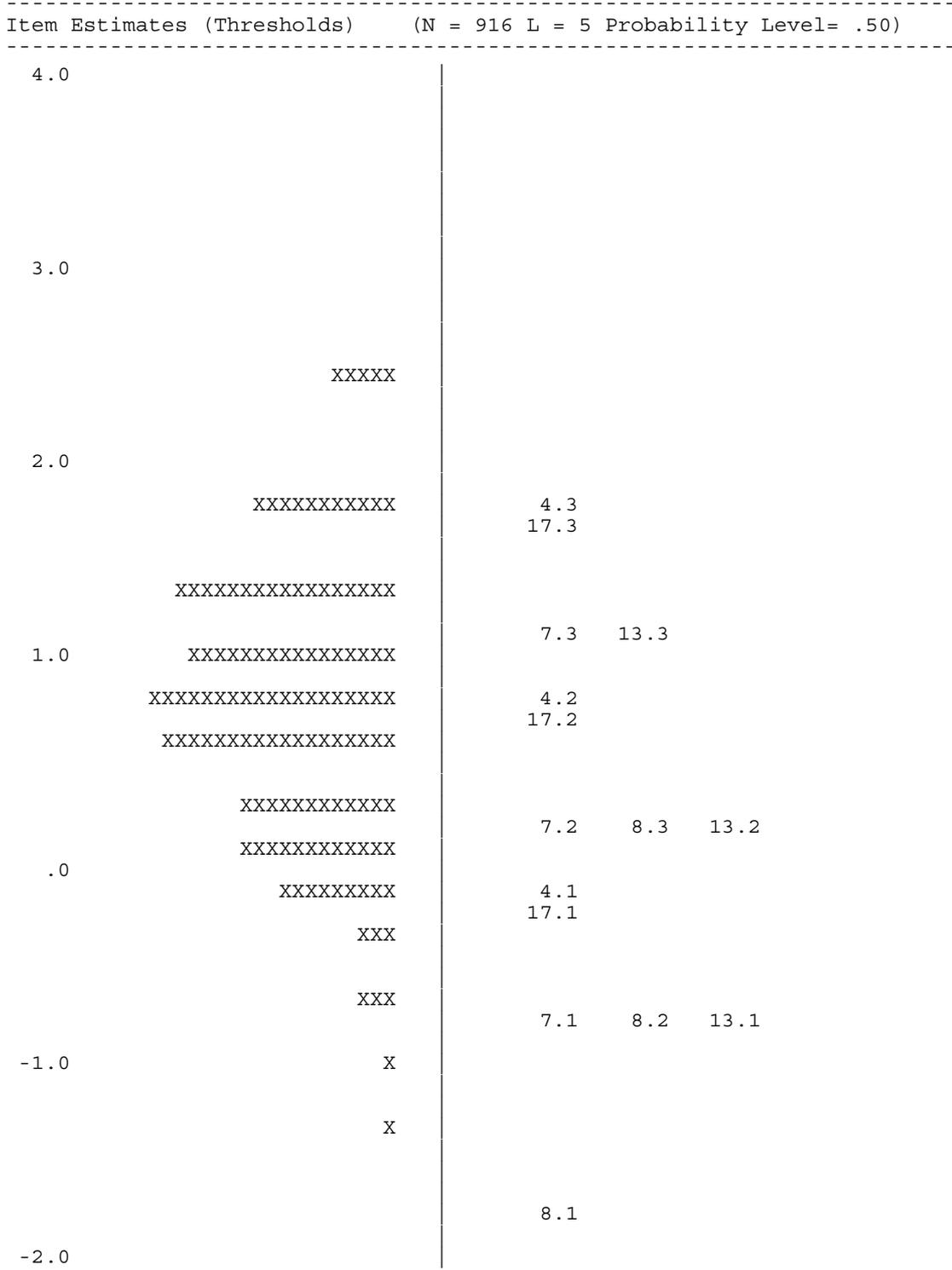
```

Figure 13 Item fit for task specific confidence subscale

General attitude subscale

The 16 items of the general attitude subscale were subjected to Rasch rating scale analysis using QUEST (Adams & Khoo, 1993). Four items were excluded from further analysis as they were not included on the Stage 2 questionnaire. Two further items were excluded from this analysis because they were scored differently on the Stage 2 questionnaire. The results of the final item estimates indicated a mean item difficulty 0.21 (SD .66). The reliability of the estimates 1.00 indicated that the scores obtained from the general attitude subscale can be regarded as stable.

The item thresholds for the general attitude subscale (see Figure 14) show there exists a clear distinction between the categories. The threshold values indicated that all the items included in the subscale had a satisfactory spread for the assessment of the different levels of general attitude.



Each X represents 7 students
=====

Figure 14 Item thresholds for general attitude subscale

The results of the initial item fit analysis for the general attitude subscale revealed that 5 of the remaining 10 items had INFIT MNSQ coefficients outside the acceptable range. Further inspection of these items and their statistics indicated there was one negatively worded statement, which worked better framed in a positive manner. The other items were basically opinions which did not discriminate well so these items were all excluded from the final analysis.

The results of the final item fit analysis (see Figure 15) for the general attitude subscale revealed that the 5 items comprising the subscale had INFIT MNSQ coefficients in the range 0.80 to 1.02 which is within the acceptable range. The subscale mean INFIT MNSQ was 0.90 (SD .09). These results showed that the items fitted the Rasch model, thus they were all measuring the same trait. This being the case it was considered the items combined to form an appropriate general attitude subscale. For the purposes of this study the sum of these item scores has been taken to present a total general attitude score.

```

-----
(N = 916 L = 5 Probability Level= .50)
-----
INFIT
MNSQ      .63      .71      .83      1.00      1.20      1.40
-----+-----+-----+-----+-----+-----
  4 item 4      .      .      .      *      .
  7 item 7      .      *      .      .      .
  8 item 8      .      .      *      .      .
 13 item 13     .      *      .      .      .
 17 item 17     .      .      *      .      .
-----

```

Figure 15 Item fit for general attitude subscale

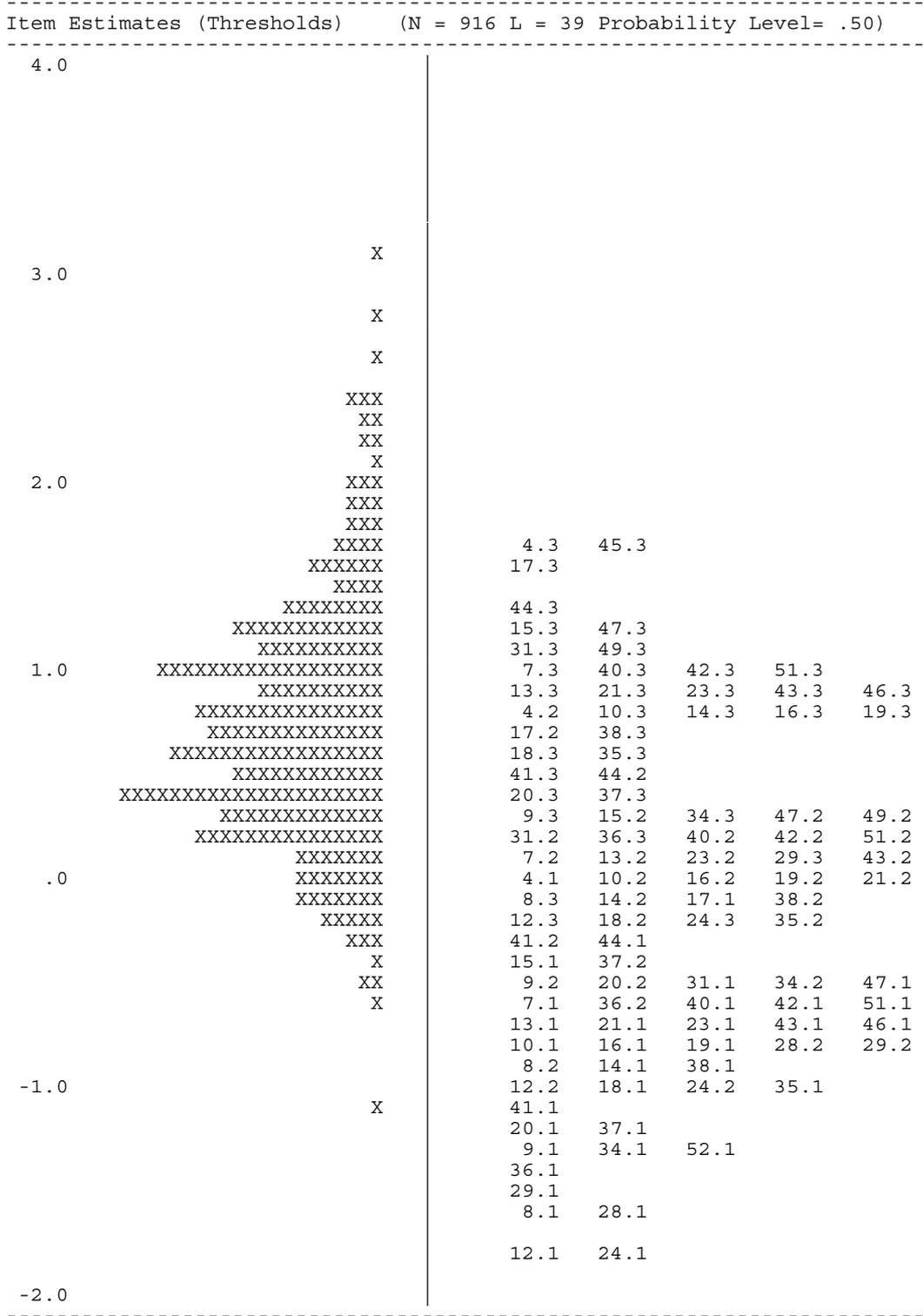
Self-efficacy scale

A total of 39 items and five subscales made up the self-efficacy scale. The mean INFIT MNSQ for the scale was 1.01 (SD .40).

The items in the questionnaires were of the kind *I can do well in maths if I want to*. Agreeability in this study is graded on degrees of agreeability, that is, how much does the student agree with the item. All items were recoded 3,2,1,0 so that the highest degree of student agreeability is 3. Note, some items had a negative construction and were reverse-scored.

The item thresholds for the self-efficacy scale (see Figure 16) show that items 4 and 45 were the most difficult to endorse, that is, they come under the response *Very true*. The easiest items to endorse were items 12 and 24. The steps between .3, .2 and .1 of each item is reasonable showing there exists a clear distinction between the categories.

The threshold values indicated that all the items included in the scale had a satisfactory spread for the assessment of the different levels of self-efficacy. This being the case it was considered the items combined to form an appropriate self-efficacy scale compatible with the Rasch model. For the purposes of this study the sum of these item scores has been taken to present a total self-efficacy score.



Each X represents 4 students
Some thresholds could not be fitted to the display

Figure 16 Item thresholds for self-efficacy scale

Validation of mathematics assessment

The Progressive Achievement Tests in Mathematics Revised (PATMaths) were designed by the Australian Council for Educational Research (1997) to provide a broad general estimate of student achievement in mathematics. These tests were revised on the basis of the Australian National Profiles (Curriculum Corporation, 1994). The revised test items were piloted with approximately 2,000 Victorian students and further refined on information obtained from the piloting.

In order to construct norm-referenced tables, showing percentile ranks and stanines, the tests were used in a 1997 standardization study. Using a stratified sampling scheme the norming sample comprised 100 primary schools and 100 secondary schools from across the whole of Australia. Public, Catholic, and Private schools were chosen proportional to the student population enrolled in each. The sample obtained was modified to ensure a range of school sizes representing both metropolitan and rural schools. At each year level (Year 3-8) the testing was conducted in one class in each participating school. Approximately equal numbers of males and females were represented in each sector and attempted each test form. No information has been provided with respect to the ethnic representation in the norming sample so it is not clear whether the published norms are appropriate for all cultural groups.

Internal reliability estimates are provided for each of the tests ranging from 0.87 to 0.92 and the test developers maintain each contains 'adequate content validity for the purposes for which they were designed' (Australian Council for Educational Research, 1997, p. 14, Teacher Manual).

With the PATMaths Revised tests it is claimed that only items fitting the Rasch model are included, but no statistical evidence of model fit is given to support this. Because of the lack of evidence the researcher has for the purposes of this study subjected the PATMaths revised items to Rasch measurement using QUEST (Adams & Khoo, 1993). The results of the final item estimates indicated a mean item difficulty 0.00 (SD 1.25). A mean of 0 for the item difficulty was obtained because the calibration was standardized on item difficulty. The reliability of the estimates 0.99 is relatively high so the scores obtained from the mathematics assessment instrument can be regarded as satisfactorily stable. The threshold values (see Figure 17) show that all the items included in the scale had a satisfactory spread for the assessment of the different levels of mathematics ability.

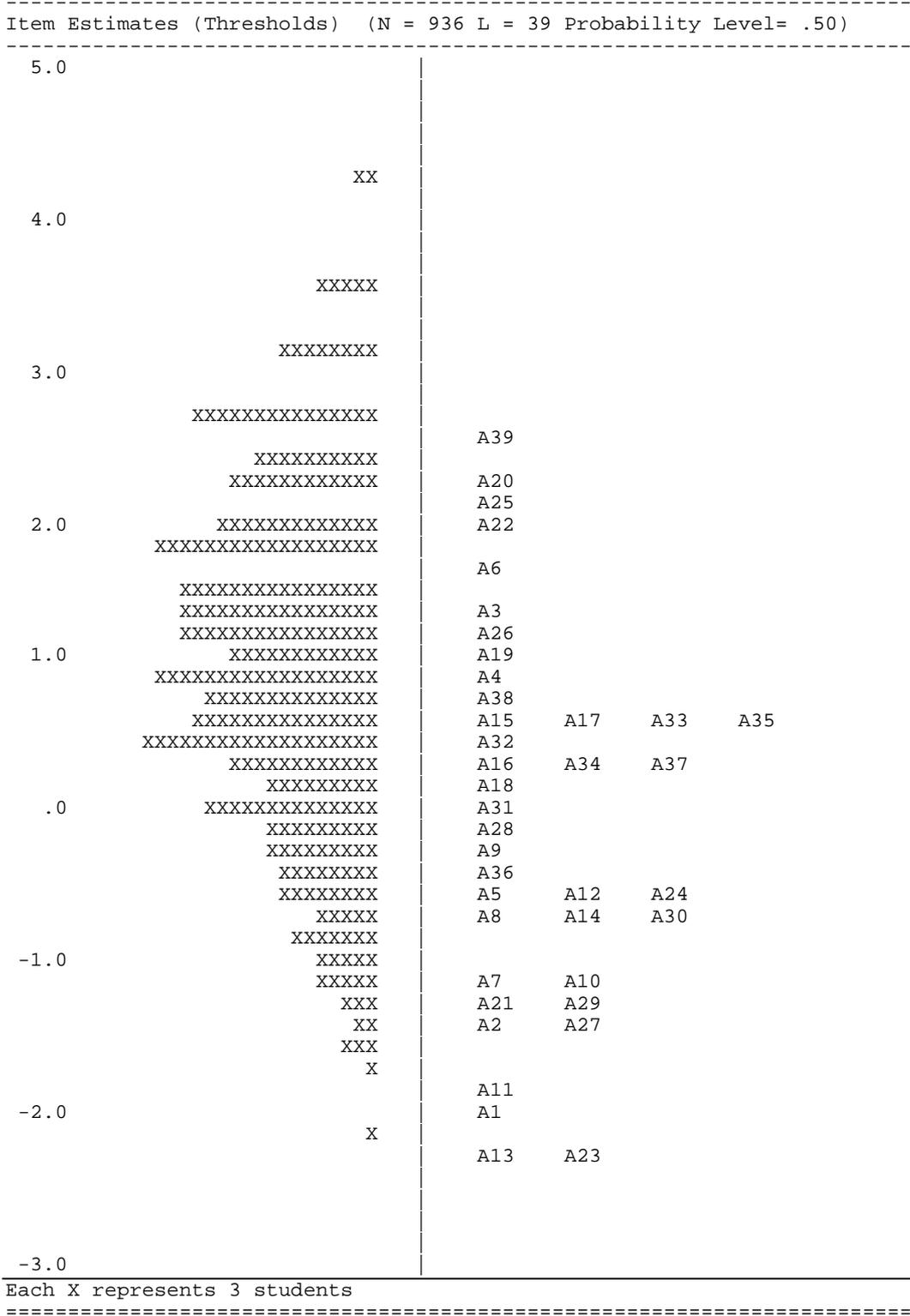


Figure 17 Item thresholds for PATMaths revised

The results of the item fit analysis (see Figure 18) for the PATMaths Revised items revealed that all 39 items had INFIT MNSQ coefficients in the range

0.83 to 1.24 which is within the acceptable range. The mean INFIT MNSQ was 0.99 (SD .09). These results confirmed that the PATMaths Revised items fitted the Rasch model as claimed by the authors ACER, thus were all measuring the same trait. This being the case it was considered the items combined to form an appropriate mathematics assessment instrument. For the purposes of this study the score of each individual student has been taken to present a total mathematics score from which the growth of that student's mathematical achievement has been measured.

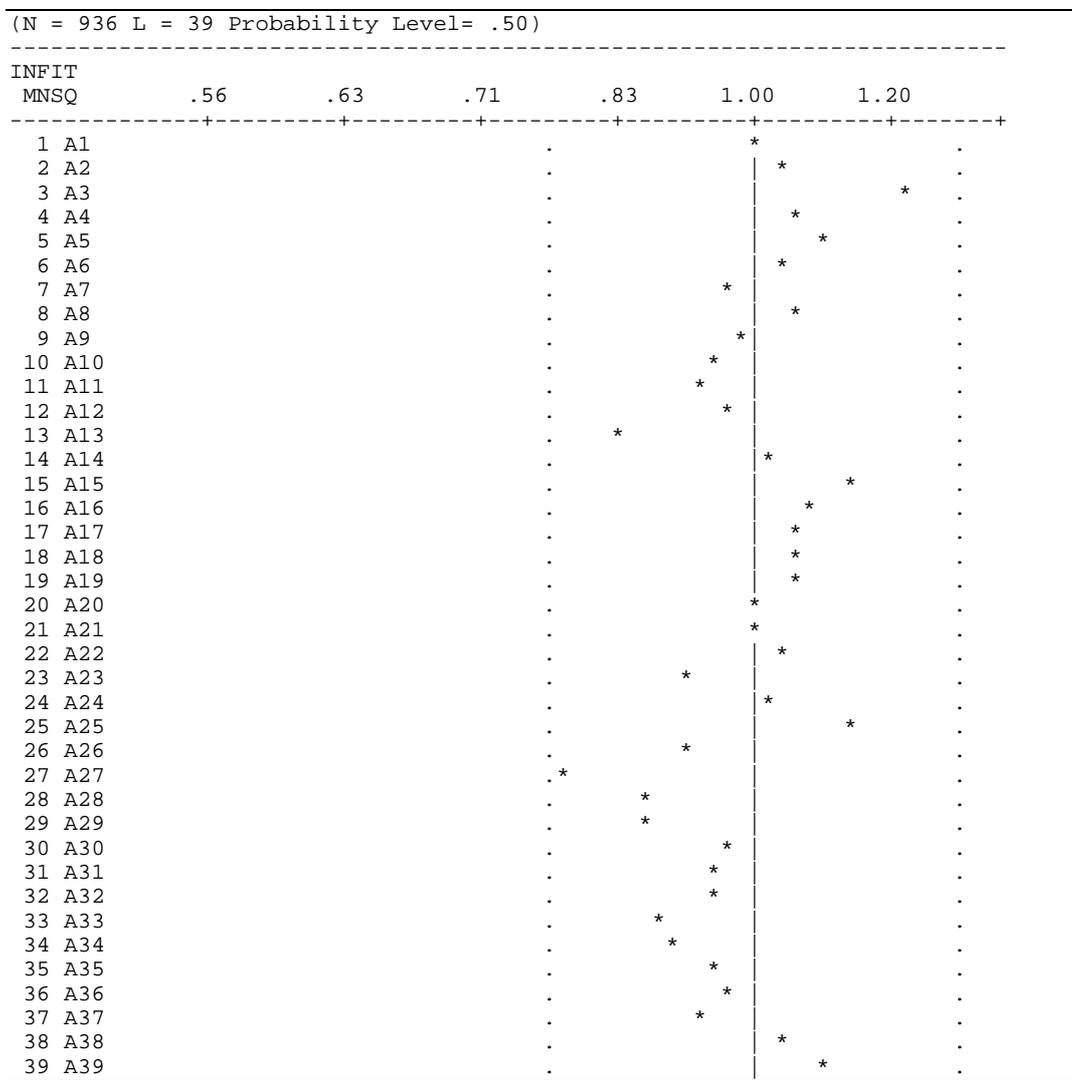


Figure 18 Item fit for PATMaths revised

Summary

This chapter discussed the validation analyses of the questionnaires and mathematics assessment instruments employed in this study. Data checking for unmatched entries was described. No outliers were found. Missing data were examined and seen to be randomly distributed. The reason for the variation in the number of control students between Stage 1 and Stage 2 of the study was explained as being the result of a school fire. Students removed from the study, the scoring of skipped items and the number of missing responses in the mathematics assessment were also explained.

Details of threshold values and the INFIT MNSQ coefficients for each subscale of the questionnaire items were described furnishing evidence that the calibration of the subscales was compatible with Rasch rating scale analysis. Rasch measurement was used to examine the mathematics assessment instrument and the statistics were shown to fit the description given by the test developers (ACER). The questionnaires and mathematics assessment instruments were thus shown to be reliable and valid within the confines of the study.

In the next chapter the data analysis of the questionnaires is examined, and a discussion of students' self-efficacy over the first year of secondary school is included.

Chapter 6: Self-efficacy cross-group comparisons

Introduction

This chapter details and discusses the student questionnaire data analysis from each stage of the study. As the questionnaires were designed to measure five interrelated subdomains of student self-efficacy over the first year of secondary school, data analyses were carried out to examine the effects on the scale and each subscale to measure any change between Stage 1 and Stage 2. Comparisons between the control and experimental group of students are presented as well as gender differences.

Research questions

In examining the effect of the experimental intervention on student self-efficacy the following subsidiary research questions were explored:

1. Do students' self-efficacy beliefs in mathematics change as they move through the first year of secondary school?
2. Does students' perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics change as they move through the first year of secondary school?
3. Do the changes in the self-efficacy beliefs of the students in the control group differ from those in the experimental group?
4. Do the changes that occur in the self-efficacy beliefs of the female students differ from those of the male students?

Statistical methods

Descriptive statistics of the groups for each stage of the study were calculated and t-tests were used for pairwise comparisons. (Note that only two groups are compared and therefore ANOVA was not necessary.) All available data were used in the independent t-tests. For the paired (repeated measures) t-tests only cases for which both scores were available were used. Findings were cross-checked using all available data and similar results were found. Thus, if a significant difference was found between the control and experimental groups using only students for which pre- and posttest data were available, a difference of the same order was also found if all students in each group were included in the computations. Where the degrees of freedom are less than expected (from the number of students) some students were not included for such reasons as, some scores not being available, aberrant response patterns which just introduces large errors, etc.

Analysis of questionnaires: Aims

It seems the first year of secondary school is a crucial developmental stage with respect to students' mathematical self-beliefs and their attitudes towards mathematics generally. As previously noted (see Chapter 3) Bandura (1986) contends that self-efficacy beliefs powerfully influence the choices people make, the amount of effort they expend, the length of time they will persevere in adverse circumstances, and the measure of anxiety or confidence they will bring to a given situation. This is borne out by Fullarton (1998) in her study structured around the perceived control model (see Skinner et al., 1990) where she concludes that 'a large number

of students were found to hold beliefs about learning mathematics that put them at risk of disaffection after transition to secondary school' (1998, p. 219).

The data analysis of the questionnaires in this study investigated whether students' self-efficacy beliefs changed as they proceeded through the first year of secondary school with a view to determining whether the Direct Instruction experimental intervention had any effect on the experimental group's attitude towards mathematics.

Comparison of control and experimental students' on self-efficacy scale

Perhaps the most important single cause of a person's success or failure educationally has to do with the question of what he believes about himself. (Combs, 1921-1999)

In the self-appraisal of efficacy, there are many sources of information that must be processed and weighed through self-referent thought. (Bandura, 1986, p. 21)

From the results obtained from the responses to the items in the student questionnaires a self-efficacy scale was constructed (see Chapter 5). The self-efficacy scale comprised five subscales, namely, perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics. All measures on the scale and its subscales are expressed in logits (see Chapter 5).

The mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the self-

efficacy scale are shown in Table 6. The higher the mean score is, the higher the self-efficacy belief.

Table 6 Descriptive statistics for stage 1 and stage 2, self-efficacy scale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|-----|---------|------|-----|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | .92 | .66 | 318 | .97 | .74 |
| Experimental | 436 | .78 | .65 | 436 | 1.03 | .78 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group's self-efficacy beliefs were generally higher than those of the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was a statistically significant difference in the mean scores for the control and experimental groups at Stage 1 in favour of the control group [$t = 3.273$ for 865 degrees of freedom and $p = .001$ (two-tailed)].

However, at Stage 2 (Table 6) the mean of the experimental group is more than the control group's. Whereas the control group's mean increased from 0.92 to 0.97 the experimental group's mean increased considerably more, from 0.78 to 1.03.

Raw scores usually clump students around the middle scores. However, an improvement near the midpoint is not the same as an improvement towards the upper/lower end. Raw scores assume equal intervals between scores, they assume that an improvement from 50 per cent to 52 per cent is seen as the same as an improvement from say 94 per cent to 96 per cent. However, it is clear that the same improvement along the scale does not reflect the same ability. Instead of expressing performance in terms of raw scores (number correct scores), abilities were computed and expressed in logits as units, which are at the interval level through logarithmic transformations (J. Barnard, personal communication, December 11, 2002).

The control group's mean increased from 0.92 to 0.97 and the experimental groups' mean from 0.78 to 1.03. For the control group the increase is 0.05 logits and for the experimental group it is 0.25 logits. One can obtain some idea of the magnitude of the difference through a transformation: for 0.05 we have $e^{0.05} = 1.05$ which means that the ratio of correct to incorrect responses is 1.05. As percentages this could be 51/49 that is a difference of 2 per cent. For 0.25 we have $e^{0.25} = 1.28$, that is a ratio of 1.28 of correct to incorrect responses. As percentages this could be 56 correct and 44 incorrect, which is a difference of 12 per cent (J. Barnard, personal communication, February 3, 2003). Therefore, the improvement of the experimental group was considerably more than the control group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically

significant difference in the mean scores at Stage 2 for the control and experimental groups [$t = -1.661$ for 807 degrees of freedom and $p = .097$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. Although the mean of the control group increased, the increase was not statistically significant [$t = -1.527$ for 317 degrees of freedom and $p = .128$ two-tailed)]. On the other hand the increase of the experimental group's mean was statistically significant [$t = -8.346$ for 435 degrees of freedom and $p < .001$ (two-tailed)].

The findings can be summarized as follows: The experimental group had statistically significant gains in self-efficacy beliefs in mathematics as they proceeded through the first year of secondary school. The control group had no significant change in self-efficacy beliefs. There was approximately 12 per cent growth in the mean self-efficacy measures for the experimental group and for the control group approximately 2 per cent growth. At Stage 1 the control group had a significantly higher mean whilst at Stage 2 the experimental group had a higher mean showing a particularly significant result overall for the experimental group.

There follows now an examination of the various subscale components that combine to make up the self-efficacy scale.

Comparison of control and experimental students' on perceived control subscale

Among the types of thoughts that affect action, none is more central or pervasive than people's judgements of their capabilities to deal effectively with different realities. (Bandura, 1986, p. 21)

The students most likely to be disaffected are those who essentially feel that they are unable to control success and failure. (Fullarton, 1998, p. 202)

One of the subscales used to measure students' self-efficacy in mathematics was the perceived control subscale. This subscale comprised items developed to measure students': perceptions of control (e.g., I can't do well in maths); beliefs about the type of strategies required to bring about desirable results or avoid undesirable ones (e.g., Trying hard is the best way for me to do well in maths); capability beliefs in respect of enacting strategies (e.g., I can work really hard in maths).

The mean measure and the standard deviation of the distribution for the perceived control subscale were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the perceived control subscale are shown in Table 7. The higher the mean score is, the greater the perceived control.

Table 7 Descriptive statistics for stage 1 and stage 2, perceived control subscale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|-----|---------|------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | 1.03 | .94 | 318 | 1.19 | 1.12 |
| Experimental | 436 | .93 | .98 | 436 | 1.00 | 1.03 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group's perceived control was greater than that of the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1) and there was no statistically significant difference [$t = 1.261$ for 865 degrees of freedom and $p = .207$ (two-tailed)].

Note that at Stage 2 (Table 7) the mean of the control group remained higher than the experimental group's mean. Whereas the control group's mean increased from 1.03 to 1.19 the experimental group's mean increased from 0.93 to 1.00.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant difference in the mean scores for the control and experimental groups [$t = 1.589$ for 807 degrees of freedom and $p = .112$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -2.332$ for 317 degrees of freedom and $p = .021$ (two-tailed)]. Although the mean of the experimental group increased, the increase was not statistically significant [$t = -1.405$ for 435 degrees of freedom and $p = .161$ (two-tailed)].

The findings can be summarized as follows: The control group had statistically significant gains in perceived control in mathematics as they proceeded through the first year of secondary school. The increase in the experimental group was not statistically significant.

Comparison of control and experimental students' on engagement subscale

People tend to avoid engaging in a task where their efficacy is low, and generally undertake tasks where their efficacy is high. (Pajares, 2002)

Students who lack confidence in the skills they possess are less likely to engage in tasks in which those skills are required, and they will more quickly give up in the face of difficulty. (Pajares & Miller, 1997, p. 22)

One of the subscales used to measure students' self-efficacy in mathematics was the engagement subscale. This subscale comprised items developed to measure the extent to which students: exert effort (e.g., I don't try very hard in maths); pay attention (e.g., I try and learn as much as I can about the maths we do).

The mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the engagement subscale are shown in Table 8. The higher the mean score is, the more the students were engaged.

Table 8 Descriptive statistics for stage 1 and stage 2, engagement subscale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|------|---------|------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | 1.23 | 1.15 | 318 | .96 | 1.32 |
| Experimental | 436 | 1.10 | 1.18 | 436 | .98 | 1.38 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group's engagement was higher than that of the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was no statistically significant difference in the mean scores at Stage 1 for the control and experimental groups [$t = 1.717$ for 865 degrees of freedom and $p = .086$ (two-tailed)].

However, at Stage 2 (Table 8) the mean of the experimental group is marginally more than the control group's mean. Whereas the control group's mean decreased from 1.23 to 0.96 the experimental group's mean decreased from 1.10 to 0.98.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant

difference in the mean scores at Stage 2 for the control and experimental groups [$t = -.416$ for 807 degrees of freedom and $p = .677$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The decrease of the control group's mean was statistically significant [$t = 3.783$ for 317 degrees of freedom and $p < .001$ (two-tailed)]. Although the mean of the experimental group decreased, the decrease was not statistically significant [$t = -1.753$ for 435 degrees of freedom and $p = .080$ (two-tailed)].

The findings can be summarized as follows: The control group had a statistically significant decrease in engagement in mathematics as they proceeded through the first year of secondary school. The decrease in the experimental group was small and not statistically significant.

Comparison of control and experimental students' on reaction to challenge subscale

As a general rule, moderate levels of arousal facilitate deployment of skills, whereas high arousal disrupts it. This is especially true of complex activities requiring intricate organization of behavior. (Bandura, 1986, p. 407)

One of the subscales used to measure students' self-efficacy in mathematics was the reaction to challenge subscale. This subscale comprised items developed to measure how students' may deal with certain situations that may be viewed negatively by: playing down their importance (denial e.g., I tell myself it didn't matter anyway); evaluating their behaviour with a view to preventing the same mistake happening

again (positive coping e.g., I try to see where I went wrong); analysing their emotions (e.g., My mind goes blank and I am unable to think clearly).

The mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the reaction to challenge subscale are shown in Table 9. The higher the mean score is, the more positive the reaction to challenge.

Table 9 Descriptive statistics for stage 1 and stage 2, reaction to challenge subscale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|-----|---------|------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | .97 | .92 | 318 | .97 | .96 |
| Experimental | 436 | .94 | .99 | 436 | .97 | 1.02 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group had a higher positive reaction to challenge in mathematics than the experimental group. However, this difference was small.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was no statistically significant difference in the mean scores at Stage 1 for the control and experimental groups [$t = .340$ for 865 degrees of freedom and $p = .734$ (two-tailed)].

However, at Stage 2 (Table 9) the mean of the control group is equivalent to the experimental group's mean. Whereas the control group's mean remained the same the experimental group's mean increased from 0.94 to 0.97.

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The mean of the control group remained the same. Although the mean of the experimental group increased, the increase was not statistically significant [$t = -.713$ for 435 degrees of freedom and $p = .713$ (two-tailed)].

The findings can be summarized as follows: The control group had no change in reaction to challenge in mathematics as they proceeded through the first year of secondary school. The increase in the experimental group was small and not statistically significant.

Comparison of control and experimental students' on task specific confidence subscale

Specific judgements are stronger predictors of the specific performances on which the judgements are based than the broader, less contextual, less task-specific judgements. How could it be otherwise? (Pajares, 1996, p. 563)

Self-efficacy should be assessed at the optimal level of specificity that corresponds to the critical task being assessed and the domain of functioning being analysed. (Pajares, 1996, p. 547)

One of the subscales used to measure students' self-efficacy in mathematics was the task specific confidence subscale. This subscale comprised items

developed to measure students' judgments in respect of their capacity to answer particular mathematics problems correctly. For example: calculate (e.g., Divide 4518 by 9); time (e.g., Express 2.45 pm in 24-hour time); fractions and decimals (e.g., Express 0.001 as a fraction); measurement (e.g., Complete 1.86 m = ? cm); shapes (e.g., How many vertices has a triangular prism); algebra (e.g., Evaluate $3 + 2m$ if $m = 6$); number patterns (e.g., Find the missing number in the pattern 18 27 ? 45 54).

The mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the task specific confidence subscale are shown in Table 10. The higher the mean score is, the more confidence students brought to the task.

Table 10 Descriptive statistics for stage 1 and stage 2, task specific confidence subscale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|------|---------|------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | 1.18 | 1.25 | 318 | 1.47 | 1.43 |
| Experimental | 436 | .85 | 1.03 | 436 | 1.59 | 1.41 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group's task specific confidence was considerably higher than that of the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of

the experimental group (Stage 1). There was a statistically significant difference in the mean scores for the control and experimental groups at Stage 1 in favour of the control group [$t = 3.885$ for 865 degrees of freedom and $p < .000$ (two-tailed)].

However, at Stage 2 (Table 10) the mean of the experimental group is higher than the control group's mean. Whereas the control group's mean increased from 1.18 to 1.47 the experimental group's mean increased from 0.85 to 1.59.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant difference in the mean scores at Stage 2 between the control and experimental groups [$t = -.413$ for 807 degrees of freedom and $p = .680$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The mean of the control group increased and the increase was statistically significant [$t = -4.125$ for 317 degrees of freedom and $p < .001$ (two-tailed)]. The mean of the experimental group increased and the increase was statistically significant [$t = -12.898$ for 435 degrees of freedom and $p < .001$ (two-tailed)].

The findings can be summarized as follows: Though both the control and experimental groups had statistically significant gains in task specific

confidence in mathematics as they proceeded through the first year of secondary school the gain of the experimental group was more.

Comparison of control and experimental students' on general attitude subscale

A number of studies that examined transition in different subject areas found that attitudes towards mathematics declined over the period of transition. (Fullarton, 1998, p. 199)

One of the subscales used to measure students' self-efficacy in mathematics was the general attitude subscale. This subscale comprised items developed to measure students' general attitude towards mathematics for example: perceived importance (e.g., Doing well in maths is important to me); liking of (e.g., I like maths more than I like most other subjects); persistence at (e.g., I give up working on maths problems when I can't understand them).

The mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the general attitude subscale are shown in Table 11. The higher the mean score is, the more positive the general attitude.

Table 11 Descriptive statistics for stage 1 and stage 2, general attitude subscale

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|------|-----|---------|------|-----|
| | N | Mean | SD | N | Mean | SD |
| Control | 318 | .86 | .85 | 318 | .89 | .94 |
| Experimental | 436 | .87 | .86 | 436 | .85 | .83 |

Note. N = paired samples. On analysis of all available data descriptive statistics were similar.

Note that initially (Stage 1) the mean of the experimental group is marginally more than the mean of the control group, indicating that before the experiment the experimental group's general attitude towards mathematics was marginally higher than that of the control group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was no statistically significant difference in the mean scores at Stage 1 for the control and experimental groups [$t = .172$ for 865 degrees of freedom and $p = .863$ (two-tailed)].

However, at Stage 2 (Table 11) the mean of the control group is higher than the experimental group's mean. Whereas the control group's mean increased marginally from 0.86 to 0.89 the experimental group's mean decreased marginally from 0.87 to 0.85.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant difference in the mean scores at Stage 2 for the control and experimental groups [$t = -.006$ for 807 degrees of freedom and $p = .996$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. Although the mean of the control group increased, the increase was not statistically significant [$t = -.642$ for 317 degrees of freedom and $p = .521$ (two-tailed)]. And, although the mean of

the experimental group decreased, the decrease was not statistically significant [$t = .544$ for 435 degrees of freedom and $p = .587$ (two-tailed)].

The findings can be summarized as follows: The control group had no statistically significant increase in general attitude towards mathematics as it proceeded through the first year of secondary school. The decrease in the experimental group was small and not statistically significant.

A summary of the descriptive statistics for the self-efficacy scale and each of its subscales are shown in Table 12.

Table 12 Descriptive statistics for stage 1 and stage 2, self-efficacy scale and subscales

| | Group | Stage 1 | | | Stage 2 | | |
|--------------------------|-------|---------|------|-----|---------|------|-----|
| | | N | Mean | SD | N | Mean | SD |
| Self-efficacy | C | 318 | .92 | .7 | 318 | .97 | .7 |
| | E | 436 | .78 | .7 | 436 | 1.03 | .8 |
| Perceived control | C | 318 | 1.03 | .9 | 318 | 1.19 | 1.1 |
| | E | 436 | .93 | 1.0 | 436 | 1.00 | 1.0 |
| Engagement | C | 318 | 1.23 | 1.2 | 318 | .96 | 1.3 |
| | E | 436 | 1.10 | 1.2 | 436 | .98 | 1.4 |
| Reaction to challenge | C | 318 | .97 | .9 | 318 | .97 | 1.0 |
| | E | 436 | .94 | 1.0 | 436 | .97 | 1.0 |
| Task specific confidence | C | 318 | 1.18 | 1.3 | 318 | 1.47 | 1.4 |
| | E | 436 | .85 | 1.0 | 436 | 1.59 | 1.4 |
| General attitude | C | 318 | .86 | .9 | 318 | .89 | .9 |
| | E | 436 | .87 | .9 | 436 | .85 | .8 |

Note. C = control group, E = experimental group, N = paired samples. On analysis of all available data descriptive statistics were similar.

The findings can be summarized as follows: The experimental group had statistically significant gains in self-efficacy beliefs and task specific

confidence. The control group had statistically significant gains in perceived control, and task specific confidence and a statistically significant decrease in engagement. The control group had no significant change in self-efficacy beliefs. There was approximately 12 per cent growth in the mean self-efficacy measures for the experimental group. For the control group there was approximately 2 per cent growth.

Comparison of control and experimental students' by gender on self-efficacy scale

Although the Third International Mathematics and Science Study (International Association for the Evaluation of Educational Achievement (IEA), 1997) results showed that gender differences in mathematics performance have almost disappeared, according to Hanchon Graham (2000) there is contemporary research claiming that gender differences continue to persist in mathematics confidence amongst students.

This analysis of gender differences focuses on the differences within each group and across groups and only cases for which both scores were available were used.

Female v's male (control group)

The analysis of gender differences for the control group was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males. The results for the scale and each subscale are presented in Table 13.

Table 13 Descriptive statistics for scale and subscale by gender, control group

| Scale/Subscale | SEX | N | Mean diff | SD | Independent-samples t-test | | |
|--------------------------------|-----|-----|-----------|------|----------------------------|-----|-----------------|
| | | | | | t | df | Sig. (2-tailed) |
| Diff: Self-efficacy | F | 156 | .01 | .49 | | | |
| | M | 161 | .08 | .57 | -1.219 | 315 | .224 |
| Diff: Perceived control | F | 156 | .09 | 1.07 | | | |
| | M | 161 | .17 | 1.08 | -.639 | 315 | .523 |
| Diff: Engagement | F | 156 | -.44 | 1.20 | | | |
| | M | 161 | -.10 | 1.33 | -2.423 | 315 | .016 |
| Diff: Reaction to challenge | F | 156 | .03 | .79 | | | |
| | M | 161 | -.03 | 1.17 | .583 | 315 | .560 |
| Diff: Task specific confidence | F | 156 | .21 | 1.22 | | | |
| | M | 161 | .38 | 1.34 | -1.159 | 315 | .248 |
| Diff: General attitude | F | 156 | .05 | .91 | | | |
| | M | 161 | .01 | .90 | .382 | 315 | .703 |

Note. N = paired samples. The mean diff values represent differences in means between stage 1 and stage 2.

The mean change on the male self-efficacy scale is greater than that on the female scale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male self-efficacy beliefs was significant. There was no statistically significant difference in the mean change for females and males [$t = -1.219$ for 315 degrees of freedom and $p = .224$ (two-tailed)].

The mean change on the male perceived control subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female

and male perceived control was significant. There was no statistically significant difference in the mean change for females and males [$t = -.639$ for 315 degrees of freedom and $p = .523$ (two-tailed)].

The mean change on the female engagement subscale is greater than that on the male subscale indicating that although the change is negative for both genders it was more negative in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male engagement was significant. There was a statistically significant difference in the mean change for females and males in favour of the females [$t = -2.423$ for 315 degrees of freedom and $p = .016$ (two-tailed)].

The mean change on the reaction to challenge subscale is equivalent for both females and males. However, the change for females is positive whereas the change for males is negative. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male reaction to challenge was significant. There was no statistically significant difference in the mean change for females and males [$t = .583$ for 315 degrees of freedom and $p = .560$ (two-tailed)].

The mean change on the male task specific confidence subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male task specific competence was significant. There was no statistically significant difference in the mean change for

females and males [$t = -1.159$ for 315 degrees of freedom and $p = .248$ (two-tailed)].

The mean change on the female general attitude subscale is greater than that on the male subscale indicating that more positive change occurred in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male general attitude was significant. There was no statistically significant difference in the mean change for females and males [$t = .382$ for 315 degrees of freedom and $p = .703$ (two-tailed)].

The findings can be summarized as follows: The results of the control group were examined for gender differences. None of the means in self-efficacy, perceived control, reaction to challenge, task specific confidence or general attitude was significantly different by gender. On the other hand, the comparison of engagement by gender within the control group was found to be significant with the females showing more decline in engagement in mathematics as they proceeded through the first year of secondary school.

Female v's male (experimental group)

The analysis of gender differences for the experimental group was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males. The results for the scale and each subscale are presented in Table 14.

Table 14 Descriptive statistics for scale and subscale by gender, experimental group

| Scale/Subscale | SEX | N | Mean diff | SD | Independent-samples t-test | | |
|--------------------------------|-----|-----|-----------|------|----------------------------|-----|-----------------|
| | | | | | t | df | Sig. (2-tailed) |
| Diff: Self-efficacy | F | 200 | .20 | .53 | | | |
| | M | 236 | .29 | .68 | -1.627 | 434 | .105 |
| Diff: Perceived control | F | 200 | .00 | 1.01 | | | |
| | M | 236 | .13 | 1.14 | -1.272 | 434 | .204 |
| Diff: Engagement | F | 200 | -.10 | 1.35 | | | |
| | M | 236 | -.14 | 1.56 | .253 | 434 | .801 |
| Diff: Reaction to challenge | F | 200 | .08 | 1.02 | | | |
| | M | 236 | .01 | 1.16 | .665 | 434 | .507 |
| Diff: Task specific confidence | F | 200 | .59 | 1.06 | | | |
| | M | 236 | .86 | 1.29 | -2.393 | 434 | .017 |
| Diff: General attitude | F | 200 | -.04 | .84 | | | |
| | M | 236 | -.01 | 1.00 | -.296 | 434 | .768 |

Note. N = paired samples. The mean diff values represent differences in means between stage 1 and stage 2.

The mean change on the male self-efficacy scale is greater than that on the female scale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male self-efficacy beliefs was significant. There was no statistically significant difference in the mean change for females and males [$t = -1.627$ for 434 degrees of freedom and $p = .105$ (two-tailed)].

The mean change on the male perceived control subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male perceived control was significant. There was no

statistically significant difference in the mean change for females and males [$t = -1.272$ for 434 degrees of freedom and $p = .204$ (two-tailed)].

The mean change on the male engagement subscale is greater than that on the female subscale indicating that although the change is negative for both genders it was more negative in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male engagement was significant. There was no statistically significant difference in the mean change for females and males [$t = .253$ for 434 degrees of freedom and $p = .801$ (two-tailed)].

The mean change on the female reaction to challenge subscale is greater than that on the male subscale indicating that more positive change occurred in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male reaction to challenge was significant. There was no statistically significant difference in the mean change for females and males [$t = .665$ for 434 degrees of freedom and $p = .507$ (two-tailed)].

The mean change on the male task specific confidence subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male task specific competence was significant. There was a statistically significant difference in the mean change for females and males in favour of the males [$t = -2.393$ for 434 degrees of freedom and $p = .017$ (two-tailed)].

The mean change on the female general attitude subscale is greater than that on the male subscale indicating that more negative change occurred in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male general attitude was significant. There was no statistically significant difference in the mean change for females and males [$t = -.296$ for 434 degrees of freedom and $p = .768$ (two-tailed)].

The findings can be summarized as follows: The results of the experimental group were examined for gender differences. None of the means in self-efficacy, perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On the other hand, the comparison of task specific confidence by gender within the experimental group was found to be significant. Males showed more positive growth than females in task specific confidence in mathematics as they proceeded through the first year of secondary school.

Female control + female experimental v's male control + male experimental

The analysis of gender differences for the combined groups was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males. The results for the scale and each subscale are presented in Table 15.

Table 15 Descriptive statistics for scale and subscale by gender, combined groups

| Scale/Subscale | SEX | N | Mean diff | SD | Independent-samples t-test | | |
|--------------------------------|-----|-----|-----------|------|----------------------------|-----|-----------------|
| | | | | | t | df | Sig. (2-tailed) |
| Diff: Self-efficacy | F | 356 | .11 | .52 | | | |
| | M | 397 | .21 | .65 | -2.158 | 751 | .031 |
| Diff: Perceived control | F | 356 | .04 | 1.04 | | | |
| | M | 397 | .15 | 1.12 | -1.360 | 751 | .174 |
| Diff: Engagement | F | 356 | -.25 | 1.30 | | | |
| | M | 397 | -.12 | 1.47 | -1.285 | 751 | .199 |
| Diff: Reaction to challenge | F | 356 | .06 | .93 | | | |
| | M | 397 | -.01 | 1.16 | .867 | 751 | .386 |
| Diff: Task specific confidence | F | 356 | .43 | 1.15 | | | |
| | M | 397 | .67 | 1.33 | -2.639 | 751 | .008 |
| Diff: General attitude | F | 356 | .00 | .87 | | | |
| | M | 397 | .00 | .96 | .043 | 751 | .966 |

Note. N = paired samples. The mean diff values represent differences in means between stage 1 and stage 2.

The mean change on the male self-efficacy scale is greater than that on the female scale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male self-efficacy beliefs was significant. There was a statistically significant difference in the mean change for females and males in favour of the males [$t = -2.158$ for 751 degrees of freedom and $p = .031$ (two-tailed)].

The mean change on the male perceived control subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male perceived control was significant. There was no

statistically significant difference in the mean change for females and males [$t = -1.360$ for 751 degrees of freedom and $p = .174$ (two-tailed)].

The mean change on the female engagement subscale is greater than that on the male subscale indicating that although the change is negative for both genders it was more negative in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male engagement was significant. There was no statistically significant difference in the mean change for females and males [$t = -1.285$ for 751 degrees of freedom and $p = .199$ (two-tailed)].

The mean change on the female reaction to challenge subscale is greater than that on the male subscale indicating that more change occurred in the females than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male reaction to challenge was significant. Though there was no statistically significant difference in the mean change for females and males, the male change was negative [$t = .867$ for 751 degrees of freedom and $p = .386$ (two-tailed)].

The mean change on the male task specific confidence subscale is greater than that on the female subscale indicating that more positive change occurred in the males than the females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male task specific competence was significant. There was a statistically significant difference in the mean change for

females and males in favour of the males [$t = -2.639$ for 751 degrees of freedom and $p = .008$ (two-tailed)].

The mean change on the female and male general attitude subscale is equivalent. There was no statistically significant difference in the mean for females and males [$t = .043$ for 751 degrees of freedom and $p = .966$ (two-tailed)].

The findings can be summarized as follows: The results of the combined (control + experimental) groups were examined for gender differences. None of the means in perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On the other hand, the comparisons of self-efficacy and task specific confidence for gender within the combined groups were found to be significant. Males showed more positive growth than females in self-efficacy and task specific confidence in mathematics as they proceeded through the first year of secondary school.

Summary

In this chapter the results of the student responses to the questionnaires were examined to investigate whether their self-efficacy beliefs changed as they proceeded through the first year of secondary school.

An independent-samples t-test was conducted on the scale and each of its subscales. Results were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study and the significance of the difference between the two groups was compared.

A paired samples (repeat measures) t-test was conducted on the scale and each of its subscales to evaluate the significance of the differences between Stage 1 and Stage 2 for the control and experimental groups separately.

In summarizing these results the questions outlined at the beginning of this chapter will now be addressed. Further, comparison will be made with the findings of the two most recent studies (Fullarton, 1998; Hanchon Graham, 2000) in the area of mathematics and student self-belief. The focus of the Fullarton study was on engagement with learning — examining perceptions about mathematics and beliefs about learning mathematics over the transition from primary to secondary school. Fullarton's questionnaire examined such things as, students beliefs about mathematics, coping style, engagement and ratings in mathematics.

In the Hanchon Graham (2000) study a major objective was to determine if there are changes in students' mathematics self-beliefs from the beginning to the end of middle school. Independent variables included such things as, mathematics self-efficacy, mathematics anxiety, mathematics self-concept, value of mathematics, engagement in mathematics, and gender. The fact that the study was conducted from the perspective of Bandura's (1986) social cognitive theory made it particularly appropriate to consider here.

Do students' self-efficacy beliefs in mathematics change as they move through the first year of secondary school?

In this study the experimental group had statistically significant positive changes in self-efficacy beliefs in mathematics as they proceeded through the first year of secondary school. The control group had no significant change in self-efficacy beliefs.

In contrast to the present study, Hanchon Graham (2000) found in her research there was a significant overall decline in students' self-beliefs from the beginning to the end of middle school. A similar finding was reported by Fullarton (1998) in her study where she concluded that there was a significant decline in students' self beliefs over transition.

Does students' perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics change as they move through the first year of secondary school?

In this study the control group had statistically significant gains in perceived control in mathematics as they proceeded through the first year of secondary school. The increase in the experimental group was not statistically significant.

In contrast, both the Fullarton (1998) and Hanchon Graham (2000) studies reported a significant decline in student's perceived control from the beginning to the end.

In this study the control group had a statistically significant decrease in engagement in mathematics as they proceeded through the first year of secondary school. The decrease in the experimental group was small and not statistically significant.

Similarly, in both the Fullarton (1998) study and Hanchon Graham (2000) study there was a significant decline in student engagement. Hanchon Graham noted in particular that students tended to put in less effort and their persistence diminished as the study progressed.

In this study the control group had no change in reaction to challenge in mathematics as they proceeded through the first year of secondary school. The increase in the experimental group was small and not statistically significant.

Similarly, in the Hanchon Graham (2000) study it was found that the students' reaction to challenge remained stable.

In the Fullarton (1998) study it was found that generally students' coped positively, for example, they tried to see where they went wrong with a view to remedying their errors, though in contrast with Hanchon Graham's (2000) and the present study, they had become significantly less positive in Year 7.

In this study both the control and experimental groups had statistically significant gains in task specific confidence in mathematics as they proceeded through the first year of secondary school.

Similarly, Hanchon Graham (2000) also found that students tended to be overconfident rather than underconfident. She further reported that research generally has found that students tend to be overconfident about their mathematics capabilities.

Although Fullarton (1998) found that students generally overrated themselves at the outset, she noted they moderated their expectations over the transition.

In this study for the control group there was no statistically significant increase in general attitude towards mathematics as they proceeded through the first year of secondary school. The decrease in the experimental group was small and not statistically significant.

In contrast, Fullarton (1998) found that students' liking and preference for mathematics declined significantly by late Year 6 and this continued. She also reported that female students were significantly less likely to say maths was their favourite subject.

Hanchon Graham (2000) found students generally had a lower interest in mathematics, did not enjoy the subject and regarded it as less important to them. She reported 'a general loss of spirit in the area of mathematics during participants' middle school experience' (p. 125).

Do the changes in the self-efficacy beliefs of the students in the control group differ from those in the experimental group?

In this study there was a significant difference in the mean self-efficacy scores for the control and experimental groups. At Stage 1 the control group had a significantly higher mean whilst at Stage 2 the experimental group had a higher mean showing a considerably significant result overall for the experimental group.

Do the changes that occur in the self-efficacy beliefs of the female students differ from those of the male students?

In this study the results of the control group were examined for gender differences. None of the means in self-efficacy, perceived control, reaction to challenge, task specific confidence or general attitude was significantly different by gender. On the other hand, in the comparison of engagement by gender within the control group there was found to be a significant difference with the females showing more decline in engagement in mathematics as they proceeded through the first year of secondary school.

Though Fullarton (1998) reported gender differences were not significant she found the female students felt more engaged in learning mathematics than male students.

Hanchon Graham (2000) found there were no gender differences in mathematics self-beliefs through the middle school years. She noted that the 'confidence gap' between boys and girls was not apparent at this stage and that changes in school policies, etc. had perhaps closed the 'very real

gap between the way that girls and boys of similar ability rated their confidence in mathematics' (Hanchon Graham, 2000, p. 126).

In this study the results of the experimental group were examined for gender differences. None of the means in self-efficacy, perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On the other hand, the comparison of task specific confidence by gender within the experimental group was found to be significant. Males showed more positive growth than females in task specific confidence in mathematics as they proceeded through the first year of secondary school.

Similarly, Fullarton (1998) found the male students to be more confident of their ability in mathematics than female students, and though she noted both genders moderated their expectations over transition the males retained more confidence. In an earlier study, Fullarton (1993) hypothesized 'that girls who are as competent as boys in mathematics, are not as confident as boys that the answers they obtain for written mathematics tasks are correct' (p. 15). From her analysis of the data, the hypothesis was proven to be true for her sample of 452 Year 7 students from schools in Melbourne's outer eastern suburbs.

As previously stated, Hanchon Graham (2000) found there were no gender differences in mathematics self-beliefs over the middle school years.

In this study the results of the combined groups were examined for gender differences. None of the means in perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On

the other hand, the comparison of self-efficacy and task specific confidence for gender within the combined groups was found to be significant. Males showed more positive growth than females in self-efficacy and task specific confidence in mathematics as they proceeded through the first year of secondary school.

Conclusion

The analysis of the questionnaire data described in this chapter showed there was no significant change in the self-efficacy beliefs of the control group. In contrast, there was a significant gain in the self-efficacy beliefs of the experimental group, strongly suggesting that the experimental intervention was effective.

In the next chapter the data analysis of the mathematics assessment is examined, and a discussion of students' achievement over the first year of secondary school is included.

Chapter 7: Measure of mathematics achievement cross-group comparisons

Introduction

This chapter details and discusses the student mathematics achievement data analysis from each stage of the study. As the mathematics assessment was employed primarily to determine whether there was any growth in student knowledge and understanding of mathematics, data analyses were carried out to examine any difference in performance between Stage 1 and Stage 2. Comparisons between the control and experimental group of students are presented as well as comparisons within like school groups. Gender differences in mathematics achievement are also investigated.

To best examine one of the major research questions (the effect of the experimental intervention on mathematics achievement) in this study it was necessary to:

1. Identify the differences in mathematics achievement of the control group and the experimental group between Stage 1 and Stage 2.
2. Examine the differences in mathematics achievement between the control group and the experimental group.
3. Determine what effect, if any, socioeconomic factors had on mathematics achievement.
4. Determine what effect, if any, gender differences had on mathematics achievement.

Statistical methods

Descriptive statistics of the groups for each stage of the study were calculated and t-tests were used for pairwise comparisons. (Note that only two groups are compared and therefore ANOVA was not necessary.) All available data were used in the independent t-tests. For the paired (repeated measures) t-tests only cases for which both scores were available were used. Findings were cross-checked using all available data and similar results were found. Thus, if a significant difference was found between the control and experimental groups using only students for which pre- and posttest data were available, a difference of the same order was also found if all students in each group were included in the computations. Where the degrees of freedom are less than expected (from the number of students) some students were not included for such reasons as, some scores not being available, aberrant response patterns which just introduces large errors, etc.

Analysis of mathematics achievement: Aims

The effectiveness, or otherwise, of an intervention in any study is best determined by identifying, examining and comparing student achievement over a period of time. Indeed, according to Willett (1997) 'only by measuring individual change is it possible to document each person's progress and, consequently, to evaluate the effectiveness of educational systems' (p. 327).

Though a variety of factors influence the level of student achievement at any point in time, two important ones that will be considered in this study are gender and socioeconomic status.

For more than ten years now there has been much social and academic concern at the 'widening gap between the genders in academic achievement' (Rothman, 2002, p. 38). Male students, it is claimed are much worse off than females, and while this is evidently so with respect to reading comprehension (Lokan, Greenwood, & Creswell, 2001) this does not appear to be the case with mathematics.

One of the key findings in the Programme for International Student Assessment (PISA) was that 'Apart from gender in relationship to reading literacy, the most important student background variable in relation to achievement in Australia was socioeconomic status (SES), based on parents' occupations' (Lokan et al., 2001, p. xii).

Accepting then that both gender and socioeconomic status have an impact on student achievement (see also Chapter 4) like school group comparisons were conducted in order to ensure the validity of the findings. Furthermore, the sample comprised a range of like school groups (2, 5, 6 and 9) drawn from across the Melbourne metropolitan student population so that the results could be deemed representative (see Chapter 4).

This analysis will first compare the control group's achievement scores in mathematics to that of the experimental group's with a view to determining whether students' growth in knowledge and understanding in mathematics has taken place. Next, it will examine the like school group comparisons

followed by an examination of gender differences. Finally, it will be determined what effect if any the experimental intervention has had on the experimental group's achievement.

At each stage of the study, student achievement in mathematics was assessed using 39 items (Australian Council for Educational Research, 1997) addressing, number, space, measurement, and chance and data, at National profile levels 3, 4 and 5 (Curriculum Corporation, 1994). Students were required to choose the correct answer from four or five options (alternatives) in multiple-choice format questions. PATMaths 2A was used in Stage 1 and PATMaths 2B was used in Stage 2. These tests were constructed so as to be parallel both in respect of content coverage and degree of difficulty. Since the tests are parallel at the item level, a raw score on the one form can be compared directly to a raw score on the other (see Australian Council for Educational Research, 1997, Teacher Manual p. 9 where only one table was produced). This strict adherence to the definition of parallel forms makes it possible to compare raw scores to determine growth from Stage 1 to Stage 2.

Reliability refers to the consistency of scores. If a test is reliable a student would expect to attain the same score on two occasions, whilst on an unreliable test, a student's score may vary based on factors that are not related to the purpose of the assessment. Statistical methods (correlations between parallel measures) can be used to establish consistency of student performances within a test or across more than one test. Through such methods an index of reliability can be obtained — commonly referred to as the reliability coefficient. The closer this coefficient is to 1, the higher the reliability.

Kuder and Richardson published a number of formulas, which can be used to estimate the reliability coefficient of a test in one administration. Several formulae, of which formula 20 (KR-20) is the most popular, are widely used to estimate a test's internal consistency. Coefficient values of 0.9 and 0.92 were respectively reported for PATMaths 2A and PATMaths 2B (see Australian Council for Educational Research, 1997, Teacher Manual p. 12). These values indicate high internal consistency for both tests and since the KR-20 coefficients usually underestimate test-retest reliability, it was concluded that these tests are particularly suitable for the purpose of this study.

From the results obtained in the student responses to the items in the mathematics assessment at Stage 1 and Stage 2 of the study, measures of achievement were identified and examined.

Comparison of control and experimental students' mathematics achievement

In terms of curriculum, students should have attained the skills of numeracy and English literacy; such that, every student should be numerate, able to read, write, spell and communicate at an appropriate level. (Ministerial Council on Education Employment Training and Youth Affairs, 1999)

To identify the differences in mathematics achievement the mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results for the mathematics achievement measure are shown in Table 16. The higher the mean score is, the higher the mathematics achievement.

Table 16 Descriptive statistics for stage 1 and stage 2, mathematics achievement measure

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|-------|------|---------|-------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 326 | 26.50 | 7.08 | 326 | 28.90 | 6.67 |
| Experimental | 455 | 24.77 | 6.75 | 455 | 28.73 | 5.64 |

Note. N = paired samples. Mean values represent raw scores.

Note that initially (Stage 1) the mean of the control group is noticeably more than the mean of the experimental group, indicating that before the experiment the control group's mathematics achievement was higher than that of the experimental group. It also indicates that the test was more difficult for the experimental group than for the control group. The standard deviation value for the experimental group was slightly lower than for the control group. The standard deviation value indicates that although trivial the mean scores showed more variability in the control group than in the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was a statistically significant difference in the mean scores for the control and experimental groups at Stage 1 in favour of the control group [$t = 3.505$ for 891 degrees of freedom and $p = .000$ (two-tailed)].

However, at Stage 2 (Table 16) the mean of the control group is only marginally more than the experimental group's. Whereas the control group's mean increased from 26.50 to 28.90 the experimental group's mean increased considerably more, from 24.77 to 28.73. An important

point to note is that the achievement gap between the two groups mean scores would appear to be narrowing. The standard deviation value indicates that although trivial the scores showed more variability in the control group than in the experimental group. Further, the variability between Stage 1 and Stage 2 decreased marginally for the control group and decreased markedly for the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant difference in the mean scores at Stage 2 for the control and experimental groups [$t = -.473$ for 833 degrees of freedom and $p = .636$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -8.384$ for 325 degrees of freedom and $p = .000$ two-tailed)]. The increase of the experimental group's mean was statistically significant [$t = -17.797$ for 454 degrees of freedom and $p = .000$ (two-tailed)].

The results confirm that for both the control and the experimental group, growth in knowledge and understanding in mathematics has taken place. Further, the growth of the experimental group was markedly more.

The findings can be summarized as follows:

Achievement in mathematics improved significantly over the first year of secondary school for both the control group and experimental group. The results of the comparison showed that although the difference in the means for these two groups at Stage 1 were statistically significant in favour of the control group, the difference in the means for the two groups at Stage 2 were trivial and not statistically significant. This revealed a particularly significant result overall for the experimental group. These findings require further investigation in order to identify the reasons why at Stage 1, the experimental group achieved markedly lower scores than the control group.

There follows an examination of the data of the various like school groups that combine to make up the mathematics achievement measure. Four different comparison groups were considered. The first was between the control and experimental students in like school group 9. The second between the control and experimental group in like school group 6. The third between like school group 2 and like school group 5. The fourth between all like school groups excluding the like school groups considered least at risk in this study, that is, like school group 2 and like school group 5. Statistics describing community unemployment and composition of the labour force relevant to the study were obtained from the 1996 census of population and housing, basic community profile, Australia (Australian Bureau of Statistics, 1997).

Comparison of control and experimental students' mathematics achievement: Like-school group 9

Analysis of the data from PISA shows that there is a significant relationship between the results of the student assessments and the student's SES ... and demonstrates that students with lower levels of SES are more likely to have lower achievement levels. (Lokan et al., 2001, p. 162).

This part of the analysis concentrates on like school group 9 participants examining the mean measure of the results of the mathematics assessment, the standard deviation of the distribution, and any changes occurring over the period of the study (Stage 1 – Stage 2).

The inclusion of this group (like school group 9) provides an internal comparison of the students in the lower socioeconomic levels of metropolitan Melbourne, those considered most at risk. The students in like school group 9 are medium-high LOTE (>0.26) and high EMA/Youth Allowance (>0.43). The experimental group has noticeably more LOTE speakers than the control group (see Figure 19, p. 191).

The unemployment rate in the location of the schools in this study in like school group 9 ranges from 15 – 19 per cent. Occupational groups for those employed in the population can be collapsed into 4 major groups: professional/managerial (19%), clerical/sales/service (29%), trades (16%), and production labourers (36%).

The mathematics responses for the mathematics achievement measure were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results are shown in Table 17. The higher the mean score is, the greater the mathematics achievement.

Table 17 Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 9

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|-------|------|---------|-------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 125 | 27.20 | 6.11 | 125 | 29.10 | 6.18 |
| Experimental | 146 | 23.62 | 6.94 | 146 | 28.10 | 6.01 |

Note. N = paired samples. Mean values represent raw scores.

Note that initially (Stage 1) the mean of the control group is considerably more than the mean of the experimental group, indicating that before the experiment the control group's mathematics achievement was higher than that of the experimental group. It also indicates that the test was more difficult for the experimental group than for the control group. The standard deviation value for the experimental group was larger than for the control group. The standard deviation value indicates that the students' scores showed more variability in the experimental group than in the control group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was a statistically significant difference in the mean scores in favour of the control group [$t = 4.470$ for 291 degrees of freedom and $p = .000$ (two-tailed)].

Note that at Stage 2 (Table 17) the mean of the control group remained higher than the experimental group's mean. Whereas the control group's mean increased from 27.20 to 29.10 the experimental group's mean increased more from 23.62 to 28.10. This revealed a particularly significant result overall for the experimental group. An important point to note is that

the achievement gap between the two groups mean scores would appear to be narrowing. The standard deviation value indicates that although trivial the scores showed more variability in the control group than in the experimental group. Further, the variability between Stage 1 and Stage 2 increased for the control group and decreased for the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was no statistically significant difference in the mean scores for the control and experimental groups [$t = .173$ for 282 degrees of freedom and $p = .863$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -4.575$ for 124 degrees of freedom and $p = .000$ (two-tailed)]. The increase of the experimental group's mean was statistically significant [$t = -11.141$ for 145 degrees of freedom and $p = .000$ (two-tailed)].

The results confirm that for both the control and the experimental group, growth in knowledge and understanding in mathematics has taken place. Further, the growth of the experimental group was considerably more.

The findings can be summarized as follows:

Achievement in mathematics improved significantly over the first year of secondary school for both the control and experimental group in like school group 9. The results of the comparison showed that although the

difference in the means for these two groups at Stage 1 were statistically significant in favour of the control group, the difference in the means for the two groups at Stage 2 were not statistically significant. This revealed a particularly significant result overall for the experimental group. The PATMaths mean score difference between Stage 1 and Stage 2 provided a measure of growth confirming that growth in knowledge and understanding in mathematics for the experimental group was more than that of the control group. Further investigation into these findings to identify the reasons why at Stage 1 the experimental group achieved markedly lower scores than the control group revealed that one of the control group schools in like school group 9 included a class comprised exclusively of the school's highest ability students. Note, this school was a late inclusion in the study (see chapter 4) and their sample comprised less than 25 per cent of their Year 7 population; the other 75 per cent did not have parental permission to participate.

Comparison of control and experimental students' mathematics achievement: Like-school group 6

When included together with measures of many other factors in analyses of contextual variables in relation to achievement, SES was still found to be dominant in accounting for differences in scores. (Lokan et al., 2001, p. xii)

This part of the analysis concentrates on like school group 6 participants examining the mean measure of the results of the mathematics assessment, the standard deviation of the distribution, and any changes occurring over the period of the study (Stage 1 – Stage 2).

In relation to the current study this like school group is considered less at risk than the group previously examined (see Figure 19, p. 191). The students in like school group 6 are medium-high LOTE (>0.26) and high EMA/Youth Allowance (>0.28 to <0.43). The control group and the experimental group are comparable in both respects.

The unemployment rate in the location of the schools in this study in like school group 6 ranges from 8 – 9 per cent. Occupational groups for those employed in the population can be collapsed into 4 major groups: professional/managerial (26%), clerical/sales/service (35%), trades (16%), and production labourers (23%).

The mathematics responses for the mathematics achievement measure were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results are shown in Table 18. The higher the mean score is, the greater the mathematics achievement.

Table 18 Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 6

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|-------|------|---------|-------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 116 | 23.72 | 7.99 | 116 | 26.64 | 7.55 |
| Experimental | 235 | 25.08 | 6.55 | 235 | 29.36 | 5.47 |

Note. N = paired samples. Mean values represent raw scores.

Note that initially (Stage 1) the mean of the experimental group is more than the mean of the control group, indicating that before the experiment the experimental group's mathematics achievement was higher than that

of the control group. It also indicates that the test was more difficult for the control group than for the experimental group. The standard deviation value for the control group was larger than for the experimental group. The standard deviation value indicates that the students' scores showed more variability in the control group than in the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1) and there was no statistically significant difference [$t = -1.229$ for 378 degrees of freedom and $p = .220$ (two-tailed)].

Note that at Stage 2 (Table 18) the mean of the experimental group remained higher than the control group's mean. Whereas the control group's mean increased from 23.72 to 26.64 the experimental group's mean increased considerably more from 25.08 to 29.36. An important point to note is that the achievement gap between the two groups mean scores would appear to be widening. The standard deviation value indicates that the scores showed more variability in the control group than in the experimental group. Further, the variability between Stage 1 and Stage 2 decreased for the control group and decreased considerably more for the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was a statistically significant difference in the mean scores for the control and experimental groups in favour of the experimental group [$t = -3.998$ for 371 degrees of freedom and $p = .000$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -4.686$ for 115 degrees of freedom and $p = .000$ (two-tailed)]. The increase of the experimental group's mean was statistically significant [$t = -14.620$ for 234 degrees of freedom and $p = .000$ (two-tailed)].

The results confirm that for both the control and the experimental group, growth in knowledge and understanding in mathematics has taken place. Further, the growth of the experimental group was considerably more.

The findings can be summarized as follows:

Achievement in mathematics improved significantly over the first year of secondary school for both the control and experimental group in like school group 6. The results of the comparison showed that the difference in the means for these two groups at Stage 1 were not statistically significant, but at Stage 2, the difference in the means for the two groups were statistically significant in favour of the experimental group. This revealed that although the mathematics achievement for both groups was significant, the growth in knowledge and understanding in mathematics for the experimental group was considerably more than that of the control group.

Comparison of control and experimental students' mathematics achievement: Like school group 2 and group 5

This part of the analysis concentrates on like school group 2 and like school group 5 participants examining the mean measure of the results of the mathematics assessment, the standard deviation of the distribution, and any changes occurring over the period of the study (Stage 1 – Stage 2).

The inclusion of these groups (like school group 2 and group 5) provides an internal comparison of the students in the middle socioeconomic levels of metropolitan Melbourne, those considered least at risk in this study (see Figure 19, p. 191). The students in like school group 2 (experimental) and group 5 (control) are low LOTE (>0.04 – <0.26) and low to medium EMA/Youth Allowance (<0.43). The experimental group has more LOTE speakers than the control group.

The unemployment rate in the location of the schools in this study in like school group 2 and like school group 5 ranges from 6 – 7 per cent. Occupational groups for those employed in the population can be collapsed into 4 major groups: professional/managerial (43%), clerical/sales/service (34%), trades (12%), and production labourers (11%).

The mathematics responses for the mathematics achievement measure were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study. The results are shown in Table 19. The higher the mean score is, the greater the mathematics achievement.

Table 19 Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 2 (experimental) and group 5 (control)

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|-------|------|---------|-------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 85 | 29.26 | 5.68 | 85 | 31.71 | 6.77 |
| Experimental | 74 | 26.03 | 6.77 | 74 | 27.96 | 5.25 |

Note. N = paired samples. Mean values represent raw scores.

Note that initially (Stage 1) the mean of the control group is considerably more than the mean of the experimental group, indicating that before the experiment the control group's mathematics achievement was much higher than that of the experimental group. It also indicates that the test was more difficult for the experimental group than for the control group. The standard deviation value for the experimental group was larger than for the control group. The standard deviation value indicates that the students' scores showed more variability in the experimental group than in the control group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was no statistically significant difference in the mean scores [$t = 1.809$ for 182 degrees of freedom and $p = .072$ (two-tailed)].

Note that at Stage 2 (Table 19) the mean of the control group remained higher than the experimental group's mean. Whereas the control group's mean increased from 29.26 to 31.71 the experimental group's mean increased from 26.03 to 27.96. The standard deviation value indicates that

the scores showed more variability in the experimental group than in the control group. Further, the variability between Stage 1 and Stage 2 increased considerably for the control group and decreased considerably for the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was a statistically significant difference in the mean scores in favour of the control group [$t = 4.902$ for 176 degrees of freedom and $p = .000$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -7.243$ for 84 degrees of freedom and $p = .000$ (two-tailed)]. The increase of the experimental group's mean was statistically significant [$t = -3.410$ for 73 degrees of freedom and $p = .001$ (two-tailed)].

The results confirm that for both the control and the experimental group, growth in knowledge and understanding in mathematics has taken place. Further, the growth of the control group was markedly more.

The findings can be summarized as follows:

Achievement in mathematics improved significantly over the first year of secondary school for both the control and experimental group in like school group 2 and group 5. The results of the comparison showed that although the difference in the means for these two groups at Stage 1 were

not statistically significant, the difference in the means for the two groups at Stage 2 were statistically significant in favour of the control group. This revealed that although the mathematics achievement for both groups was significant there was a particularly significant result overall for the control group. The mean score difference provided a measure of growth confirming that growth in knowledge and understanding in mathematics for the control group was more than that of the experimental group. It is interesting to observe a different trend on this occasion and that is in relation to the achievement gap between the two groups mean scores widening with the experimental group falling behind. It is worth noting that the post testing in the experimental school was implemented in less than desirable conditions. Due to an internal school misunderstanding the experimental group only agreed to participate in the Stage 2 data collection at the last minute.

The reason the control group achieved markedly higher scores at Stage 1 than the experimental group might be explained by the fact there was only one control school in this particular comparison, and students entering transition year there were traditionally considered above average in mathematics by the school's mathematics coordinator.

The results of the comparisons in this section showed that when compared with all like school groups in the study, the most gains in mathematics were achieved by the experimental group (like school group 9) and that the control students in this group made the least gains in mathematics achievement. Moreover, these like school group 9 students belong to the most at risk group involved in the study. The investigation also revealed that the second to most gains in mathematics were achieved by the

experimental group in like school group 6 and like school group 6 students would be considered the second to most at risk in this study. It is interesting to observe that this trend was not the same for the experimental group least at risk in this study but, as previously stated, the circumstances under which post-testing took place was far from satisfactory. Hence, further comparison excluding the groups least at risk, that is, like school group 2 and group 5, was considered important and is undertaken in the next section.

Comparison of control and experimental students' on mathematics achievement: Like school group 6 and group 9

To identify the differences in mathematics achievement the mean measure and the standard deviation of the distribution were computed separately for the students in the control and experimental like school groups 6 and 9 for Stage 1 and Stage 2 of the study. The results for the mathematics achievement measure are shown in Table 20. The higher the mean score is, the higher the mathematics achievement.

Table 20 Descriptive statistics for stage 1 and stage 2, mathematics achievement measure like school group 6 and group 9

| Group | Stage 1 | | | Stage 2 | | |
|--------------|---------|-------|------|---------|-------|------|
| | N | Mean | SD | N | Mean | SD |
| Control | 241 | 25.53 | 7.28 | 241 | 27.91 | 6.98 |
| Experimental | 381 | 24.52 | 6.72 | 381 | 28.88 | 5.71 |

Note. N = paired samples. Mean values represent raw scores.

Note that initially (Stage 1) the mean of the control group is more than the mean of the experimental group, indicating that before the experiment the control group's mathematics achievement was higher than that of the experimental group. It is interesting to observe that the difference between the mean of the control group and that of the experimental group is less than any of the comparisons conducted so far in this chapter. Hence, these groups appear to be well matched. The standard deviation value for the control group was slightly more than for the experimental group. The standard deviation value indicates that although trivial the mean scores showed more variability in the control group than in the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 1). There was a statistically significant difference in the mean scores for the control and experimental groups at Stage 1 in favour of the control group [$t = 2.104$ for 671 degrees of freedom and $p = .036$ (two-tailed)].

However, at Stage 2 (Table 20) the mean of the experimental group is more than the control group's. Whereas the control group's mean increased from 25.53 to 27.91 the experimental group's mean increased considerably more, from 24.52 to 28.88. An important point to note is that the achievement gap between the two groups mean scores would appear to have widened. The standard deviation value indicates that the scores showed more variability in the control group than in the experimental group. Further, the variability between Stage 1 and Stage 2 decreased marginally for the control group and decreased more for the experimental group.

An independent-samples t-test was conducted to compare the significance of the difference between the mean of the control group with the mean of the experimental group (Stage 2). There was a statistically significant difference in the mean scores at Stage 2 in favour of the experimental group [$t = -2.804$ for 655 degrees of freedom and $p = .005$ (two-tailed)].

A paired samples (repeated measures) t-test was conducted to evaluate the significance of the differences between Stage 1 and Stage 2 for each of the control and experimental groups separately. The increase of the control group's mean was statistically significant [$t = -6.464$ for 240 degrees of freedom and $p = .000$ two-tailed)]. The increase of the experimental group's mean was statistically significant [$t = -18.374$ for 380 degrees of freedom and $p = .000$ (two-tailed)].

The results confirm that for both the control and the experimental group, growth in knowledge and understanding in mathematics has taken place. Further, the growth of the experimental group was markedly more.

The findings can be summarized as follows:

Achievement in mathematics improved significantly over the first year of secondary school for both the control and experimental like school group 9 and group 6. It was observed that the difference between the mean score of the control group and that of the experimental group was less than any of the comparisons reported so far in this chapter. The results of the comparison showed that although the difference in the means for these two groups at Stage 1 were statistically significant in favour of the control group, the difference in the means for the two groups at Stage 2 were

statistically significant in favour of the experimental group. This revealed that although the mathematics achievement for both groups was significant there was a particularly significant result overall for the experimental group. The mean score difference provided a measure of growth confirming that growth in knowledge and understanding in mathematics for the experimental group was more than that of the control group.

Comparison of control and experimental students' mathematics achievement by gender

It is particularly interesting to consider ... results by gender for two reasons. The first is the progress made towards gender equity in mathematics The second reason is the recent concern and topical debate in Australia about a decline in boys' achievement in many academic areas relative to that of girls. (Lokan et al., 2001, p. 33)

This analysis of gender differences focuses on the differences across groups (that is the combined control and experimental group), and within each group individually.

Comparisons of mathematics achievement: Male v's female

The analysis of gender differences across both the control and experimental groups for the students who participated in both Stage 1 and Stage 2 of the study was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males. The results are presented in Table 21.

Table 21 Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender

| SEX | N | Mean diff | SD |
|---------|-----|-----------|------|
| Females | 365 | 3.36 | 5.03 |
| Males | 415 | 3.25 | 4.96 |

Note. N = paired samples.

The mean change on the female mathematics achievement measure is marginally more than that on the male mathematics achievement measure indicating that the females achieved marginally more growth in mathematics than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male mathematics achievement was significant. There was no statistically significant difference in the mean change for females and males [$t = .324$ for 778 degrees of freedom and $p = .746$ (two-tailed)].

The analysis of gender differences for within the control group was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males. The results are presented in Table 22.

Table 22 Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender, control group

| SEX | N | Mean diff | SD |
|---------|-----|-----------|------|
| Females | 162 | 2.84 | 5.38 |
| Males | 168 | 2.20 | 5.14 |

Note. N = paired samples.

The mean change on the female mathematics achievement measure is greater than that on the male mathematics achievement measure indicating that the females achieved more growth in mathematics than the males. An independent-samples t-test was conducted to determine whether the difference between the mean change in female and male mathematics achievement was significant. There was no statistically significant difference in the mean change for females and males [$t = 1.100$ for 328 degrees of freedom and $p = .272$ (two-tailed)].

The analysis of gender differences for within the experimental group is presented in Table 23.

Table 23 Differences in means between stage 1 and stage 2 for mathematics achievement measure by gender, experimental group

| SEX | N | Mean diff | SD |
|---------|-----|-----------|------|
| Females | 208 | 3.96 | 4.80 |
| Males | 247 | 3.96 | 4.71 |

Note. N = paired samples.

The mean change on the mathematics achievement measure is equivalent for both females and males indicating that the females and males achieved the same growth in mathematics.

The findings can be summarized as follows:

The results of the mean score gains in mathematics achievement were examined for gender differences across groups (combined control and experimental) and within each group individually. None of the mean score

gains was significantly different by gender. Further, the comparison by gender within the experimental group was found to be at equivalent levels.

Like school group by gender

SES and gender together are a powerful combination — PISA data showed the much greater probability of having low reading skills for boys coming from low SES backgrounds than for boys coming from an average or higher SES background. (Lokan et al., 2001, p. xii)

The analysis of gender within like school groups was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males for the control and experimental groups separately. The results are presented in Table 24.

Table 24 Descriptive statistics for mathematics achievement measure by gender, like school groups

| Like sch group | SEX | GROUP | N | Mean diff | SD | Independent-samples t-test | | |
|----------------|-----|-------|-----|-----------|------|----------------------------|-----|-----------------|
| | | | | | | t | df | Sig. (2-tailed) |
| 9 | F | C | 57 | 1.51 | 4.75 | | | |
| | | E | 77 | 4.74 | 5.26 | -3.663 | 132 | .000 |
| 9 | M | C | 67 | 2.12 | 4.58 | | | |
| | | E | 69 | 4.19 | 4.39 | -2.691 | 134 | .008 |
| 6 | F | C | 52 | 3.67 | 7.06 | | | |
| | | E | 96 | 3.81 | 4.10 | -0.152 | 146 | .879 |
| 6 | M | C | 64 | 2.30 | 6.37 | | | |
| | | E | 139 | 4.60 | 4.72 | -2.876 | 201 | .004 |
| 2 & 5 | F | C | 48 | 2.65 | 2.71 | | | |
| | | E | 35 | 2.66 | 5.30 | -0.013 | 81 | .990 |
| 2 & 5 | M | C | 37 | 2.19 | 3.60 | | | |
| | | E | 39 | 1.28 | 4.42 | 0.978 | 74 | .331 |

Note. N = paired samples, C = control group, E = experimental group. The mean diff values represent differences in means between stage 1 and stage 2.

In like school group 9 the mean change on the female mathematics achievement measure for the experimental group is much greater than that of the control group indicating that the experimental group females achieved more growth in mathematics than the control group females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female control and female experimental mathematics achievement was significant. There was a statistically significant difference in the mean change for females in favour of the experimental group [$t = -3.663$ for 132 degrees of freedom and $p = .000$ (two-tailed)].

In like school group 9 the mean change on the male mathematics achievement measure for the experimental group is much greater than that of the control group indicating that the experimental group males achieved more growth in mathematics than the control group males. An independent-samples t-test was conducted to determine whether the difference between the mean change in male control and male experimental mathematics achievement was significant. There was a statistically significant difference in the mean change for males in favour of the experimental group [$t = -2.691$ for 134 degrees of freedom and $p = .008$ (two-tailed)].

In like school group 6 the mean change on the female mathematics achievement measure for the experimental group is marginally more than that of the control group indicating that the experimental group females achieved marginally more growth in mathematics than the control group females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female control and

female experimental mathematics achievement was significant. There was no statistically significant difference in the mean change for females [$t = -0.152$ for 146 degrees of freedom and $p = .879$ (two-tailed)].

In like school group 6 the mean change on the male mathematics achievement measure for the experimental group is much greater than that of the control group indicating that the experimental group males achieved more growth in mathematics than the control group males. An independent-samples t-test was conducted to determine whether the difference between the mean change in male control and male experimental mathematics achievement was significant. There was a statistically significant difference in the mean change for males in favour of the experimental group [$t = -2.876$ for 201 degrees of freedom and $p = .004$ (two-tailed)].

In like school group 2 and group 5 the mean change on the female mathematics achievement measure for the experimental group is only slightly more than that of the control group indicating that the experimental group females achieved slightly more growth in mathematics than the control group females. An independent-samples t-test was conducted to determine whether the difference between the mean change in female control and female experimental mathematics achievement was significant. There was no statistically significant difference in the mean change for females [$t = -0.013$ for 81 degrees of freedom and $p = .990$ (two-tailed)].

In like school group 2 and group 5 the mean change on the male mathematics achievement measure for the control group is greater than

that of the experimental group indicating that the control group males achieved more growth in mathematics than the experimental group males. An independent-samples t-test was conducted to determine whether the difference between the mean change in male control and male experimental mathematics achievement was significant. There was no statistically significant difference in the mean change for males [$t = 0.978$ for 74 degrees of freedom and $p = .331$ (two-tailed)].

The findings can be summarized as follows:

The results of the like school groups were examined for gender (same sex). For females within like school group 6, 2 and 5 none of the means was significantly different. On the other hand, within like school group 9, there was a significant difference in favour of the experimental group females. For males within like school group 2 and 5 the mean was significantly different in favour of the control group. On the other hand, for males within like school group 9 and 6 there was a significant difference in favour of the experimental group.

Summary

In this chapter the results of the student responses to the mathematics assessment were examined to determine growth in student knowledge and understanding of mathematics as they proceeded through the first year of secondary school. Data analyses were conducted to examine any difference in performance between Stage 1 and Stage 2. Comparisons between the control and experimental group of students were presented

as well as comparisons within like-school groups. Gender differences in mathematics achievement were also investigated.

In summarizing these results the issues outlined at the beginning of this chapter will now be addressed. Where appropriate comparison will be made with the findings of the Hanchon Graham (2000) study and two major international studies reporting on Australian achievement in mathematics. One of these reports, the Mathematics and Science Achievement of Junior Secondary Students in Australia (Zammit et al., 2002), analyses and interprets the data collected as part of The Third International Mathematics and Science Study (TIMSS) 1998. This assessed the mathematics and science achievements of students in their second year of high school. The other report the Programme for International Student Assessment (PISA) surveys student mathematical and scientific literacy skills with a major focus on the ability of 15 year-old students to 'apply their knowledge and skills to real-life problems and situations, rather than how much curriculum-based knowledge they possess' (Lokan et al., 2001, p. vii). Although the student age groups in these studies differ from those in the current study, because of their status on the education platform it will be helpful here to consider their findings in relation to socioeconomic status and gender factors.

Identifying the differences in mathematics achievement of the control group and the experimental group between Stage 1 and Stage 2 of the study.

In the current study the control group's mathematics achievement mean increased from 26.50 to 28.90 and the experimental group's mean increased considerably more, from 24.77 to 28.73.

In contrast, Hanchon Graham (2000) observed in her US research there was a significant overall decline in students' grades in mathematics from the beginning to the end of middle school. This decline was evident by the end of Year 7.

Examining the differences in mathematics achievement between the control group and the experimental group.

Achievement in mathematics improved significantly for both the control and experimental groups. The difference in the means for these two groups at Stage 1 was statistically significant in favour of the control group; the difference in the means for the two groups at Stage 2 was not statistically significant. This revealed a considerable improvement in the experimental group overall.

Determining what effect, if any, socioeconomic factors had on mathematics achievement.

One of the recent key findings for Australia from the PISA report was that apart from gender, socioeconomic status was the most significant background variable when it came to student achievement (Lokan et al., 2001).

The results of the comparisons in the current study (see Table 25) showed that when compared with all like school groups, the most growth in mathematics achievement was that of the experimental group in like school group 9, while the growth of the control students in that like school group was the least. Significantly, these like school group 9 students belong to the most at risk group involved in the study. This experimental group outcome is in direct contrast with the PISA report, which found 'students with lower levels of SES are more likely to have lower achievement levels' (Lokan et al., 2001, p. 162).

Table 25 Summary statistics for stage 1 and stage 2, mathematics achievement measure

| Like school group | group | N | Stage 1 | | Stage 2 | | Mean diff |
|-------------------|-------|-----|---------|------|---------|------|-----------|
| | | | Mean | SD | Mean | SD | |
| 2+5+6+9 | C | 326 | 26.50 | 7.08 | 28.90 | 6.67 | 2.40 |
| | E | 455 | 24.77 | 6.75 | 28.73 | 5.64 | 3.96 |
| 9 | C | 125 | 27.20 | 6.11 | 29.10 | 6.18 | 1.90 |
| | E | 146 | 23.62 | 6.94 | 28.10 | 6.01 | 4.48 |
| 6 | C | 116 | 23.72 | 7.99 | 26.64 | 7.55 | 2.92 |
| | E | 235 | 25.08 | 6.55 | 29.36 | 5.47 | 4.28 |
| 2 and 5 | C | 85 | 29.26 | 5.68 | 31.71 | 6.77 | 2.45 |
| | E | 74 | 26.03 | 6.77 | 27.96 | 5.25 | 1.93 |
| 6+9 | C | 241 | 25.53 | 7.28 | 27.91 | 6.98 | 2.38 |
| | E | 381 | 24.52 | 6.72 | 28.88 | 5.71 | 4.36 |

Note. N = paired samples, C = control group, E = experimental group. The mean diff values represent differences in means between stage 1 and stage 2.

Interestingly, the Longitudinal Surveys of Australian Youth (LSAY) reported that ‘as a school’s percentage of students from other language backgrounds increased, its scores on tests of reading comprehension and mathematics decreased’ (Rothman, 2002, p. 37). As can be seen from Figure 19, the experimental group here, however, had noticeably more LOTE speakers than any other group in the study, which was again in direct contrast with both LSAY and the PISA (Lokan et al., 2001) findings.

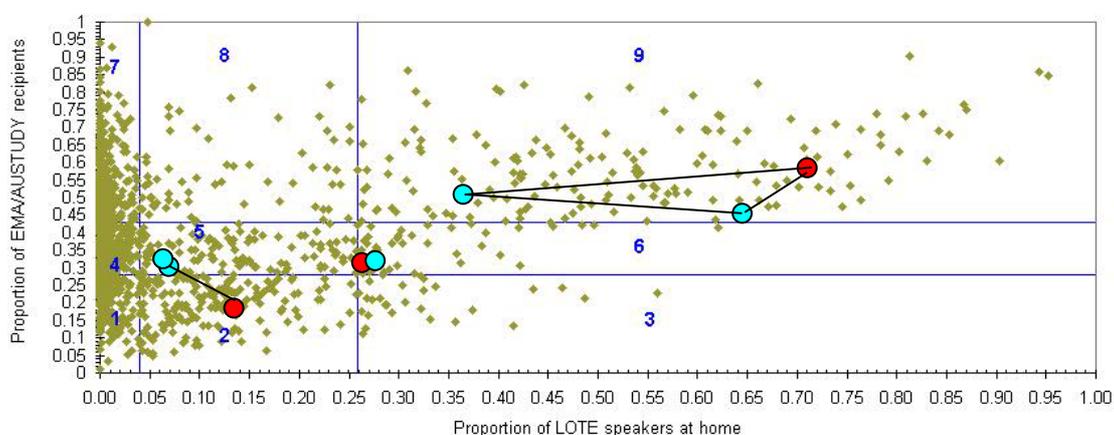


Figure 19 Like school group LOTE statistics

Key: ● = Experimental school, ● = Control school

Note. Adapted from *VCE Benchmarks 2001*, Standards and Accountability Department, Department of Education, Employment and Training, Victoria. Copyright 2001 State of Victoria.

The control group results, on the other hand, are directly in line with the general trend of the PISA study, which found ‘students with lower levels of SES are more likely to have lower achievement levels’ (Lokan et al., 2001, p. 162). Summary socioeconomic statistics for the particular schools in this study are shown in Table 26.

Table 26 Summary statistics for like school group LOTE, EMA/Youth allowance, community employment

| Like school group | 9 | 6 | 2 and 5 |
|-------------------------|-------|---------------|---------|
| LOTE | >0.26 | >0.26 | <0.26 |
| EMA/Youth allowance | >0.43 | >0.28 – <0.43 | <0.43 |
| Unemployment | 15–19 | 8–9 | 6–7 |
| Professional/managerial | 19 | 26 | 43 |
| Clerical/sales/service | 29 | 35 | 34 |
| Trades | 16 | 16 | 12 |
| Production labourers | 36 | 23 | 11 |

Note. Values expressed as per cent.

That the outcomes in like school group 9 were not happenstance is borne out by the fact that the second to most gains in mathematics achievement in the study were those of the experimental group in like school group 6 (see Table 25), the second to most at risk group in the study.

Another major background variable in the PISA (Lokan et al., 2001) study with respect to achievement was socioeconomic status based on the occupations of the parents. Paradoxically, in this study students from the experimental group whose parents had the highest incidence of management and professional roles (see Table 26) made the least gains of the experimental group (see Table 25). However, as previously stated the circumstances under which post-testing took place was far from satisfactory.

Determining what effect, if any, gender differences had on mathematics achievement.

As can be seen from current international studies gender equity is a hot topic in mathematics and there is much debate in Australia centring round a supposed decline in boys' achievement in many academic areas relative to that of girls. In the current study the focus was on the differences existing across groups, that is, the combined control and experimental groups, and within each group.

The analysis of gender differences across both the control and experimental groups for the students who participated in both Stage 1 and Stage 2 of the study was determined by comparing the value of the difference between the two means (Stage 1 and Stage 2) for females and males.

The results of the mean score gains in mathematics achievement were examined for gender differences across groups (combined control and experimental) and within each group individually. Although females overall made more progress than males, none of the mean score gains was significantly different by gender. This is in line with TIMSS (Zammit et al., 2002), which found that gender differences in mathematics performance have almost disappeared.

Similarly, in the current study the comparison by gender within the experimental group was found to be at equivalent levels. With respect to mathematics in Australia, TIMSS found that 'there was no difference in achievement between Australian boys and girls' (Zammit et al., 2002, p. 28).

The results of the like school groups were examined by gender, girls vs girls, boys vs boys. For females within like school group 6, 2 and 5 none of the means was significantly different. On the other hand, within like school group 9, the group most at risk, there was a significant difference in favour of the experimental group females.

For males within like school group 2 and 5 the mean was significantly different in favour of the control group. On the other hand, for males within like school groups 9 and 6, the groups most at risk, there was a significant difference in favour of the experimental group. These findings are in direct contrast with PISA, which although making the point with reading skills, found there was a 'much greater probability of having low reading skills for boys coming from low SES backgrounds than for boys coming from an average or higher SES background' (Lokan et al., 2001, p. xii).

Significantly, in the current study the most gains by male students were achieved by like school group 9, the group, which was most at risk.

Conclusion

We saw in chapter 1 how it is generally accepted that students' grades fall off considerably in almost any school transition (Midgley et al., 1989). And although the subject has attracted little research, the particular discontinuity that occurs between primary and secondary school, which has long been recognised as a problem, is commonly referred to as 'the transition problem' (Clarke, 1989, p. 2). In direct contrast with this generally accepted view it is evident from the findings in the current study that there was significant growth in mathematics achievement and

knowledge for both the control group and the experimental group as measured by the PATMaths assessment instruments.

The results of the comparison showed that although the difference in the means for these two groups at Stage 1 were statistically significant in favour of the control group, the difference in the means for the two groups at Stage 2 were trivial and not statistically significant. This revealed a particularly significant achievement overall for the experimental group.

The most significant finding, however, came with mathematical achievement of the experimental group students who were most at risk. They achieved the most growth in the entire study. And this despite the fact that both the international studies discussed in this chapter found that of all the variables socioeconomic status was the dominant negative factor with respect to student achievement.

In the light of these findings it is difficult not to conclude that the experimental intervention has had a positive effect on mathematical achievement in the experimental group and more especially with the students traditionally regarded as those most at risk.

In the next chapter the relationship between self-efficacy and mathematics achievement is examined.

Chapter 8: Relationship between self-efficacy and mathematics achievement

Introduction

In the third chapter the literature and research relating to the role self-efficacy plays in the learning process was reviewed and the various constructs that come into play in predicting and mediating academic performance. This chapter details and discusses the investigation of the relationship between student self-efficacy beliefs and mathematics achievement. More explicitly it explores the actual effects students' self-efficacy (the student-level variable) has had on their mathematics achievement over time. Data analyses were carried out to explore this relationship, and comparisons between the control and the experimental groups of students are examined as well as comparisons within each group. Low ability student differences are also investigated.

Research questions

The research questions that guided the investigation into the relationship between student self-efficacy beliefs and achievement in this study are as follows:

What are the relationships between self-efficacy and mathematics achievement?

Are these the same for the control and experimental groups?

Statistical methods

Since the focus of the study was on change, which 'requires that observations are made for at least two points in time' (Keeves, 1997, p. 138), only the students who participated in both Stage 1 and Stage 2 were used in the analysis.

There is no single method capable of assessing the potential impact, which student self-efficacy beliefs may have on mathematics achievement over time. Thus, the data analysis was to a large extent exploratory. Descriptive statistics of student change in self-efficacy beliefs between each stage of the study were calculated (the mean difference between Stage 1 and Stage 2) and plotted against a mathematics achievement scale (derived from the Stage 2 mathematics achievement data) to see if a relationship existed between the two variables. It was considered unsound statistically to use the mean difference between Stage 1 and Stage 2 mathematics achievement data for this analysis as the potential ceiling effect would give a distorted view of the relationship. For example, high achieving students who scored 38 out of a possible 39 at Stage 1 could only possibly gain 1 score point at Stage 2. This growth of just one score point would naturally lead to the assumption that there was little growth when in fact the opposite is likely to be the case since the difference of one point towards the top (and bottom) of the scale is much more than a difference of one point in the middle of the scale.

Relationships analysis: Aims

Though a variety of factors can have an influence on the level of student achievement, the focus in this analysis will be on the relationship between self-efficacy beliefs and mathematics achievement. The gathered data will be analysed in an attempt to gauge the true extent to which these self-efficacy beliefs actually do determine student achievement.

Accepting that self-efficacy beliefs have an impact on student achievement (see Chapter 3) a mathematics achievement scale was devised and a comparison with mathematics achievement and change in self-efficacy beliefs was carried out to determine the association.

Developing a mathematics scale

The mathematics assessment instrument used in the study provides norm-referenced stanine scores. These stanine scores permit student performance in a particular test to be compared with the test scores of other students on the same test. Using a single digit scale, stanines divide the raw scores into nine levels ranging from a low of 1 through to a high of 9.

It is generally assumed that the achievement underlying test performance can be appropriately represented by the normal curve. Accepting this, it would follow that with a large group of students the bulk of them would fall into the average achievement level with small amounts falling into the very low or very high levels. Because stanines are based on the normal distribution the nine scale points hold unequal proportions of students. The proportion of students in each stanine level is shown in Table 27.

Table 27 Percentage of students in each stanine score category

| Stanine | Percentage of students |
|---------|------------------------|
| 1 | 4.01 |
| 2 | 6.55 |
| 3 | 12.10 |
| 4 | 17.47 |
| 5 | 19.74 |
| 6 | 17.47 |
| 7 | 12.10 |
| 8 | 6.55 |
| 9 | 4.01 |

Stanines are regarded most appropriate for reporting results in broad terms (Australian Council for Educational Research, 1998) and are considered suitably precise for all practical purposes in this study. The calibration of the Stage 2 mathematics data into stanines allowed the construction of a three-category scale. The students showing the greatest mathematics achievement, those in stanine 6, 7, 8, and 9, were placed in the high category (3) of the scale. Students showing the least mathematics achievement, those in stanine 1 and 2 were placed in the low category (1), and those in stanine 3, 4, and 5 were placed in the middle category.

Pajares (1996) stresses that in order to accurately predict academic outcomes from students' self-efficacy beliefs, 'self-efficacy judgments should be consistent with and tailored to the domain of functioning and/or task under investigation' (p. 550). With this in mind, an analysis of the self-efficacy subscale, task specific confidence, was carried out separately to avoid 'a general sense of efficacy' (Pajares, 1996, p. 547).

Within the abovementioned three categories, this analysis will first compare the control group's relationship between self-efficacy and mathematics achievement. Then, it will compare the three categories within the experimental group. Next, each group's relationship between task specific confidence and mathematics achievement will be explored. Finally, the lowest performing students' relationship between self-efficacy and mathematics achievement will be investigated.

Relationship: Self-efficacy and mathematics achievement

The descriptive statistics for the relationship between self-efficacy beliefs and mathematics achievement for the control students who participated in both Stage 1 and Stage 2 of the study are shown in Table 28. The mean measure differences in students self-efficacy beliefs were very low while the mean measure difference in category 2 shows no change. The comparison between categories shows very small differences. Thus, since the differences are trivial it is doubtful whether self-efficacy beliefs can be considered as having made an impact on students mathematics achievement.

Table 28 Descriptive statistics control group by category, self-efficacy variable

| Category | Variable | N | Mean diff | SD |
|----------|---------------|-----|-----------|-----|
| 1 | Self-efficacy | 42 | .04 | .47 |
| 2 | Self-efficacy | 141 | .00 | .56 |
| 3 | Self-efficacy | 122 | .09 | .52 |

Note. The values represent differences in means (expressed in logits) between stage 1 and stage 2. Category 1 = PATMath stanine 1,2 (low ability); 2 = stanine 3,4,5 (below average-average ability); 3 = stanine 6,7,8,9 (above average ability); N = paired samples.

The descriptive statistics for the relationship between self-efficacy beliefs and mathematics achievement for the experimental students who participated in both Stage 1 and Stage 2 of the study are shown in Table 29. In direct contrast to the control group, the mean measure difference in students self-efficacy beliefs for each category shows that the students self-efficacy beliefs became more positive between Stage 1 and Stage 2. The most change is in category 3 and the difference between each of the other two categories is quite small.

Table 29 Descriptive statistics experimental group by category, self-efficacy variable

| Category | Variable | N | Mean diff | SD |
|----------|---------------|-----|-----------|-----|
| 1 | Self-efficacy | 37 | .24 | .53 |
| 2 | Self-efficacy | 264 | .22 | .61 |
| 3 | Self-efficacy | 130 | .31 | .67 |

Note. The values represent differences in means (expressed in logits) between stage 1 and stage 2. Category 1 = PATMath stanine 1,2 (low ability); 2 = stanine 3,4,5 (below average-average ability); 3 = stanine 6,7,8,9 (above average ability); N = paired samples.

In response to Pajares (1996) caution, to increase accuracy of prediction, there follows an examination of the task specific confidence data and their relationship to mathematics achievement.

Relationship: Task specific confidence and mathematics achievement

The descriptive statistics for the relationship between task specific confidence and mathematics achievement for the control students who participated in both Stage 1 and Stage 2 of the study are shown in Table 30. It is interesting to note that the mean differences in students' perceptions of their abilities between the categories descends from low

ability students (category 1) having the most growth in confidence in their ability to successfully perform mathematics tasks, to high ability students (category 3) having the least gains in confidence in their ability to successfully perform mathematics tasks. However, the difference between each of the categories is trivial. This trend is in reverse to that anticipated by the research literature and thus it is doubtful that task specific confidence can be considered as having made an impact on mathematics achievement in this case.

Table 30 Descriptive statistics control group by category, task specific confidence variable

| Category | Variable | N | Mean diff | SD |
|----------|--------------------------|-----|-----------|------|
| 1 | task specific confidence | 42 | .40 | 1.38 |
| 2 | task specific confidence | 141 | .24 | 1.24 |
| 3 | task specific confidence | 122 | .23 | 1.16 |

Note. The values represent differences in means (expressed in logits) between stage 1 and stage 2. Category 1 = PATMath stanine 1,2 (low ability); 2 = stanine 3,4,5 (below average-average ability); 3 = stanine 6,7,8,9 (above average ability); N = paired samples.

The descriptive statistics for the relationship between task specific confidence and mathematics achievement for the experimental students who participated in both Stage 1 and Stage 2 of the study are shown in Table 31. It is interesting to note that in direct contrast to the control group, the mean differences in students' perceptions of their abilities between the categories descends from high ability students (category 3) having the most growth in confidence in their ability to successfully perform mathematics tasks to low ability students (category 1) having the least gains in confidence in their ability to successfully perform mathematics tasks. Since there is a marginal difference between category 1 and 2 and a considerable difference between category 2 and 3 it seems probable that

task specific confidence had a positive impact on mathematics achievement. It is also worthy of note that the gain in category 1 for the experimental group is more than the gain in the control group category 2 (0.23), category 3 (0.24) and category 1 (0.40).

Table 31 Descriptive statistics experimental group by category, task specific confidence variable

| Category | Variable | N | Mean diff | SD |
|----------|--------------------------|-----|-----------|------|
| 1 | task specific confidence | 37 | .47 | .83 |
| 2 | task specific confidence | 264 | .61 | 1.16 |
| 3 | task specific confidence | 130 | 1.08 | 1.31 |

Note. The values represent differences in means (expressed in logits) between stage 1 and stage 2. Category 1 = PATMath stanine 1,2 (low ability); 2 = stanine 3,4,5 (below average-average ability); 3 = stanine 6,7,8,9 (above average ability); N = paired samples.

Because at risk students form an ancillary focus in this study, there now follows an examination of the data of the lowest achieving students.

Relationship: Self-efficacy, task specific confidence, and mathematics achievement for at risk students

In order to facilitate an examination of the data of the lowest performing students, a more finely tuned breakdown of the mathematics scale was needed to identify the particular students falling into this class. Based on the normal distribution curve and representing approximately 4 per cent of the population, students achieving at stanine 1 can be fairly described as very low in ability. A further calibration of the mathematics data permitted these stanine 1 students to be identified.

The descriptive statistics for the relationship between self-efficacy beliefs and mathematics achievement for the control and experimental students

who participated in both Stage 1 and Stage 2 of the study are -0.20 and 0.43 respectively. In comparing the two values the difference is considerable indicating that while self-efficacy beliefs for the experimental group showed considerable gain, the control group showed the reverse. The experimental group also showed a considerable gain in task specific confidence (0.92) with the control group again showing the reverse (-0.19). While it should be noted that there were only twelve students fitting this category in the control group and three in the experimental group, it is also noteworthy that the control group outnumbered the experimental group.

Summary and discussion of results

One of the findings in TIMSS (see Zammit et al., 2002) was that students who believed they usually did well in mathematics did better than those who did not share that belief. Specifically 'self-efficacy was found to have the highest correlation of any student characteristic variable with achievement' (Zammit et al., 2002, p. 132). Similarly PISA found 'Students who were more self-assured about accomplishing certain tasks ... were more likely to perform better than students with less positive beliefs in their own capabilities' (Lokan et al., 2001, p. 157). As with this study, the questionnaires asked students about their perceptions of their ability in mathematics. For the experimental group in the current study these data are consistent with both the TIMSS and PISA findings.

The data also shows that for the control group, while all ability levels displayed growth in mathematics knowledge and understanding there was little change in their self-efficacy beliefs. It seems likely, however, that these students started out with inflated views about their abilities and their

abilities simply caught up with their expectations. This is borne out by Fullarton when she reports 'both males and females moderated their expectations of themselves over the transition to secondary school, perhaps reflecting a more realistic idea about what they could achieve' (1998, p. 143). It would also seem likely that the experimental group started out with similarly inflated expectations, yet, in direct contrast, their change in self-efficacy beliefs occurred across all ability groups with the most appearing in those students of high ability.

In line with the researcher's expectations, within the experimental group, the low ability group (category 1) showed the least growth in task specific confidence. In contrast, within the control group, the low ability group (category 1) showed the most growth in task specific confidence, a puzzling finding which does not fit the research trend and one for which the researcher has no explanation.

When comparing the least achieving students in the study, stanine 1 control students with stanine 1 experimental students, the experimental group showed gains across both the self-efficacy variable and the task specific confidence variable, whereas the control group showed a decline in both.

Conclusion

An examination of the experimental group's data in this study shows that across all ability levels students with more firmly held self-efficacy beliefs in mathematics achieved more highly. These data confirm previous research findings (for reviews see Bandura, 1986; Pajares, 1996) and indicate the value of self-efficacy beliefs in the prediction of students'

mathematics performance. In this study the highest levels of achievement in mathematics were reported in those students with the most gain in self-efficacy beliefs. These data also confirm the findings related to Direct Instruction research (see chapter 2) that found Direct Instruction was effective not only for regular students but also those at risk. This is borne out by the fact that the most gain in task specific confidence was seen in the experimental group, and on closer analysis the lowest performing students here made considerable gain in their self-efficacy beliefs and task specific confidence while the control group showed the reverse.

As previously stated in chapter 3 the most influential source of self-efficacy information is personal performance attainment (Bandura, 1977, 1986; Pajares, 1997) because it is based on personal experience of mastery (Pajares, 1997). The experimental intervention actually required the students to daily monitor their progress towards mastery of academic goals. This daily monitoring had the effect of modifying the students' self-efficacy beliefs in that as their goals were attained students realised they were capable of performing certain tasks, thus enhancing their confidence with respect to future learning.

The conclusion was reached in Chapter 3 that if we can determine how students come to estimate their ability we should be well equipped to provide interventions aimed at improving their self-efficacy beliefs so they are better able to exploit their talents and potential. In the following chapter it will be opportune to examine student responses to the short answer questions included in the questionnaires with a view to gaining further insight into their views and attitudes towards learning with a particular focus on the mathematics classroom.

Chapter 9: Self-ratings in mathematics and perceptions of the mathematics lesson

The previous chapter dealt with the relationship between self-efficacy and mathematics achievement. This chapter will examine student responses to the self-ratings in mathematics and short answer items with a view to gaining further insight into their attitude towards learning with a particular focus on the mathematics lesson. To enable the researcher better generalise the views and attitudes of students towards mathematics the decision was made to compare data, which were collected at different points in time, using the same instrument and the same groups of students as well as different students at the same year level. Since, one section of one questionnaire was identical to that presented in the Fullarton (1998) study, it was possible to examine their views and attitudes towards mathematics and compare them with those of the students in the current study thus enhancing its generalisability.

Stage 1 Ratings in mathematics

This section compares the current study data to that of Fullarton (1998) with respect to how students perceived their mathematics ability and how they thought other people perceived the student's mathematics ability. The student questionnaires in both studies included a set of items where students were required to rank on a scale of A = excellent, to E = weak, how good they thought they were at mathematics, how good they would like to be, where their teacher, parents and classmates would place them on the scale, and how good they and their parents would like them to be at mathematics. Fullarton termed this part of the questionnaire the *How Good? Measure*.

The control and experimental students who took part in Stage 1 of the current study were grouped together (CS) for comparison with the students who took part in Stage 2 of the Fullarton study (FS). Fullarton collected data at three stages and it was considered that Stage 2 (late Year 6) was the closest match to Stage 1 (beginning Year 7) of the current study. The mean measure and the standard deviation of the distribution is presented in Table 32. The higher the mean measure, the higher the self-rating. For example, a mean measure of 5 = excellent, 3 = average and 1 = weak.

Table 32 Descriptive statistics for students' self-rating, stage 1 current study and stage 2 Fullarton study

| | | Stage 1 Current study | | | Stage 2 Fullarton | | |
|------|-----|-----------------------|------|------|-------------------|------|------|
| | | N | Mean | SD | N | Mean | SD |
| HGS | All | 901 | 3.68 | .87 | 302 | 3.61 | .78 |
| | F | 432 | 3.51 | .85 | | 3.48 | .77 |
| | M | 469 | 3.84 | .85 | | 3.74 | .77 |
| HGSW | All | 899 | 4.71 | .66 | 302 | 4.71 | .52 |
| | F | 429 | 4.70 | .66 | | 4.64 | .55 |
| | M | 470 | 4.72 | .66 | | 4.78 | .48 |
| HGT | All | 899 | 3.60 | .94 | 302 | 3.51 | .86 |
| | F | 433 | 3.49 | .89 | | 3.45 | .80 |
| | M | 466 | 3.70 | .98 | | 3.57 | .92 |
| HGP | All | 898 | 3.81 | .95 | 302 | 3.73 | .92 |
| | F | 432 | 3.69 | .90 | | 3.70 | .88 |
| | M | 466 | 3.92 | .98 | | 3.75 | .96 |
| HGPW | All | 901 | 4.68 | .67 | 302 | 4.65 | .56 |
| | F | 433 | 4.66 | .67 | | 4.55 | .60 |
| | M | 468 | 4.71 | .68 | | 4.75 | .56 |
| HGC | All | 888 | 3.71 | .98 | 302 | 3.53 | .98 |
| | F | 427 | 3.64 | .92 | | 3.47 | .81 |
| | M | 461 | 3.77 | 1.03 | | 3.58 | 1.11 |

Key: HGS=How good are you at maths? HGSW=How good would you like to be?
HGT=How good does your teacher think you are? HGP=How good do your
parents think you are? HGPW=How good would your parents like you to be?
HGC=How good do your classmates think you are?

One of the five items used in the self-rating measure was 'How good are you at maths?' (HGS). At the beginning of Year 7 the students in the current study rated themselves as having slightly above average ability in mathematics. When the HGS mean measure of the CS and the FS are compared, the mean measure of the FS is just below the mean measure of the CS, in fact, the difference (0.07) is trivial. The standard deviation value for the FS (.78) was smaller than for the CS (.87). The standard deviation value indicates that the students' measures showed more variability in the CS than in the FS. Thus a majority of the students rated themselves as having above average ability in mathematics, but as can be seen from the analysis of the mathematics assessment for the CS (see chapter 7), for many students this is a maladaptive belief.

When the mean measure of the CS and the FS are compared on the second item, 'How good would you like to be at maths?' (HGSW), the mean measure for the CS and the FS is equivalent. The mean measure for this item was the highest overall, and showed that all students preferred to be within the range of, above average, and, excellent.

For the third item, 'Where would your teacher put you on this scale?' (HGT) the mean measure was the lowest overall, and the two groups indicated slightly above average. The mean measure of the FS is just below the mean measure of the CS; the difference (0.09) is trivial.

When the mean measure of the CS and the mean measure of the FS are compared on the fourth item, 'Where would your parents put you on this scale?' (HGP), both groups rated slightly above average. The mean

measure of the FS is just below the mean measure of the CS; the difference (0.08) is trivial.

When the mean measure of the CS and the FS are compared on the last item, 'Where would your classmates put you on this scale?' (HGC) the mean measure for the CS is more than that for the FS. The CS mean measure is slightly more (0.03) than the mean measure for their HGS rating whereas the mean measure for the FS is less (0.08) than the mean measure for their HGS rating.

The comparison between the HGS mean measure of female and male students showed that the male students in both groups were more confident of their ability in mathematics than the female students. Fullarton (1998) reported 'both males and females moderated their expectations of themselves over the transition to secondary school, perhaps reflecting a more realistic idea about what they could achieve. Males, however, remained more ambitious than females' (p. 137). Across both studies males had higher teacher ratings (HGT), higher classmate ratings (HGC), higher parent ratings (HGP) and higher parent expectations (HGPW) than the female students.

Stage 1 short answer items

The students involved in Stage 1 of the study responded to five short answer items, which were included to extract more detail on students' enjoyment of mathematics, their worries and expectations concerning transition to secondary school, and whether they thought maths classes would be different in secondary school than they were in primary school.

It is well accepted that qualitative data should speak for themselves and that any analysis should remain as close as possible to the original recorded data (Coffey & Atkinson, 1996). With Stage 1 of this study yielding over 5,000 open student responses it was important to adopt an approach well suited to analysing large amounts of accumulated data. Preliminary scanning was undertaken to see what themes and patterns emerged for it is by looking for 'patterns, themes, and regularities, as well as contrasts, paradoxes, and irregularities that one can move toward generalizing and theorizing from the data' (Coffey and Atkinson, 1996, p. 45).

An initial coding of the responses helped in the identification process with these codes being reviewed and revised from time to time until stable patterns and themes could be discerned. This provided a method of categorizing the responses thereby reducing the data. Next, the coded data were transformed into meaningful data where the emphasis was placed not only on the positive patterns but also on the negative exceptions.

The responses of the students in the control and experimental groups for Stage 1 of the study were combined for comparison with the responses of the 510 students who answered identical items in Stage 2 of the Fullarton (1998) study.

With the first open response item students were asked, 'Do you like maths? Explain why or why not'. Overall, 32 per cent of students responded 'Yes' they liked maths for a variety of reasons such as it was fun, it was challenging, they were good at it, etc. Overall, this response fell

into two sub categories: enjoyment and capability. Under the sub category of, enjoyment, the following student comments were typical: I like numbers and you never get bored; I like a challenge and it is fun. Under the sub category of, capability, the following student comments were typical: I think I am good at this subject; I find it simple to do; my teacher (sic) always gets me to help my friends. Fullarton (1998) reported that 37 per cent of students responded 'Yes' in this category.

A further 19 per cent of students responded 'Yes but' explaining why their yes was qualified. A large proportion of these qualifications related to concerns with fractions and decimals. Overall, this response fell into one category, task specific confidence. The following student comments were typical: if it is timestables (sic) I love it but when it comes to fractions I don't like it as much; I have difficult (sic) with my fractions and stuff (sic); it's alright but sometimes I get worried when it comes to fractions; it is exciting but I don't understand fractions, decimal and algebra (sic) much; my classmates are just too good and I like adding and multiply (sic) but I'm scared about fractions. Fullarton's (1998) study did not include this category.

Overall, 38 per cent of students responded 'No' they did not like maths for a variety of reasons. Some of the stronger comments in this category serve to illustrate just how badly some students are affected by their failure to understand maths. Overall, the responses fell into two sub categories: anxiety and self-image. Under the sub category of, anxiety, the following student comments were typical: my mind goes blank during maths classes; maths to me is very boring and a nightmare; sometimes I don't know the questions and I'm scared to ask cause people will laugh (sic) at me; because

I am not very good at it I get nervous sometimes; I'm not confident enough I just get embarrassed to ask for help; I find it stressful (sic) and hard; I get nervous and really scared if I get it wrong.

Under the sub category of, self-image, the following student comments were typical: maths to me is very boring and a nightmare; it is hard and I am dumb (sic); sometimes I really can't do a question that everyone else can and I feel like a dumb!; I'm not excellent at maths and I'm afraid that I'll be dumb at it and people will tease me; it is very difficult for me and I really really try but I don't get anywhere. The following student comments were typical of those that were unclarified; maths comes easy to me but I definitely (sic) don't like it. I hate math!!

Fullarton (1998) reported that 37 per cent of students responded to the open ended items that they didn't like maths. And, on this issue it is pertinent to quote Fullarton stating that 'a number of strongly negative affective comments were made by students about why they didn't like maths, and it should be noted with some concern that many of these involved fear, embarrassment and ridicule' (1998, p. 142).

Overall, 11 per cent of students were undecided. In this category the following student comments were typical: sometimes I do and sometimes I don't; we had a teacher and she made maths look very bad so I'm not sure how I feel; I don't hate it but I don't like it either.

For the last open response item students were asked, 'Do you think that maths classes will be different in secondary school than they are in primary school? If so, how do you think things will be different?' Overall, 62 per cent of students thought maths classes would be more difficult. The bulk of these students cited algebra as their primary concern and many of them had negative preconceptions of secondary maths teachers. In this category the following student comments were typical: yes it will get harder and harder and they won't help you; yes the teacher won't care if you try or if you don't try; at primary school we work from sheets at sec school we will work from a book.

Fullarton (1998) reported 58 per cent of students thought maths would be harder. Interestingly, she posted algebra as a specific student concern, and noted a number of comments reflecting that students had a 'fairly low opinion of secondary mathematics teachers'. Further, she commented that students' positive comments 'vastly outnumbered' negative comments (p. 146).

Stage 2

The Stage 2 student questionnaire again included a section developed to source more information on student's perceived ability and their attitude towards mathematics. In order to compare the control group and the experimental group's liking of mathematics and their personal ability rating in mathematics, simple descriptive statistics were calculated.

Stage 2 Ratings in mathematics

Students were again asked to rank on a scale of A = excellent, to E = weak, how good they thought they were in mathematics and where their classmates would place them on the scale.

Table 33 shows the results for the CS and the FS. At the end of Year 7 the control group (CG) in the current study rated themselves as having slightly above average ability in mathematics, and the experimental group (EG) even more so. When the HGS mean measure of EG and FS are compared, the mean measure of FS is more (0.01) than the EG and more (0.2) than the CG. Further, the FS males have the highest overall rating on this measure. Thus, a majority of the students rated themselves as having above average ability in mathematics and, as can be seen from the analysis of the mathematics assessment for CS (see chapter 7), for some students this remained a maladaptive belief.

When the mean measure of the CG, EG and FS are compared on the item, 'Where would your classmates put you on this scale?' (HGC), the mean measure for EG is more. Further, EG females have the highest overall rating on this measure whereas the CG females have the lowest overall rating. The results are compared with the Stage 3 Fullarton study.

Table 33 Descriptive statistics for students' self-rating, stage 2 current study and stage 3 Fullarton study

| | | Control Group | | | Experimental Group | | | Stage 3 Fullarton | | |
|-----|-----|-----------------------|------|------|-----------------------|------|-----|-------------------|------|------|
| | | Stage 2 Current study | | | Stage 2 Current study | | | Stage 3 Fullarton | | |
| | | N | Mean | SD | N | Mean | SD | N | Mean | SD |
| HGS | All | 362 | 3.53 | .87 | 474 | 3.62 | .91 | 302 | 3.63 | .85 |
| | F | 174 | 3.33 | .80 | 212 | 3.49 | .84 | | 3.49 | .79 |
| | M | 188 | 3.71 | .90 | 262 | 3.57 | .84 | | 3.77 | .88 |
| HGC | All | 362 | 3.54 | .99 | 474 | 3.64 | .91 | 302 | 3.58 | .93 |
| | F | 174 | 3.49 | .91 | 212 | 3.72 | .96 | | 3.56 | .81 |
| | M | 188 | 3.58 | 1.05 | 262 | 3.70 | .96 | | 3.60 | 1.04 |

Key: HGS=How good are you at maths? HGC=How good do your classmates think you are?

The mathematics lesson

Because observation of student behaviour in mathematics lessons offers particularly beneficial insights into their performance, information concerning how they felt about specific aspects of classroom activity was collected in the questionnaire. This section examines the perceptions and attitudes of the students in the current study towards those specific aspects.

The Stage 2 questionnaire included items on the mathematics lesson, developed to elicit information from students on their preferred learning styles. The categories ranged from a method where the emphasis was on teacher-directed learning to an open section, which provided for any other learning procedures (e.g., student-directed learning) that may have been operating in the classroom. Students were asked to indicate from 5 choices that part of the mathematics lesson they liked the most and that part they liked the least. The choices included, *class text book*, *worksheets*, *correcting homework*, *other*, and the experimental intervention *EMM*. They were asked to choose from the categories, that part of the mathematics lesson they had *learned the most* from and that part of the lesson they had *learned the least*. Further, they were asked to provide some detail about their feelings (explain why they thought this is).

An excellent response rate was obtained on this part of the questionnaire, though there were more missing responses from the control group than the experimental group. Percentages of missing responses are included in the tables and in the item discussion. In order to compare the control group's perceptions and attitudes towards aspects of the mathematics

lesson with those of the experimental group, simple descriptive statistics were calculated. Data in tables are reported in percentages.

The 380 control and 483 experimental students who participated in the Stage 2 data collection were asked, 'Do you like maths?' The scale ranged from, *Yes very much*, to, *No*. The frequency distribution of the responses of the students in Table 34 indicates that 3 per cent of the control students and 1 per cent of the experimental students did not give any response. Eleven per cent of the control students responded, *No*, while a further 11 per cent responded, *Not very much*, resulting overall in 22 per cent of students indicating that they did not like mathematics. In contrast, 5 per cent of the experimental students responded, *No*, while 9 per cent responded, *Not very much*, resulting overall in 14 per cent of students indicating that they did not like mathematics. This shows that 8 per cent more of the control students expressed a dislike of mathematics. Meanwhile, 22 per cent of the control students responded, *Yes very much*, while a further 22 per cent responded, *Yes*, resulting overall in 52 per cent of students indicating that they did like mathematics. In contrast, 24 per cent of the experimental students responded, *Yes very much*, while a further 32 per cent responded, *Yes*, resulting overall in 56 per cent of students indicating that they did like mathematics. This shows that 4 per cent more of the experimental students expressed a liking for mathematics. The remaining students responded, *Sort of*, indicating a relatively undecided view.

Table 34 Frequency distribution % of the students' liking of mathematics stage 2

| | Control | | Experimental | |
|---------------|---------|-----|--------------|-----|
| | N | % | N | % |
| No | 40 | 11 | 25 | 5 |
| Not very much | 43 | 11 | 44 | 9 |
| Sort of | 87 | 23 | 140 | 29 |
| Yes | 115 | 30 | 155 | 32 |
| Yes very much | 82 | 22 | 114 | 24 |
| Total | 367 | 97 | 478 | 99 |
| Missing | 13 | 3 | 5 | 1 |
| Total | 380 | 100 | 483 | 100 |

The reasons for the differing student perceptions and attitudes with respect to mathematics may become apparent in the examination of the following analysis.

The students were asked, 'Which part of the maths lesson do you like the most?' The frequency distribution of the responses of the students in Table 35 indicates that 7 per cent of the control students and 1 per cent of the experimental students gave no response. For this item the most favourable response for both the control and experimental students was, *Class text book*, 33 per cent and 31 per cent respectively. The second to most favourable response for the control students was, *Worksheets* (31%), although this would appear to be more a reflection of their primary school mathematics experience rather than secondary school experience. The second to most favourable response for the experimental students (27%) was the intervention *EMM*. It is interesting to note that although the intervention did not play a part in the control groups curriculum, 5 per cent of these students choose the intervention. Further investigation revealed that for most of these students the intervention was part of their Year 6

curriculum. *Correcting homework*, was the response for about 10 per cent of the students overall. Fifteen per cent of the control students and 11 per cent of the experimental students chose the category, *Other*. Students responding in this category generally reported an unfavourable attitude towards mathematics in the previous item. Further, in their description of, *Other*, these students reported such activities as group work, working on projects, working with calculators and working on the computer. It is worthy of note that students of all ability levels were represented in every category with the exception of, *Correcting homework*, where the majority of students were those of low ability.

Table 35 Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you like the most?

| | Control | | Experimental | |
|---------------------|---------|-----|--------------|-----|
| | N | % | N | % |
| Intervention (EMM) | 19 | 5 | 132 | 27 |
| Class Text Book | 126 | 33 | 150 | 31 |
| Worksheets | 119 | 31 | 97 | 20 |
| Correcting Homework | 33 | 9 | 47 | 10 |
| Other | 58 | 15 | 52 | 11 |
| Total | 355 | 93 | 478 | 99 |
| Missing | 25 | 7 | 5 | 1 |
| Total | 380 | 100 | 483 | 100 |

The students were asked, 'From which part of the maths lesson do you think you have learned the most?' The frequency distribution of student responses in Table 36 indicates that 7 per cent of the control students and 2 per cent of the experimental students gave no response. For this item the most favourable response for both the control and experimental students was, *Class text book*, 66 per cent and 47 per cent respectively. Meanwhile, the second to most favourable response for the control

students was, *Worksheets* (9%). The second to most favourable response for the experimental students was the intervention *EMM* (38%). As already noted, the intervention played no part in the control group's Year 7 curriculum, yet 6 per cent of these students chose it. *Correcting homework*, was the least favourable response (5%) for the control students overall, while for the experimental students, *Correcting homework* (3%), and, *Other* (3%), were equally unfavourable.

Table 36 Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you think you have learned the most?

| | Control | | Experimental | |
|---------------------|---------|-----|--------------|-----|
| | N | % | N | % |
| Intervention (EMM) | 21 | 6 | 185 | 38 |
| Class Text Book | 249 | 66 | 228 | 47 |
| Worksheets | 33 | 9 | 29 | 6 |
| Correcting Homework | 20 | 5 | 16 | 3 |
| Other | 32 | 8 | 14 | 3 |
| Total | 355 | 93 | 472 | 98 |
| Missing | 25 | 7 | 11 | 2 |
| Total | 380 | 100 | 483 | 100 |

In order to elicit further information regarding their responses to the item relating to which part of the mathematics lesson they thought they had learned the most, students were asked, 'Why do you think this is?' Students that choose the intervention *EMM*, expressed opinions that fell into 3 sub categories, namely, teacher factors, student factors and environmental factors. The following student comments were typical — I lern (sic) faster even if it is difficult because the teacher shows us how to do it; I put in the effort to listen; because we do 20 different topics and I understand all of them; because after getting it wrong you no (sic) what you did wrong then it becomes easier to get the right answer; because

when I was at primary school I didn't know these things now after EMM I do; it teaches you to think faster and to listen; we get taught new things each lesson; when I get an answer wrong I find out; it teaches us and explains things better; it's really really quiet and you can concentrate; you get to iron out your bugs each lesson so the next day you don't get it wrong.

The following student comments were typical from those who chose *Class text book*: because you have helpful hints in front of you; it's more easier (sic); because the book explains it in a simpler way; my dad can help me from the book.

The following student comments were typical from those who chose *Worksheets*: because it's easier; because we did them at primary school and now we do them when the teacher is away.

The following student comments were typical from those who chose *Correcting homework*: because when I get the answer wrong I find out the right one and try to see how I got a different one.

The following student comments were typical from those who chose *Other*: because I'm good on computers; my friends can help me.

In direct contrast to the previous item students were asked, 'From which part of the maths lesson do you think you have learned the least?' The frequency distribution of the student responses in Table 37 indicates that 17 per cent of the control students and 2 per cent of the experimental students gave no response. For this item the biggest response for both the

control and experimental students was, *Correcting homework* (37%). The following student comments were typical: because we don't have much homework; because if I get the answer right I'm not learning anything.

Meanwhile, the second biggest response for the control and experimental students was, *Worksheets*, 26 per cent and 28 per cent respectively. The following student comments were typical: because it's not detailed like the text book; it just fills in time.

Class text book, was the third biggest response for both groups (14%). The following student comments were typical: because lots of things aren't explained; it's too hard for me.

Thirteen per cent of students chose the intervention *EMM*. For the experimental students the following comments were typical: because the teacher goes to (sic) fast; its too much information; it's too hard; I already no (sic) how to do it; because when I'm away I can't catch up.

Further, it is interesting to note here that there was a spread of student abilities across all responses with the exception of, *Worksheets*, where the majority of students were those of above average ability.

Table 37 Frequency distribution % of the students' stage 2 responses to the item: Which part of the maths lesson do you think you have learned the least?

| | Control | | Experimental | |
|---------------------|---------|-----|--------------|-----|
| | N | % | N | % |
| Intervention (EMM) | | | 62 | 13 |
| Class Text Book | 48 | 13 | 68 | 14 |
| Worksheets | 97 | 26 | 133 | 28 |
| Correcting Homework | 146 | 38 | 181 | 37 |
| Other | 21 | 6 | 30 | 6 |
| Total | 312 | 83 | 474 | 98 |
| Missing | 68 | 17 | 9 | 2 |
| Total | 380 | 100 | 483 | 100 |

The students who participated in Stage 2 data collection were asked, 'Compared to primary school how do you think you are doing in maths?' The frequency distribution of the responses of the students in Table 38 indicates that 6 per cent of the control students and 2 per cent of the experimental students gave no response. Two per cent of the control students responded, *A lot worse*, while a further 6 per cent responded, *A little worse*, resulting in all 8 per cent of students indicating that compared to primary school they were doing worse in mathematics. Two per cent of the experimental students responded, *A lot worse*, while a further 5 per cent responded, *A little worse*, resulting in all 7 per cent of students indicating that compared to primary school they were doing worse in mathematics.

Meanwhile, 47 per cent of the control students responded, *A lot better*, while a further 28 per cent responded, *A little better*, resulting in all 75 per cent of students indicating that compared to primary school they were doing better in mathematics. In contrast, 55 per cent of the experimental students responded, *A lot better*, while a further 27 per cent responded, *A*

little better, resulting in all 82 per cent of students indicating that compared to primary school they were doing better in mathematics. This shows that the experimental students perceived improvement compared to primary school was 7 per cent more than that of the control students.

The remaining students responded, *About the same*.

Table 38 Frequency distribution % of the students' stage 2 responses to the item: Compared to primary school how do you think you are doing in maths?

| | Control | | Experimental | |
|-----------------|---------|-----|--------------|-----|
| | N | % | N | % |
| A lot worse | 8 | 2 | 11 | 2 |
| A little worse | 21 | 6 | 24 | 5 |
| About the same | 43 | 11 | 61 | 9 |
| A little better | 108 | 28 | 129 | 27 |
| A lot better | 178 | 47 | 267 | 55 |
| Total | 358 | 94 | 473 | 98 |
| Missing | 22 | 6 | 10 | 2 |
| Total | 380 | 100 | 483 | 100 |

In order to elicit further information regarding their responses to the item relating to their ability in mathematics compared to primary school, students were asked, 'Why do you think this is?' Their opinions fell mainly into two categories, namely, teacher factors and difficulty factors.

The following student comments were typical: they teach you more in secondary school; we did baby stuff (sic) in primary school; it's better work in secondary school; we didn't learn much in primary school; because the teachers teach more in secondary school; I have a better teacher; I never understood the teachers in primary school; we spend a lot more time on maths at secondary school; because it took me a while to understand what

me (sic) teacher taught me and to click to what she said; I don't really know maybe it's the teacher; because my teacher explains things better; because the teachers helps (sic) more and makes things easier to understand.

Summary

In this chapter the student responses to the self-ratings in mathematics and short answer items were examined to gain further insight into their attitude towards learning with a particular focus on the mathematics lesson. Because one section of the questionnaire was identical to that presented in the Fullarton study, comparison between the students in both studies was carried out with the current study control and experimental groups being combined.

The first comparison (Stage 1 current study/Stage 2 Fullarton study) showed that overall, students in both studies rated themselves as having slightly above average ability. On being questioned as to how good they would like to be at mathematics the male students overall aspired higher than the female students. And though over the transition from primary to secondary school both male and female students toned down their expectations, the male students' expectations stayed higher, and they were more positive they would be given higher ratings by teachers, parents and peers.

Overall, 32 per cent of students in the current study said that they liked mathematics while Fullarton reported 37 per cent. Their open responses mainly came under the categories of enjoyment and capability. A further

19 per cent of the current study students also said that they liked mathematics but qualified this with concerns about fractions and decimals.

In both studies about 38 per cent of students responded that they did not like mathematics. Overall, the responses fell into two sub categories: anxiety and self-image. It was of concern in both studies that many comments for their dislike of the subject 'involved fear, embarrassment and ridicule'. Further, 11 per cent of students in the current study were undecided.

In describing their thoughts on how different maths classes were likely to be at secondary school, about 60 per cent of student in both studies thought they would be more difficult. Across both studies a large number of students' explanations suggested they had a low opinion of secondary teachers. Further, Fullarton reported that students' positive comments about maths classes at secondary school vastly outnumbered negative comments and this was the same for the current study.

At the final stages of both studies the vast majority of the students rated themselves as having slightly above average ability in mathematics and it was noted in the current study that for some this remained a maladaptive belief.

At Stage 2, overall, 22 per cent of the control group indicated they did not like mathematics. In contrast, 14 per cent of the experimental group indicated they did not like mathematics. Overall, 52 per cent of control group indicated they did like mathematics while 56 per cent of the experimental group indicated they did like mathematics.

It was anticipated that the reasons for the differing student perceptions and attitudes with respect to mathematics may have become apparent in the students responses to questions related to the mathematics lesson. Although the researcher is aware (from discussion with teachers participating in study) that some mathematics classrooms in this study had many features associated with student-directed learning, for example, independence in learning, students exploring and attempting to solve problems, this did not appear in the students' perceptions and attitudes as expressed in their open responses. In describing the parts of the mathematics lesson students had learned least or most from, very few students chose, *Other*. One should not necessarily conclude, however, that students have not learned from these situations. It is critical to remember that students are constructing their understanding through experiences, thus, learning may often be hidden with students remaining completely unaware that learning has actually taken place. About 32 per cent of students described the *Class text book* as the part of the mathematics lesson they liked the most. The second to most favourable response for the experimental students (27%) was the intervention *EMM*, while *Worksheets* was the dominant response for the control students.

In describing the part of the maths lesson they thought they had learned the most from, the most favourable response for both the control and experimental students was the *Class text book*, 66 per cent and 47 per cent respectively. Meanwhile, the second to most favourable response for the control students was *Worksheets* (9%). The second to most favourable response for the experimental students was the experimental intervention *EMM* (38%).

Students that choose the experimental intervention *EMM*, expressed reasons that fell into 3 sub categories: teacher factors, student factors and environmental factors. Students chose, *Class text book*, because: you have helpful hints in front of you; it's more easier (sic); because the book explains it in a simpler way; my dad can help me from the book.

In describing the part of the maths lesson they thought they had learned the least from, the biggest response for both the control and experimental students was, *Correcting homework* (37%). Meanwhile, the second biggest response for the control and experimental students was, *Worksheets*, 26 per cent and 28 per cent respectively. *Class text book*, was the third biggest response for both groups (14%). Thirteen per cent of students chose the experimental intervention *EMM*.

Meanwhile, the second biggest response for the control and experimental students was, *Worksheets*, 26 per cent and 28 per cent respectively. The following student comments were typical: because it's not detailed like the text book; it just fills in time.

Class text book, was the third biggest response for both groups (14%). The following student comments were typical: because lots of things aren't explained; it's too hard for me.

Thirteen per cent of students chose the experimental intervention *EMM*. For the experimental students the following comments were typical: because the teacher goes to (sic) fast; it's too much information; it's too hard; I already no (sic) how to do it; because when I'm away I can't catch up.

Further, it is interesting to note that in this category there was a spread of student abilities across all responses with the exception being those students who chose, *Worksheets*, where the majority were of above average ability.

In describing, 'Compared to primary school how do you think you are doing in maths?' overall, 8 per cent of control students indicated that compared to primary school they were doing worse in mathematics while 7 per cent of experimental students indicated that compared to primary school they were doing worse in mathematics.

Overall, 75 per cent of control students indicated that compared to primary school they were doing better in mathematics while 82 per cent of the experimental students indicated that compared to primary school they were doing better in mathematics. Thus, the experimental students perceived improvement compared to primary school was 7 per cent more than that of the control students perhaps as a result of the experimental intervention. The remaining students responded, *About the same*.

Chapter 10: Summary and discussion

Summary

At the very outset of the first chapter in this study it was seen that the current state of mathematics teaching in schools in the Western world, at least, is in an unhealthy condition, and the adult population generally has the most fragile of grasps on the subject. The fact that words such as anxiety, apprehension, fear and dislike, are perfectly acceptable in describing peoples' common feelings towards the general topic of mathematics is worrisome in itself. But the problem is compounded when year after year vast numbers of our students leave school to take up employment as young adults without being able to communicate mathematically. In that first chapter it was argued strongly that the sociological implications of this ongoing situation for the future direction of the Australian workforce must inevitably lead to distinct and possibly insoluble problems. This is more especially so when the global business environment, almost daily is becoming more increasingly technology-based. It was revealed that recent studies point to a breakdown in mathematics education, which was highlighted in the 2002 UNICEF report. There, the latest cross-national surveys into educational performance showed that nearly 30 per cent of Australian Year 8 students were incapable of using basic mathematical skills in everyday situations. Further, the LSAY (Lamb, 1997) report showed that it was these very teenagers, the ones with poor numeracy skills, who are most likely to be unemployed both in the short and long term. It was also seen that although we have these many thousands of students with poor mathematics skills

there was no evidence to show this was their fault. On the contrary, the evidence showed that most of our students' problems relate to curriculum and methods of teaching. Problematically though, there was no consensus as to which aspects of curriculum and what teaching methods were to blame.

With all this in mind one of the major objectives in this study was to determine what effects, if any, a Direction Instruction intervention might have on students' mathematical achievement in the regular mathematics classroom. Further, recognising the importance of self-efficacy in the prediction of academic performance, it was stated that another major objective in the study would be to determine the effect, if any, the experimental intervention might have on students' mathematical self-efficacy.

The study was determined to focus on Year 7, as this particular student transition year has long been seen as a major problem zone, with mathematics ending up a major casualty. Since this transition problem was recognised as being compounded for at risk students, they were to provide a subsidiary focus. It was suggested that secondary teachers in mathematics classes had little time to cater for individual differences, and as the streaming of students according to ability was no longer acceptable, persistent failure for those at risk was inevitable. UNICEF (2002) reported that when these students were ranked academically alongside their peers, in the worst affected countries they lagged approximately five years behind their middle-achieving peers. Though Australia performed slightly better than the US, Germany, New Zealand and Belgium (which were the worst)

it was still ranked in the bottom half of the 24 countries included in the study.

In the second chapter the researcher explored the student-directed versus teacher-directed learning debate. The literature and research pertaining to the constructivist approach to learning was reviewed alongside that relating to Direct Instruction. The origins and elements of these teaching practices were described and potential benefits and problems arising out of both were analysed. There it was seen that the current popular trend in education is called Constructivism and almost all the teacher education programs in this country are based on constructivist approaches. Student-directed learning practices like Constructivism, it was shown, are founded on the notion that children's learning needs are unique and best served by allowing them to pursue their individual interests. As was discussed, Constructivism, or student-directed learning, has been around for more than two thousand years with Plato (1955) instructing parents to 'let your children's lessons take the form of play' (para. 537). It was, however, seen that the idea that real learning cannot take place under the control and direction of a teacher is almost universally attributed to Piaget (1970) when he said that the premature teaching of children something they could have discovered for themselves, prevents them from ever completely understanding it. The current writer queried how much children are actually capable of discovering for themselves, and how long the discovery process should take.

On the other hand it was shown there were educators (e.g., Harris & Graham, 1996; Jones et al., 1997; Swanson & Hoskyn, 1998; Westwood, 2000) saying it is too much to expect children to seek out basic knowledge

in literacy and numeracy unaided. They argued that if teachers did not impart knowledge to the children, how could they be expected to discover for themselves complex theories and concepts that have taken some of the great intellects of the world centuries to put together? It was seen that student-directed learning practices are no simple thing to implement successfully — simply providing students with resources and a location for collaborative learning situations will not in itself facilitate successful learning. Listening and responding to student constructions was seen to be difficult and time-consuming and it was shown the teachers themselves needed to thoroughly understand the body of knowledge behind the discipline. It was argued that while there are obvious potential benefits to be gained from properly implemented student-directed learning it is not a complete theory for learning everything. Some educators (e.g., Dick, 1992; Gagne et al., 1993; Resnick, 1987; Westwood, 2000) thought a constructivist approach was inappropriate for every aspect of learning and that certain learning processes involved in acquiring facts and concepts were quite different from the learning processes involved in the development of intellectual skills and strategies. In their opinion a constructivist approach could not guarantee children would acquire fluency and automaticity with basic number and computation.

It was suggested that students differed with respect to exactly what benefits, and how much meaning they could extract from situations where there was no explicit instruction. Indeed, some authorities (e.g., Grossen, 1993; Hempenstall, 1997; Matthews, 2000) argued that many students learned better when they were explicitly taught and when the curriculum was structured. Research (e.g., Carnine, 1997; Gersten, 1985; Gersten &

Maggs, 1982; Harris & Graham, 1996; Kavale, 1990; Mastropieri et al., 1997; White, 1988) was discussed showing many students with special needs actually require detailed and explicit instruction in order to acquire a wide variety of cognitive and metacognitive strategies. Those researchers thought that students with special needs required more structured and explicit instruction, not less. They argued some students just get confused and frustrated with discovery approaches to learning, and that learners differ greatly in their need for teacher direction. It was seen that while the traditional approaches to learning concentrate on getting students to process information efficiently and effectively, constructivist approaches are directed towards students awareness of, and reflection on the learning experience. It was argued that since learners all have their own unique perspective of the experience, the concept of global learner is not part of the constructivist perspective. It was stated that more capable students often experience problems when exposed to a constructivist approach, and perfectly capable students can become frustrated when required to involve themselves in discussion, analysis and reflection.

Direct Instruction, or teacher-directed instruction, was described in chapter 2 as the antithesis of Constructivism. Being a highly structured system of teacher-student interactions directed by the teacher it was seen by many constructivists (e.g., Reetz & Hoover, 1992; Wakefield, 1997) as inappropriate for this day and age. They argued it was boring and repetitive, and neither necessary nor desirable. One of the most serious drawbacks with Direct Instruction was seen to be the large amount of work it requires from the teachers both in respect of the acquisition of new skills and in the teaching process. Another major criticism with Direct Instruction

was shown to be the restrictive nature of the teaching model itself, and the high degree of specificity required by the actual teaching behaviour. Critics (see Stebbins et al., 1977) were seen to be concerned that the tightly controlled instruction might discourage children from freely expressing themselves and consequently damage their self-esteem. It was accepted that Direct Instruction is indeed specific: the lessons are scripted and teacher behaviour is carefully defined. The current writer, however, argued that Direct Instruction teaching innovations should be malleable enough to allow teachers to adapt them to suit their own classrooms. While it was shown that some of the criticism aimed at Direct Instruction was vague (see Boomer, 1988) and unsubstantiated (see Kuder, 1991), and that much of it revolved around philosophical issues like the teacher's role and the real nature of the learning process, further discussion was outside the ambit of this thesis.

Discussion of Project Follow Through showed that it was instigated to identify teaching models that could elevate the academic performance of America's underprivileged schools from the 20th to the 50th percentile. It was seen that student-directed learning models as well as teacher-directed learning models received a thorough evaluation in Project Follow Through (for the original Follow Through report see Stebbins et al., 1977), the largest, most expensive research in the history of education. There, Abt Associates (1977) results were shown illustrating that student-directed learning had consistently more negative outcomes than those achieved in traditional education on all measures of basic skills, cognitive development, and self-esteem. Further, across multiple implementations and settings, it was shown to be vastly inferior to traditional education.

Despite this, however, it was seen that the student-directed learning adherents have been accused of failing to learn from the empirical data. It was shown that throughout North America student-directed learning practices, which previously came under the rubric, open education, were reappearing under such nomenclature as developmentally appropriate practices and constructivism. In fact the education model now being promoted for elementary school reform in the US is called developmentally appropriate practice yet it was seen to embody exactly the same student-directed learning themes that performed so dismally in Project Follow Through.

The situation in England was shown to be no different. For more than twenty years student-directed learning was England's official educational policy except it was called progressive education, and there, in what was the lengthiest and most comprehensive implementation of student-directed learning practices on record, it was an equally dismal failure.

It was apparent that the analysts of the Project Follow Through evaluation data unanimously agreed that teacher-directed instruction resulted in stronger academic outcomes than the student-directed models. In the final analysis (Stebbins et al., 1977) students being taught under the Direct Instruction model scored close to the 50th percentile in every subject, while the other student-directed models', students consistently scored beneath the 20th percentile.

It is perhaps appropriate here to finish this summation of the student-directed versus teacher-directed learning debate with a snippet from the New York Newsday's education forum.

Project Follow Through, America's longest, costliest and perhaps, most significant study of public school teaching methods quietly concluded this year. The good news is that after 26 years, nearly a billion dollars, and mountains of data, we now know which are the most effective instructional tools. The bad news is that the education world couldn't care less. (Tashman, 1994, p. 36)

On the brighter side, however, it was seen that various scholars (e.g., W. Bishop, 1999; Casazza, 1993; Harris & Graham, 1996; Westwood, 2000) are now seriously arguing that there are definite advantages in achieving a balance between teacher-directed and student-directed learning. Certain constructivists (e.g., Battista, 1999; Inagaki et al., 1998) were even seen to be openly acknowledging that students require basic skills to solve certain problems or understand certain concepts, and that to acquire these skills they need teacher-directed learning. Some freely acknowledged there were certain things that had to be learned in a purely mechanical way.

It was concluded in chapter 2 that in the acquisition of basic skills an instructional approach was more suitable and while explicit teaching might not be necessary for the acquisition of these skills, it was generally accepted that most mathematical understanding could be successfully acquired through its implementation. Importantly, it was argued that the combined approach was especially relevant for children with special needs. Definite advantages were seen in striking a balance between teacher-directed learning and student-directed learning and it was argued the method of instruction best suited to the type of learning should be

adopted and that the approach chosen for any particular learning situation would ideally be based on a *fitness for purpose* appraisal.

The third chapter examined the literature on the role of self-efficacy in the learning process analysing the various constructs that come into play in predicting and mediating academic performance. In that chapter it was shown that Bandura (1986) had placed the concept of confidence in learning securely into the academic realm; under the title of self-efficacy he set out to analyse and utilise the role confidence plays in the learning process. Bandura's (1986) Social Cognitive Theory was discussed and it was seen that what we do in any situation is governed in large part by what we think we can do. Self-efficacy beliefs were shown to be personal judgements on personal capacity and distinctions were made between self-efficacy beliefs and self-concept. The relationships between self-efficacy and engagement, motivation, self-regulation and modeling were also discussed. Self-efficacy was distinguished from self-concept in that self-efficacy beliefs were of a context specific nature involving judgements of capacity to carry out certain tasks, whilst self-concept was concerned with more general judgements like self-worth in respect of performance. It was seen that there is a strong relationship between self-concept and self-efficacy and that as a consequence distinctions between the two have often been fuzzy. The point was also made that the more precisely self-concept comes to be measured, the more closely it will equate to self-efficacy and the harder it will be to determine whether the two are in fact distinguishable.

The four major sources of self-efficacy were described as being: performance attainments, vicarious experiences, verbal persuasion, and

physiological states. And although they are called sources of self-efficacy they do not directly translate into self-efficacy beliefs, rather people make judgements on their own interpretations of their actions. The most influential of these sources was shown to be personal performance attainment because it was based on actual mastery experience. It was demonstrated that while students learn they monitor their progress towards academic goals and this has the effect of modifying self-efficacy beliefs. As students attained their goals they became aware they could perform particular tasks and as a result their confidence in their ability to learn in the future was enhanced — their performances were seen to be providing students with dependable data to measure their self-efficacy.

Where students lack personal experience they were seen to often look to their peers who did have such experience using them as models to estimate their own capabilities. It was shown to be generally accepted that vicarious learning actually speeds up the learning process with the potentially beneficial side effect of protecting the learner from being physically engaged in negative experiences.

Verbal persuasion as a means in itself of engendering self-efficacy was seen to be of limited potency. This was thought to be only truly effective when the persons being persuaded had good cause to believe they were capable of actually performing the task in question.

It was seen that the amount of time and effort students are prepared to put into solving an academic problem depends largely on their perception of their own academic ability. Students with high self-efficacy beliefs, it was

shown, persevere longer and harder and this effort has been termed by educators (e.g., Fullarton, 1998; Skinner et al., 1990), engagement.

Self-efficacy and motivation were seen to be related in that progress in the learning process enhanced self-efficacy, which in turn enhanced student motivational constructs. Though if students felt they couldn't achieve their goals their motivational constructs were diminished. A belief in one's personal competence was shown to be a major element in motivation and this was thought to be best displayed in perseverance. It was, however, apparent that even highly effective and skilled individuals may fail to act in accordance with their self-belief and ability if they lacked the incentive. Generally speaking, for goals to enhance motivational constructs, it was argued they should ideally be relatively specific so progress towards them could be easily measured. General goals, it was concluded, failed to boost motivation whereas specific goals did. Similarly goals set too far in the future failed to motivate as much as shorter term goals.

Self-efficacy and self-regulation were seen to be related in that students with effective self-regulatory skills were shown to not only use their time more efficiently and work more effectively, but they saved themselves unnecessary stress. It was concluded that ideally students' thoughts and behaviours, whilst engaged in the learning process, would be systematically oriented towards attaining their learning goals.

Modeling was said to refer to the 'cognitive, affective, and behavioural changes' (Schunk, 2001, p. 128) that derived from the observation of models. And self-efficacy and modeling were seen to be related in that although students were often unaware of it, they were actually acquiring

certain knowledge from the simple observation of a model. It was argued that by observing model behaviour and the consequences thereof, students were able work out what would happen if they were to behave in a similar fashion to the model. It was seen that some educators even claimed that observing a model successfully perform a particular task could enhance students' feelings of self-efficacy towards that task though they had previously been incapable of performing it. It was also noted that through the observation of others students might tune into mental processes that might otherwise have remained suppressed.

The correlation between academic performance and self-efficacy beliefs was stated in chapter 3 to be higher in mathematics than any other academic area. It was seen as significant that self-efficacy was accepted by some researchers (e.g., Pajares & Kranzler, 1995) as having a stronger influence on performance in the mathematics domain than general mental ability, long regarded as the most powerful predictor of academic outcomes. And it was reported that across all levels of ability students with stronger self-efficacy computed mathematically with greater accuracy.

Chapter 4 presented the researcher's personal perspective then generally described the gathering of information about student self-efficacy and achievement in mathematics so as to measure the effect of the experimental intervention. Details concerning the instruments and procedures used to implement the research were outlined, and descriptions of the schools participating in the study were provided. Also described was the quantitative research method whereby data were collected from students participating in the study at two stages during their first year of secondary school.

The researcher's focus on mathematics during students first year of secondary school was shown to be influenced by several interrelated factors: firstly, the problem teachers face at the beginning of the school year when confronted by the generally inadequate academic standard of their new class; secondly, class sizes were large and time to prepare Year 6 students for secondary school was limited; thirdly, the need to establish firm maths foundations in all students, including those at risk; fourthly, the generally negative attitude and low self-efficacy of upper primary students when it came to mathematics.

It was decided that because the general purpose of this study was to examine the effects of the researcher's own teaching intervention, the quantitative method of research, and more specifically a pretest-posttest control group design, was the best suited. It was argued that the quantitative approach would allow the researcher to view the facts objectively, uncontaminated by a personal perspective, and that the differences between pretest and posttest measures produced by the experimental treatment should be highly reliable.

It was stated that the project submission was approved by the Standing Committee on Ethics in Research Involving Humans (Appendix 1) at a meeting A8/2001 on 4 December 2001 and permission was also obtained from the Department of Education, Employment and Training to conduct research in Government schools (Appendix 4). Purposeful sampling was discussed and it was seen that this occurs when samples are selected in a nonrandom manner, based on member characteristics relevant to the particular research problem

Since at risk students formed a subsidiary focus of this study, and because it is well accepted these students come from socially disadvantaged circumstances, the matter of socioeconomic status was discussed and was a relevant consideration in the selection of schools.

Whilst it was concluded that no consensus existed regarding the definition and measurement of socioeconomic status it was shown in that chapter that it was generally accepted that the working class is amongst the least privileged and low status groups such as these have been specifically identified as being at risk in Australian schools.

The Australian Bureau of Statistics (1997) figures were discussed and it was shown that schools were selected from suburbs in the Melbourne metropolitan area where the occupational status was primarily manual labour with unemployment rates predominantly between 8 – 19 per cent. And the four like school groups (9, 6, 5, 2), representing the best available spread of the different Melbourne school populations, were those chosen by the researcher to participate in the study.

So that variability across school systems could be better controlled it was shown that the research design would be confined to Government secondary schools. Thus, the three schools in like school group 9 were situated in distinctly working class suburban areas in the Southern and Western regions where unemployment ranged from 15 – 19 per cent. The two schools in like school group 6 were situated in largely working class suburban areas in the Northern and Western regions where unemployment ranged from 8 – 9 per cent. The three remaining schools were either like school group 2 or 5 and situated in relatively middle-class

suburban areas in the Eastern and Western regions where unemployment ranged from 6 – 7 per cent. It was stated that a total of 8 schools comprising 54 classrooms and 967 students would participate in the study.

It was stipulated that throughout the study the following instruments would be employed: the Student Questionnaires, the PATMaths Achievement Tests (Australian Council for Educational Research, 1997), and the Experimental Intervention (Farkota, 2000).

The student questionnaires to be implemented at Stage 1 and Stage 2 of the study were discussed and seen to comprise items which were constructed into subscales: perceived control, reaction to challenge, task specific confidence, engagement and general attitude. Scores for the subscales were to be combined to form a total self-efficacy score.

The Progressive Achievement Tests in Mathematics Revised — PATMaths Revised, (Australian Council for Educational Research, 1997) was designated as the assessment instrument to be used in the study to measure student achievement in mathematics. Test 2A would be used at Stage 1 and Test 2B at Stage 2

The experimental intervention was discussed and seen to be a mental maths program specifically designed around Mathematics — A Curriculum Profile for Australian Schools (Curriculum Corporation, 1994). Comprising 20 different strands it was described as a daily program for the entire class requiring 15 minutes to implement, then 5 minutes of feedback diagnosis and correction procedures. The scheduled time was to be 15–20 minutes

and the program was to be implemented at the beginning of the regular mathematics lesson a minimum of 4 times per week.

Chapter 5 discussed the validation analyses of the questionnaires and mathematics assessment instruments employed in this study. Data checking for unmatched entries was described. No outliers were found. Missing data were examined and seen to be randomly distributed. The reason for the variation in the number of control students between Stage 1 and Stage 2 of the study was explained as being the result of a school fire. Students removed from the study, the scoring of skipped items and the number of missing responses in the mathematics assessment were also explained.

The questionnaire subscales and mathematics assessment items were shown to be validated by Rasch measurement and the results were recorded in this chapter.

To determine the most appropriate selection of items for calibration of the student questionnaire subscales it was seen that the researcher undertook a Rasch rating scale analysis (Andrich, 1997; Andrich & Masters, 1988), where the items were divided into five subscales: perceived control, engagement, reaction to challenge, task specific confidence and general attitude towards mathematics. These stated five subscales made up the self-efficacy scale and so as to provide classical indices alongside the Rasch based information the QUEST computer program (1993) was employed. To determine item difficulty and to see whether or not the items fitted the subscales, the results of the indices and the Rasch based information were shown to have been examined. This process was

important because although Fullarton (1998) validated the scales using a principal component and reliability analysis, not all those items were used in the current study, therefore the makeup of the subscales differed slightly.

Thresholds indicating the item difficulty for probability levels were set at 0.5, and to determine whether or not the items discriminated in a similar way between students the infit mean square statistic (INFIT MNSQ) was shown to have been considered. Details of threshold values and INFIT MNSQ coefficients for the subscales of the questionnaire items were described showing calibration of the subscales to be compatible with Rasch rating scale analysis. Estimates reliability for each subscale were seen to be relatively high and the scores obtained could thus be regarded as satisfactorily stable.

It was seen that the 7 items of the perceived control subscale were subjected to Rasch rating scale analysis using QUEST and the results indicated a mean item difficulty -0.23 (SD $.45$). The threshold values showed that all the subscale items had a satisfactory spread for assessing the varying levels of perceived control. An examination of the item fit analysis for the perceived control subscale provided evidence that the items fitted the Rasch model.

The results of the engagement subscale indicated a mean item difficulty -0.60 (SD $.47$) and the threshold values showed that all the subscale items had a satisfactory spread for assessing the varying levels of engagement. An examination of the item fit analysis for the engagement subscale revealed that 5 of the 9 items were outside the acceptable range

and these 5 items were excluded from the final analysis. An examination of the final item fit analysis for the engagement subscale provided evidence that the items fitted the Rasch model.

The results of the reaction to challenge subscale indicated a mean item difficulty -0.10 (SD $.31$) and the threshold values showed that all the subscale items had a satisfactory spread for assessing the varying levels of reaction to challenge. An examination of the item fit analysis for the reaction to challenge subscale revealed that 3 of the 10 items were outside the acceptable range or close to the upper cutting point 1.30 and these 3 items were excluded from the final analysis. An inspection of the final item fit analysis for the reaction to challenge subscale provided evidence that the items fitted the Rasch model.

The results of the task specific confidence subscale indicated a mean item difficulty 0.13 (SD $.33$) and the threshold values showed that all the subscale items had a satisfactory spread for assessing the varying levels of task specific confidence. An examination of the item fit analysis for the engagement subscale revealed that 1 of the 17 items was outside the acceptable range and was excluded from the final analysis. An inspection of the final item fit analysis for the task specific confidence subscale provided evidence that the items fitted the Rasch model.

From the 16 items in the general attitude subscale, 4 items were excluded as they were not included on the Stage 2 questionnaire and a further 2 items were excluded because they were scored differently on the Stage 2 questionnaire. The initial item fit analysis revealed that 5 of the remaining 10 items had INFIT MNSQ coefficients outside the acceptable range and

were also excluded from the final analysis. An examination of the final item fit analysis for the general attitude subscale provided evidence that the items fitted the Rasch model.

A total of 39 items and five subscales made up the self-efficacy scale. The mean INFIT MNSQ for the scale was 1.01 (SD .40) and the threshold values showed that all the items had a satisfactory spread for the assessment of the different levels of self-efficacy. The items combined to form an appropriate self-efficacy scale compatible with the Rasch model. The sum of these item scores was taken to present a total self-efficacy score.

It was stated that the Progressive Achievement Tests in Mathematics Revised (PATMaths Revised) had been designed by the Australian Council for Educational Research (1997) to provide a broad general estimate of student achievement in mathematics and was thus appropriate for the purposes of this study. Internal reliability estimates were provided for each test ranging from 0.87 to 0.92 and were guaranteed by the test developers to contain adequate content validity.

It was explained that with the PATMaths Revised assessment though the test developers claimed the items fitted the Rasch model, no statistical evidence was given to support this. The researcher thus, for the purposes of this study, carried out her own analysis in which reliability was found to be acceptable. The threshold values showed that all the items included in the scale had a satisfactory spread and the item fit analysis for the PATMaths Revised items was satisfactory. The questionnaires and

mathematics assessment instruments were thus seen to be reliable and proved valid within the confines of the study.

In chapter 6 the results of the student responses to the questionnaires were examined to investigate whether their self-efficacy beliefs changed as they proceeded through the first year of secondary school.

Year 7 students were chosen to participate in the study because the first year of secondary school is traditionally regarded as a crucial developmental stage in respect of students' mathematical self-beliefs and their attitudes towards mathematics generally. Studies (e.g., Fullarton, 1998; Hanchon Graham, 2000) have found the views held by these students in respect of learning mathematics to be generally negative.

An independent-samples t-test was conducted on the scale and each of its subscales. Results were computed separately for the students in the control and experimental groups for Stage 1 and Stage 2 of the study and the significance of the difference between the two groups was compared. A paired samples (repeat measures) t-test was conducted on the scale and each of its subscales to evaluate the significance of the differences between Stage 1 and Stage 2 for the control and experimental groups separately.

Student self-efficacy beliefs were examined to see what influence they had on the choices they made, the effort they expend, the time they will persevere in adverse circumstances, and the measure of anxiety or confidence they will bring to a given situation. It was found students' self-efficacy beliefs in mathematics changed as they moved through the first

year of secondary school for the experimental group but the control group virtually remained constant. More specifically the experimental group had statistically significant ($p < .001$) gains in self-efficacy beliefs in mathematics as they proceeded through the first year of secondary school. The experimental groups' mean increased 0.25 logits. The control group's mean increased 0.05 logits. The control group had no significant ($p = .128$) change in self-efficacy.

It was found the changes in the self-efficacy beliefs of the students in the control group differed from those in the experimental group. More specifically there was a significant difference in the mean self-efficacy scores for the control and experimental groups. At Stage 1 the control group had a significantly higher mean whilst at Stage 2 the experimental group had a higher mean showing a particularly significant result overall for the experimental group. There was approximately 12 per cent growth in the mean self-efficacy measures for the experimental group. For the control group there was approximately 2 per cent growth.

Students' perceived control was examined to see what types of thoughts affected their actions. They were required to judge statements about the extent to which they thought outcomes in mathematics were due to effort, for example, 'Trying hard is the best way for me to do well in maths', or unknown strategies, for example, 'I don't know how to keep myself from doing badly in maths', and the extent to which they felt they were able to influence these factors. It was found that the control group had statistically significant gains in perceived control in mathematics as they proceeded through the first year of secondary school whereas the increase in the experimental group was not statistically significant.

Students were examined to see what levels of engagement they exercised in mathematics classes. They were required to judge statements indicating their inclination to engage, for example, 'I dread having to do maths', and, 'When I'm in maths classes I usually just act as though I'm working'. It was found that the control group had a statistically significant decrease in engagement in mathematics as they proceeded through the first year of secondary school, whereas the decrease in the experimental group was small and not statistically significant.

Students were examined to see how they reacted to challenge. They were required to judge statements such as, 'I tell myself that I'll do better next time', and, 'I say I didn't care about it anyway', and, 'I worry that the other students will think I'm dumb'. Items in the scale ranged from feeling at ease, for example, 'Maths doesn't scare me at all'; to feeling distinct anxiety, for example, 'When I'm in maths classes I usually feel uncomfortable and nervous'. It was found that the control group had no change in reaction to challenge in mathematics as they proceeded through the first year of secondary school, and the increase in the experimental group was small and not statistically significant.

Students were examined in task specific confidence because specific judgements are considered more accurate predictors of specific performances than those which are broader and/or less task specific. It was seen that it is deemed critical that self-efficacy be assessed at the optimal level of specificity. To address the context specific nature of the study the researcher developed 17 mathematics items with a view to discerning how confident students felt about responding correctly to these items in a mathematics assessment. The items were similar to those that

they would subsequently be presented with in the mathematics assessment. Students were asked not to attempt to solve the problems but simply to provide confidence judgments as to how successful they thought they would be at solving each problem on a four point Likert-type scale ranging from very confident to not at all confident. For example, 'Express 0.05 as a percentage', and, 'Draw an angle of 135° '. It was found that both the control and experimental groups had statistically significant gains in task specific confidence in mathematics as they proceeded through the first year of secondary school.

Students' attitudes towards mathematics generally were examined to see whether they did in fact decline over the period of transition. Students were asked to relate to statements concerning their attitude towards mathematics, perceived importance of mathematics, persistence at and understanding of mathematics, which were ranked on a four point Likert-type scale ranging from very true to not at all true. Examples of such items were: 'I give up working on maths problems when I can't understand them', and, 'I usually understand the work we do in maths'. It was found that for the control group there was no statistically significant increase in general attitude towards mathematics as it proceeded through the first year of secondary school, and the decrease in the experimental group was small and not statistically significant.

Gender was examined to see whether the changes that occurred in the self-efficacy beliefs of the female students differed from those of the male students. It was found for the control group, that none of the means in self-efficacy, perceived control, reaction to challenge, task specific confidence or general attitude was significantly different by gender. On the other hand,

in the comparison of engagement by gender there was found to be a significant difference with the females showing more decline in engagement in mathematics as they proceeded through the first year of secondary school.

It was found for the experimental group that none of the means in self-efficacy, perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On the other hand, the comparison of task specific confidence by gender was found to be significant. Males showed more positive growth than females in task specific confidence in mathematics as they proceeded through the first year of secondary school.

It was found for the combined groups that none of the means in perceived control, engagement, reaction to challenge or general attitude was significantly different by gender. On the other hand, the comparison of self-efficacy and task specific confidence for gender was found to be significant. Males showed more positive growth than females in self-efficacy and task specific confidence in mathematics as they proceeded through the first year of secondary school.

From analysis of the questionnaire data in chapter 6 it was concluded there was no significant change in the self-efficacy beliefs of the control group. By contrast, there was a significant gain in the self-efficacy beliefs of the experimental group, strongly suggesting that the Direct Instruction intervention was effective.

In chapter 7 the results of the student responses to the mathematics assessment were examined to determine growth in student knowledge and understanding of mathematics as students proceeded through the first year of secondary school. Data analyses were conducted to examine any difference in performance between Stage 1 and Stage 2. Comparisons between the control and experimental group of students were presented as well as comparisons within like-school groups. Gender differences in mathematics achievement were also investigated.

Since it was recognised that the effectiveness of an intervention in any study is best determined by identifying, examining and comparing student achievement over a period of time, only the students who participated in both stages of the study were used in the analysis. Descriptive statistics of the groups for each stage of the study were calculated and t-tests were used for pairwise comparisons.

Though a variety of factors influence the level of student achievement at any point in time, it was decided that only the two most important; gender and socioeconomic status, would be investigated. To this end like school group comparisons in student achievement were carried out. Furthermore, the sample comprised a range of like school groups drawn from across the Melbourne metropolitan student population so that the results could be deemed representative.

In identifying the differences in mathematics achievement of the control group and the experimental group between Stage 1 and Stage 2 with a view to determining whether students' growth in knowledge and understanding had taken place, it was found that the control group's

achievement mean increased from 26.50 to 28.90 while the experimental group's mean increased considerably more, from 24.77 to 28.73.

In examining the differences in mathematics achievement between the control group and the experimental group it was found that achievement in mathematics improved significantly for both the control and experimental groups. The difference in the means for these two groups at Stage 1 was statistically significant in favour of the control group; the difference in the means for the two groups at Stage 2 was not statistically significant. This revealed a considerable improvement in the experimental group overall.

In determining what effect, if any, socioeconomic factors had on mathematics achievement it was found that when compared with all like school groups, the most growth in mathematics achievement appeared in the experimental like school group 9 (those most at risk in this study) whilst the growth of the control students in that like school group was the least. Similarly resultant growth to that of experimental like school group 9 was achieved by experimental like school group 6, the second to most at risk like school group in the study. On the other hand it was found that the control group results were directly in line with the general trend of the PISA (Lokan et al., 2001) study where it was demonstrated that students with lower levels of SES were more likely to have lower achievement levels.

Another major background variable in the PISA (Lokan et al., 2001) study with respect to achievement was stated to be socioeconomic status based on the occupations of the parents. However, in this study, students from the experimental group whose parents had the highest incidence of

management and professional roles made the least gains of the experimental group.

In determining what effect, if any, gender differences had on mathematics achievement the focus in this study was on the differences existing across groups, that is, the combined control and experimental groups, and within each group. It was found that although females overall made more progress than males across groups (combined, control and experimental), none of the mean score gains was significantly different by gender. This was in line with the TIMSS (Zammit et al., 2002) finding that gender differences in mathematics performance have almost disappeared. The comparison by gender within the experimental group was found to be at equivalent levels. This was in line with the Australia TIMSS finding.

The results of the like school groups were examined by gender, girls vs girls, boys vs boys. It was found for females within like school groups 6, 2 and 5 that none of the means was significantly different. On the other hand, within like school group 9, the group most at risk, there was a significant difference in favour of the experimental group females. It was found for males within like school groups 2 and 5 that the mean was significantly different in favour of the control group. On the other hand, for males within like school groups 9 and 6, the groups most at risk, there was a significant difference in favour of the experimental group. It was noted (see p. 194) that these findings were in direct contrast with PISA (Lokan et al., 2001).

In the light of the findings in chapter 7 it was concluded that the experimental intervention had a positive effect on mathematical

achievement in the experimental group and more especially with the students traditionally regarded as those most at risk.

Chapter 8 reported on the relationship between student self-efficacy beliefs and mathematics achievement. More explicitly it explored the actual effects students' self-efficacy has had on their mathematics achievement over time. Data analyses were carried out to explore this relationship, and comparisons between the control and the experimental groups of students were examined as well as comparisons within each group. Low ability students' differences were also investigated.

The research questions in this study that guided the investigation into the relationship between student self-efficacy beliefs and achievement were stated to be as follows:

What are the relationships between self-efficacy and mathematics achievement?

Are these the same for the control and experimental groups?

Since there is no single method capable of assessing the impact student self-efficacy beliefs may have on mathematics achievement over time, the data analysis in this study was shown to be to a large extent exploratory.

Descriptive statistics of student change in self-efficacy beliefs between each stage of the study were calculated (the mean difference between Stage 1 and Stage 2) and plotted against a mathematics achievement scale (derived from the Stage 2 mathematics achievement data) to see if a relationship existed between the two variables.

Though a variety of factors were seen to influence the level of student achievement the focus in this analysis was to be on the relationship between self-efficacy beliefs and mathematics achievement. And the gathered data were analysed in an attempt to gauge the extent to which these self-efficacy beliefs affected student achievement. Accepting that self-efficacy beliefs impact on student achievement, a mathematics achievement scale was devised and comparisons with mathematics achievement and change in self-efficacy beliefs was undertaken to determine the association.

Stanines were regarded most appropriate for reporting results in broad terms and considered suitably precise for all practical purposes in this study. The calibration of the Stage 2 mathematics data into stanines allowed a mathematics scale to be constructed across the following three categories. Students having the greatest mathematics achievement, those in stanine 6, 7, 8, and 9, were placed in the high category (3) of the scale. Students having the least mathematics achievement, those in stanine 1 and 2 were placed in the low category (1), and those in stanine 3, 4, and 5 were placed in the middle category. Within these categories, the analysis first compared the control group's relationship between self-efficacy and mathematics achievement, then the three categories within the experimental group. Next, each group's relationship between task specific confidence and mathematics achievement was explored, and finally, the lowest performing students' relationship between self-efficacy and mathematics achievement was investigated.

For the control group, the data showed that while all ability levels displayed growth in mathematics knowledge and understanding there was little

change in their self-efficacy beliefs. It was deemed likely, however, that these students started out with inflated views on their abilities and their abilities simply caught up with their expectations. It was seen Fullarton (1998) reached a similar conclusion reporting 'both males and females moderated their expectations of themselves over the transition to secondary school, perhaps reflecting a more realistic idea about what they could achieve' (p. 143). Similarly it is thought that the experimental group started out with inflated expectations, however, in direct contrast, their change in self-efficacy beliefs occurred across all ability groups with the most gains appearing in those students of high ability.

In line with the literature reviewed herein and the researcher's expectations it was seen that within the experimental group, the low ability group (category 1) showed the least growth in task specific confidence. By way of contrast, within the control group, the low ability group (category 1) showed the most growth in task specific confidence, a finding which does not fit the research trend and one the researcher, cannot explain.

In the comparison of the least achieving students in the study, stanine 1 control students with stanine 1 experimental students, the experimental group showed gains across both the self-efficacy variable and the task specific confidence variable, whereas the control group showed a decline in both.

The experimental group's data for this study showed that across all ability levels students with more firmly held self-efficacy beliefs in mathematics achieved more highly, confirming previous research findings (for reviews see Bandura, 1986; Pajares, 1996), and highlighting the value of self-

efficacy beliefs in the prediction of students' mathematics performance. The highest levels of achievement in mathematics were seen to be in those students with the most gain in self-efficacy beliefs. These data also backed up the Direct Instruction research (Adams & Engelmann, 1996) findings confirming that Direct Instruction was effective not only for regular students but also for those at risk. It was regarded as noteworthy that the most gains in task specific confidence were seen in the experimental group. Further, the least achieving (stanine 1) students in this group made considerable gains in their self-efficacy beliefs and task specific confidence while for the control group it was the reverse.

As shown in chapter 3 the most influential source of self-efficacy information is personal performance attainment because it is based on personal experience of mastery. The intervention (the experimental treatment in this study) required the students monitor their progress towards mastery of academic goals on a daily basis. This had the effect of modifying students' self-efficacy beliefs so that as they attained their goals they realised they could actually perform the tasks they were being set and this had the effect of enhancing their confidence for future tasks.

In chapter 9 the student responses to the self-ratings in mathematics and short answer items were examined to gain further insight into their attitude towards learning with a particular focus on the mathematics lesson. At Stage 1, students overall rated themselves as having slightly above average ability at mathematics and on being questioned how good they would like to be the male students overall aspired higher than the female students. It was seen that over the transition from primary to secondary school both male and female students lowered their expectations though

the male students' expectations stayed higher and they were more positive that they would get higher ratings from teachers, parents and peers.

At Stage 1, 32 per cent of students said that they liked mathematics, and it was noteworthy that most of the remainder expressed concerns about fractions and decimals. The 38 per cent of students who did not like mathematics fell into two sub categories: anxiety and self-image. Again it was noteworthy, and concerning, that many comments involved fear, embarrassment and ridicule.

About 60 per cent of students thought maths classes would be more difficult at secondary school with an alarming number voicing low opinion of secondary teachers.

At Stage 2, 22 per cent of the control group did not like mathematics and 14 per cent of the experimental group expressed a similar opinion. Overall, 52 per cent of control group liked mathematics and 56 per cent of the experimental group expressed a similar opinion.

In describing, 'Compared to primary school how do you think you are doing in maths?' overall, 8 per cent of control students thought they were doing worse while 7 per cent of experimental students expressed a similar view. Overall, 75 per cent of control students thought that compared to primary school they were doing better in mathematics with 82 per cent of the experimental students expressing a similar view.

Discussion

A major aim of this research was to evaluate the effectiveness of a Direct Instruction intervention on the mathematical achievement of Year 7 students.

At the pretest stage the investigation found there was a significant difference in mathematics achievement between the control and experimental groups in favour of the control group. This indicates that the groups were not well matched in mathematical ability. Although the control group performed significantly better at this stage the advantage was substantially diminished at the posttest stage. While both groups made statistically significant gains in mathematics achievement, the fact that the posttest revealed no significant difference between the groups, highlights the particularly significant result overall for the experimental group. The advance made by the experimental group in mathematics achievement was particularly significant and clear evidence of the positive impact the Direct Instruction intervention had on their mathematical ability.

The strength of this finding is more pronounced in the evaluation of the like school group analysis. At the pretest stage investigation of the students at the lower socioeconomic level (like school group 9), those most at risk in this study, found a significant difference in mathematics achievement between the control and experimental groups in favour of the control group. This indicates that the groups were not well matched in mathematical ability. While the control group performed significantly better at this stage the advantage was substantially diminished at the posttest stage. Although both groups made statistically significant gains in

mathematics achievement, the fact that the posttest revealed no significant difference between the groups, highlights the particularly significant result overall for the experimental group. In fact the gain for this group was the biggest overall in the study. The advance made by the experimental group in mathematical achievement was significant and clear evidence of the positive impact the Direct Instruction intervention had on their mathematical ability.

Schools belonging to like school group 9 have large proportions of students for whom the main language spoken at home is not English as well as many who receive the Education Maintenance Allowance or Commonwealth Youth Allowance. This combination of educational and demographic factors places demands on these students resulting in them being considered most at risk. The findings in this study strongly suggest that these students do not need to fail academically. The study shows they can be taught. The Direct Instruction intervention had a consistently positive academic effect on this group, a finding directly in line with those of the National Evaluation of Follow Through (Abt Associates, 1977). In that enormous study a planned variation design was employed 'to provide a broad-range comparison of educational alternatives for teaching the disadvantaged and find out "what works" ' (Becker & Engelmann, 1996, p. 33). As was seen in the current study's literature review, the final report of the National Evaluation of Project Follow Through showed that the Direct Instruction Model (University of Oregon) was most successful in assisting disadvantaged children in catching up with their middle-class peers in academic skills. The results on display in Figure 20 illustrate graphically the abysmal results achieved by all students in the study except those taught by Direct Instruction. In their report on the sponsor

findings from Project Follow Through, Becker and Engelmann argue ‘The popular belief that it is necessary to teach different students in different ways is, for the most part, a fiction. The requirements for sequencing an instructional program are determined by what is to be taught, not who’ (Becker & Engelmann, 1996, p. 39).

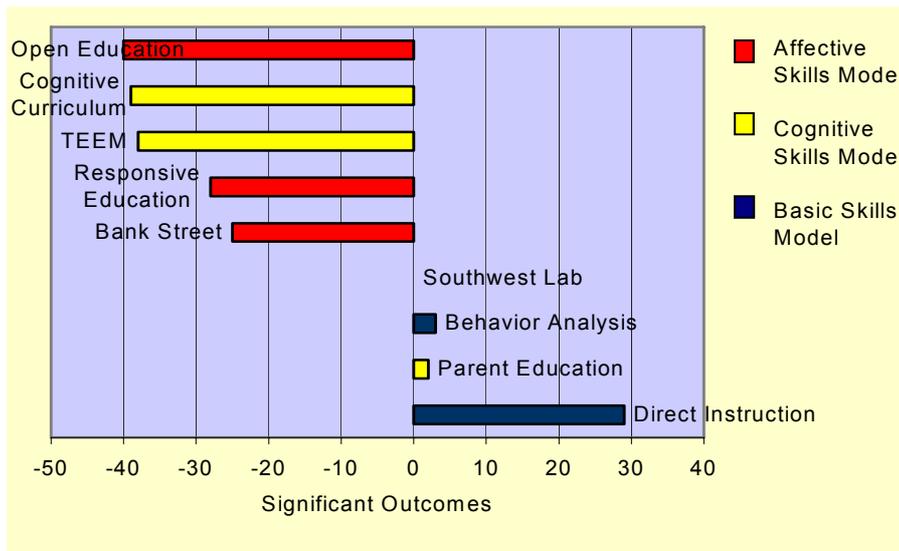


Figure 20 Significant outcome comparisons of basic skills achievement across Follow Through models

Note. Adapted from Adams & Engelmann, 1996, p. 72. Follow Through models’ scores are compared to control groups — if the Follow Through model scored higher than the control group on a variable, then the index was a positive number. If the control group scored higher the index was negative. If there was no difference between the two groups the score was zero.

Critics of Direct Instruction predict that most Direct Instruction achievement gains will disappear over time. It is their contention that Direct Instruction students are spoon fed through lessons with instructions that are easy to follow and they receive significant reinforcement not found in the normal classroom. They argue that once these students go into the normal classroom where they no longer receive Direct Instruction any gains they have made will dissipate.

Naturally enough Direct Instruction supporters think differently, they give different reasons why Direct Instruction results might decrease over time. They say, because Direct Instruction students are taught more during available time than those in a traditional program, upon leaving the Direct Instruction classroom they would be provided with relatively fewer opportunities to learn and would naturally enough learn less.

Subsequent studies on Direct Instruction Follow Through students, however, showed both camps were wrong. The Direct Instruction students at both 'primary and secondary levels, show strong continuing effects in terms of academic performance at the primary level, and better attendance, fewer grade retentions, and increased college acceptance at the high school level' (Becker & Engelmann, 1996, p. 41).

At the pretest stage in the like school group 6 analysis, those second to most at risk in this study, there was no significant difference in mathematics achievement between the control and experimental groups. This indicates that the groups were well matched in mathematical ability. This was not the situation at the posttest stage. Although both groups made significant gains in mathematics achievement, the posttest revealed a significant difference between the control and experimental groups showing a particularly significant result overall in favour of the experimental group. The particularly significant advances made by the experimental group in mathematical achievement is clear evidence of the positive impact the Direct Instruction intervention had on the mathematical ability of these students.

The strength of this finding was diminished in the combined like school group 2 and 5 analysis. At the pretest stage investigation of these students constituting the middle socioeconomic level (those least at risk in this study) showed there was no significant difference on mathematics achievement between the control and experimental groups. This indicates that the groups were well matched in mathematical ability, though this was not the case at the posttest stage. Although both groups made significant gains the posttest revealed a significant difference between the control and experimental groups, showing a particularly significant result overall in favour of the control group. The significant advances in mathematical achievement made by the experimental group were eclipsed by those of the control group. As previously adverted to the experimental group may have been disadvantaged as the control group, which in this case consisted of just one school, was known to have a history of above average achievement and the post testing for the experimental group was carried out in less than desirable conditions.

For the sake of integrity further comparison excluding the groups least at risk, that is like school group 2 and group 5, was undertaken. At the pretest stage in the combined like school group 9 and 6 analysis, there was no significant difference on mathematics achievement between the control and experimental groups, indicating that the groups were well matched in mathematical ability (they were in fact the best match overall). This was not the situation at the posttest stage. Although both groups made significant gains in mathematics achievement, the posttest revealed a significant difference between the control and experimental groups showing a particularly significant result overall in favour of the

experimental group. The particularly significant advances made by the experimental group in mathematical achievement is clear evidence of the positive impact the Direct Instruction intervention had on the mathematical ability of this mix of at risk and regular students.

Mixed-ability grouping has been the subject of much criticism largely 'because it has failed to stretch high ability pupils sufficiently and is not always able to provide sufficient support for the less able' (Hallam & Toutounji, 1996, p. 1). However, the findings in this study clearly show that the benefits of the experimental Direct Instruction intervention in the regular classroom were beneficial to students of all abilities. These results are consistent with other research using Direct Instruction programs in regular classroom populations (e.g., Sexton, 1989; Tarver & Jung, 1995; Vitale & Romance, 1992). Tarver and Jung (1995) found 'Direct Instruction is much more likely to meet the needs of diverse students grouped together in the regular classroom than are constructivist teaching approaches' (p. 56). In light of the findings in the current study this writer concurs with Tarver and Jung (1995) in calling for a more responsive approach to the identification and promotion of effective instruction: 'Results should speak louder than rhetoric' (p. 56).

Another major aim of this research was to evaluate the effectiveness of a Direct Instruction intervention on the mathematical self-efficacy of Year 7 students.

At the pretest stage the investigation found there was a significant difference between the control and experimental groups' mathematical self-efficacy in favour of the control group. This indicates that the groups

were not well matched in their self-efficacy beliefs. Although at this stage the control group possessed significantly more positive self-efficacy beliefs on a range of dimensions the advantage was not maintained at the posttest stage. The posttest revealed a significant difference between the control and experimental groups' mathematical self-efficacy beliefs in favour of the experimental group, showing a particularly significant result overall for this group. There was no change in the mathematical self-efficacy beliefs of the control group. The significant advances made by the experimental group provides clear evidence that the Direct Instruction intervention had a positive impact on their self-efficacy beliefs.

It was stated in chapter 2 that some Direct Instruction critics thought the tightly controlled instruction inherent in Direct Instruction might discourage children from expressing themselves freely and consequently have a detrimental effect on their self-esteem. With due respect, what little research has been done in this area indicates precisely the contrary. Although there are few studies that have measured affective variables in relation to Direct Instruction, those few that have, report it as having had positive effects on self-concept and attitudes towards learning. Further, the evaluation of Project Follow Through (Abt Associates, 1977) shows the Direct Instruction Model (University of Oregon) performed best on measures of self-esteem. It specifically found that the child-centred models 'that emphasized improving students' self-esteem produced students with the poorest self-esteem' (Adams & Engelmann, 1996, p. 84). Though educators often accuse Direct Instruction of ignoring the whole child by concentrating on academic achievement at the expense of personal development, the Follow Through data clearly show this is nonsense (see

Figure 21). The Direct Instruction model was actually found to be more effective than any other in elevating student self-esteem. Bearing this out, the results in the current study show that the Direct Instruction intervention had the effect of substantially elevating students' self-efficacy beliefs across the entire spectrum of ability levels.

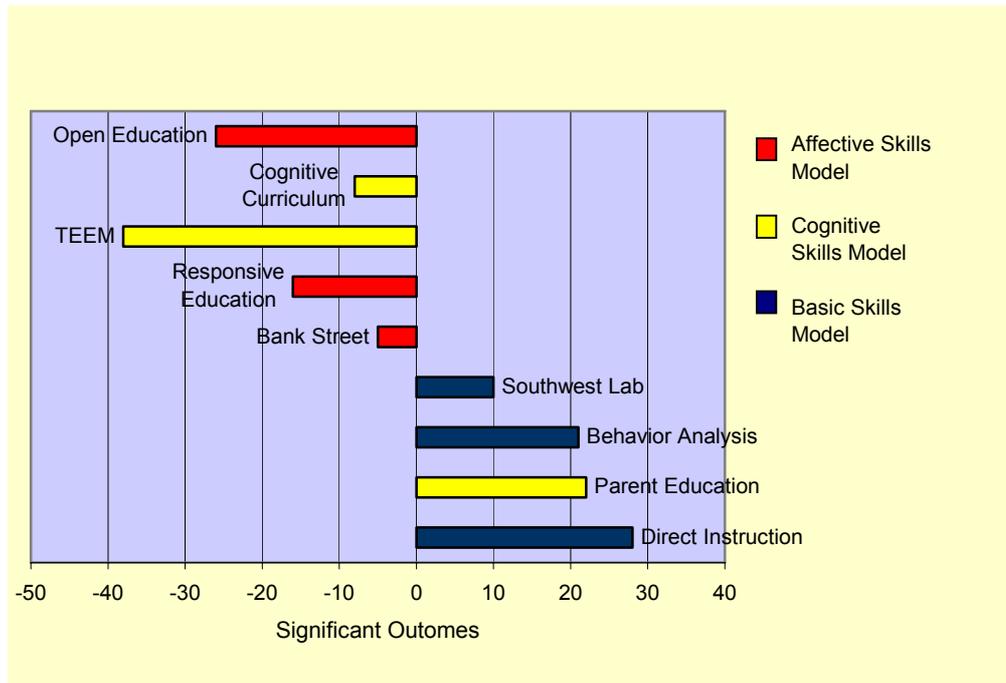


Figure 21 Significant outcome comparisons of self-esteem across Follow Through models

Note. Adapted from Adams & Engelman, 1996, p. 72. Follow Through models' scores are compared to control groups. If the Follow Through model scored higher than the control group on a variable, then the index was a positive number. If the control group scored higher the index was negative. If there was no difference between the two groups the score was zero.

It is appropriate here to further discuss certain constructs (underpinning the experimental intervention) that have been shown in many research studies to be instrumental in elevating student self-efficacy. The following being some of the more important constructs: specific learning goals, instructional models, learning mastery and feedback. Research has also

shown that the information acquired from the application of these principles helps motivate students strive towards their personal best.

It was stated in chapter 3 that while students engage in the learning process they automatically monitor their progress towards academic goals and in the doing it is important for them to see progress. With the experimental Direct Instruction intervention, self-evaluation of progress was an integral and ongoing component. By providing tasks that gradually increased in difficulty the intervention provided students with clear criteria, which allowed them to assess their performance and gauge their progress independently. As the students progressed they acquired more skills and became more proficient at the self-evaluation process.

The findings in this study support the research literature reviewed in chapter 3 showing clearly that goal setting and self-efficacy are important factors in academic achievement. By providing short term goals that were specific and challenging yet attainable the Direct Instruction intervention actually enhanced students' self-efficacy. As was seen in chapter 3 setting students long term goals has met with little success. On the other hand students actually believe they can attain short term goals because they provide them with clear standards against which they can measure their progress in the immediate future. Indeed, it seems the mere whiff of progress may be enough — as Schunk (1999) informs us, the perception of progress strengthens self-efficacy and motivates students to continue to improve.

In chapter 3 it was stated that by watching models successfully perform tasks, observers not only acquire particular behaviours and skills but pick

up important information with respect to correctly sequencing actions. The experimental intervention required the teacher to actually model each skill and concept as it was being introduced thus ensuring students learned correctly. In chapter 2 it was stated that the exponents of student-directed learning expect students to construct their own procedures and take responsibility for their own learning. While the writer accepts there is merit in this approach it is, of course, greatly debatable how much children are actually capable of discovering for themselves. And, how long should this discovery process take? If students have reached the transitional door without the basic mathematical skills in place then surely time is up? The findings in the current study support the wealth of research findings showing that by observing models demonstrate skills, students' learn and through this learning their self-efficacy and academic achievement are both elevated.

As reported in chapter 3, Bandura (1986) argues that experiences of mastery are the most influential sources of our information on self-efficacy. Naturally, students retain information better when repetition and reinforcement is embedded into the learning process and distributed over a period of time. It was by exposing students to direct explanations and problem tasks in incremental portions, and providing them with systematic review of their work on a daily basis, that the experimental intervention led them to task mastery.

It was seen in chapter 3 that Schunk (1982a) argued feedback of prior performance elevated self-efficacy expectations and that this was partially responsible for increased skill in performance. In the writer's opinion assessment is an integral part of the learning process and implemented

effectively, provides a reliable basis for teacher feedback. 'Ideally, testing is part of teaching' (Adams & Engelmann, 1996, p. 14). It is a fact of life that very few of us enjoy sitting down to tests or examinations. With this in mind the lessons in the experimental intervention were specifically designed so as not to be seen by the students as formal tests. The experimental group's student responses, however, provided the teacher with reliable daily diagnostic information similar to that which could only be acquired from a formal test situation. This was important because by giving the students daily feedback on their performance they became aware of their progress, which strengthened their self-efficacy, sustained their motivation and enhanced their academic achievement. As students engaged in the experimental program they learned which actions produced positive results and were thus provided with a guide for future lessons. It was also found the anticipation of desirable results motivated these students to persevere.

In summation the research questions posed in Chapter 1 of this study have been answered as follows:

The students taught with the Direct Instruction intervention developed a higher mathematical self-efficacy than the students taught without it.

The at risk students taught with the Direct Instruction intervention developed a higher mathematical self-efficacy than the at risk students taught without it.

With the exception of the group least at risk in the study, the overall growth in knowledge and understanding in mathematics of the students taught

with the Direct Instruction intervention exceeded that of the students taught without it.

The growth in knowledge and understanding in mathematics of the at risk students taught with the Direct Instruction intervention exceeded that of the at risk students taught without it.

While a relationship was found to exist between students' self-efficacy and students' achievement in the students taught with the Direct Instruction intervention, this was not the case with the control group. Thus, the accumulated data do not allow the researcher to satisfactorily answer the question: What relationship, if any, will exist between students' self-efficacy and students' achievement? The researcher can only suggest that the approach adopted in this study was perhaps unsuitable for this particular purpose.

Bandura (1997) makes the proposition that 'Education has now become vital for a productive life' (p. 213) and in today's increasingly technologically reliant society it is impossible to argue with him. Societies pay a high price for faulty or neglectful education and poorly educated children are inevitably the first casualty of unemployment. There is ample evidence showing that unemployment leads to major social maladies such as delinquency and drug abuse, and the fact that our prisons are full of prisoners doing time for drug related crime may thus be seen to relate directly back to fault lines in the education system.

For teachers to successfully educate they must be aware of certain factors crucial to teaching which are perhaps best illustrated by analysing the role

self-efficacy plays in the learning process. The notion that students act on their perceived capability has important implications for classroom practice and this is discussed in the following section. The potential drawbacks associated with any educational innovation are also discussed.

Implications for teaching

The sources of self-efficacy have been seen to be, performance attainments, vicarious experiences, verbal persuasion, and physiological states. And it is important teachers understand the circumstances in which these constructs operate most favourably in the development of self-efficacy. By far the most influential of these sources is personal performance attainment and this is because it is based on personal experience of mastery. As students attain their goals they see they are capable of performing certain tasks and their confidence with respect to future learning is enhanced. Their performances provide dependable data on which to base their self-efficacy. Teachers should be aware, however, that while repeated success establishes a strong sense of self-efficacy, the occasional failure is not likely to impact much upon one's perception of one's abilities.

Where students lack personal experience in learning a particular concept, if possible, they look to their peers who have had such experience, using them as models by which they estimate their own capability. Naturally enough, if students see their peers successfully negotiate a task they are inclined to think they will be capable of similar performance, thus the simple knowledge in itself that their peers have performed a task successfully can enhance students' self-efficacy. It is generally accepted

that learning vicariously speeds up the learning process and can have the effect of shielding the learner from negative experiences. Teachers should thus consciously seek out opportunities for students to assess peer performance in non-competitive learning situations.

Verbal persuasion is often used as a means of inducing people to think they are capable of doing certain things, but as a means in itself of engendering self-efficacy it is of limited potency. There are certain caveats to verbal persuasion which teachers should be aware of. Verbal persuasion is most effective when students have valid cause to believe they are capable of performing the task. While there is no doubt that thoughtfully employed teacher encouragement can elevate self-efficacy, lifting beliefs of personal competence to unrealistic levels not only invokes failure it actually lowers self-efficacy. Just as positive social persuasion is capable of contributing to successful performance, derogatory comments can have the most detrimental effect, especially if that person is already lacking confidence in the area. As stated earlier, verbal persuasion as a means of engendering self-efficacy should thus be viewed as a delicate instrument to be treated with respect and applied with care.

Whilst physiological indices such as perspiring palms, hollow feelings in the stomach and elevated heart rates may be important indicators of degree of confidence at an individual level, teachers should bear in mind they vary greatly between people. Differing interpretations of arousal have differing effects on self-efficacy perceptions and what makes one person frightened may well see another fired up. Whatever effects arousal may have on self-efficacy, however, derive from past experience.

Students with high self-efficacy beliefs will persevere longer and harder even in particularly difficult circumstances. Conversely students with low self-efficacy are reluctant to engage in tasks where those skills are required and are more likely to quit early. Students need to be made aware that effort is a crucial ingredient in achieving academic success. In chapter 3 of this thesis Fullarton (1998) was seen to make the critical point that 'to foster motivation and enhance perceived control, it is important for children to be made aware of the connections between their efforts and outcomes'. She saw much merit in encouraging students to view their successes as the result of their 'high effort and ability' (p. 214).

Students' feelings of self-efficacy are enhanced as they see their learning progress and this in turn sees their motivational constructs enhanced, though if they feel they can't achieve their goals their motivational constructs will diminish. Teachers should bear in mind that for goals to promote motivation they should be set in a specific context so students can easily gauge their progress. Abstract goals fail to enhance motivational constructs and the same applies for goals set too far in the future.

Peers and authority figures such as parents, adults and teachers are powerful models for students and it is important for teachers not to underestimate the role they play and the effect they have. By observing models behave and seeing the results of that behaviour students know what they can expect to happen if they behave similarly; they can see what requirements are necessary to successfully negotiate the task.

Teachers should also be aware that their own perceived efficacy is important and depends on many factors above and beyond the physical act of teaching. Teacher effectiveness is dependent also upon such things as the ability to provide an environment conducive to learning, impose discipline, maximise time and resources, identify and eliminate negative influences. These are among the more important. Effective educators operate on the understanding that difficult students can be taught as long as the teaching technique appropriate to the circumstances is employed. Ineffective educators, on the other hand, blame the students; they are unmotivated, they are dumb.

The drawbacks in respect of educational innovations are many, not the least being they add to teachers' already heavy workloads. Teachers have become accustomed to seeing much touted educational innovations implode in their own puffery and, naturally enough, have built up a resistance. Unfortunately it takes time and a fair degree of dedication to introduce an educational innovation and it is often aborted before any benefits can be realized with the blame being laid on the program rather than the poor manner of implementation.

The role of the teacher in implementing a Direct Instruction innovation can seem a complex one but whether or not this is so is debateable. The matter really rests on the precise nature and implementation process of the innovation. The most notable feature of Direct Instruction is the high degree of specificity in terms of teaching behaviours and educators quite justifiably talk of the dilemma of explicitness. The writer has already given her opinion that innovations should never be so specific as to disallow teachers from adapting them to suit their own classrooms. However, if the

innovations lack sufficient specificity, even receptive teachers will be discouraged because they don't know precisely what is required of them.

A vital factor in the implementation of any educational innovation is the professional development of staff so as to ensure they understand the nature of the project and its manner of implementation. This is especially essential where the program is to be modified in some way to meet a school or classroom's particular needs, but it seems that teachers are more inclined to adopt a new practice if they have a sense of ownership of the program (McLaughlin, 1976).

Once the program is implemented 'the quality of implementation must be assessed periodically to determine how well the various components are working by themselves and with one another. Such probes identify the needed corrective adjustments to ensure that the program is being implemented productively' (Bandura, 1997, p. 258).

It has already been shown that despite the depth of research supporting the general success of Direct Instruction programs in a wide variety of teaching situations, the approach has not exactly been embraced in the Australian educational arena. A major reason for this, in this writer's opinion, is that most teachers have had little or no hands-on experience with Direct Instruction either as teachers or as students. Hence all the mythology surrounding the Direct Instruction approach such as 'it's only for students in special ed, or disadvantaged kids'. In fact the current study has demonstrably proved this is not the case. Direct Instruction was shown to work with all the children involved and they were spread across a particularly broad spectrum.

Another common misconception is that students simply don't like Direct Instruction. Again the current study showed this was not so with close to 30% of the experimental students nominating the Direct Instruction intervention as that part of the lesson they enjoyed the most. An analysis of the experimental intervention shows that like other Direct Instruction programs it incorporates meticulously planned distributed practice where instead of reliance being placed upon basic or rote skills, the intervention moves quickly from foundational skills to problem solving involving high order thinking.

At the end of the experimental treatment period, the teachers filled out a survey (see Appendix 7) that included an open response comment and questions such as: What have you found positive about the EMM program? What are your concerns about the program? Would you be prepared to trial the program another year? and, Other comments. Teachers were not required to identify themselves and understood that their responses were completely confidential. Therefore, teachers understood that they were welcome to give negative evaluations.

The following responses to the teacher survey provide first-hand information on how teachers who actually implemented the Direct Instruction experimental intervention perceived it.

What have you found positive about the EMM program? 'It's made the students more disciplined, they listen better and have more confidence. They are more willing to have a go'; 'They get excited about EMM, saying, Yes when the lesson gets to problem-solving'; 'They tend to focus as a group and work better together'; 'It helps the low ability students

emotionally because they are part of the group and not isolated'; 'I thought the program definitely eliminated gaps in the students' learning. Made the students consistent'; 'It helped me diagnose problems and watch student learning growth. The students followed directions better, and had better recall of procedures and rules they had learned'; 'Students settle quickly. Their listening skills improve once they realise you will not be repeating the questions. Students enjoy the challenge of trying to get 20'; 'Organised skills reinforcement'; 'Students have shown interest in completing the EMM program'.

What are your concerns about the program? 'At the early stage some of the high ability students become bored'; 'Some concerns about getting through the maths syllabus'; 'Cheating'; 'Not correcting properly'; 'Students absent need to catch-up'; 'Students who join the class through the year are a problem'; 'The program works but I don't believe every child needs it'.

Would you be prepared to trial the program another year?

All experimental schools continued the program in 2003.

Other comments included 'EMM has worked so well with our "middle band" students that in 2003 we will offer two Direct Instruction programs, EMM and Corrective Maths, for our lower ability students'; 'We are thinking of including Direct Instruction in the schools literacy program'.

These comments from the teacher survey go quite some way towards addressing the aforementioned mythology surrounding Direct Instruction because they contain, after all, first-hand information from teachers who

have actually implemented a Direct Instruction program. That said, however, this writer believes teachers who have tried the Direct Instruction approach and remain uncomfortable with it should never be forced to continue. On this issue Bandura is firm, having the following to say on the teaching profession and the implementation of educational innovations, 'Effective social mechanisms must be created for replacing leadership and staff members who remain recalcitrant to essential changes despite substantial offers of assistance' (Bandura, 1997, p. 258).

There follows a discussion of certain issues arising out of this study that have implications for future research.

Implications for research

The key objective of this research was to determine the effects of a Direct Instruction intervention on student learning in mathematics and self-efficacy during the first year of secondary school. It should be noted that research into the effectiveness of any educational program is in varying degree dependent on the quality of implementation. Where the circumstances of learning and the quality of teaching are high, effective learning is assured. Unfortunately, however, these variables are almost always outside the control of the researcher and whilst this is a well recognised problem it is less the case with an intervention which is entirely scripted.

One important direction research could be taken is into the domain of the self-efficacy beliefs of the teacher. In the researcher's opinion teachers' sense of self-efficacy, that is, their perceived capacity to help students

learn, is a critical element in effective teaching. From the literature reviewed in this study it is clear that one of the most important factors in students' academic achievement is their self-efficacy belief. It would follow that the same should apply to teachers. Indeed, Bandura (1993) regards teachers' beliefs in their ability to motivate and promote learning as one of the most important contributors to student achievement. Accepting this to be so, investigation into teachers' beliefs should pose a valuable proposition for educational research. Such research would require comprehensive examination of teacher beliefs involving enquiry into many and varied matters: what teachers say and do and why, what they do and don't believe and why, how they feel about what they do and don't do and why, and so on. If research into teachers' educational beliefs does not give insight into the relationship between these beliefs, and their practices, knowledge, and student outcomes, then in this writer's opinion the research will tell us little about effective teaching practices. As one leading educator comments:

It is easy to urge teacher educators, for instance, to make educational beliefs a primary focus of their teacher preparation programs, but how are they to do this without research findings that identify beliefs that are consistent with effective teaching practices and student cognitive and affective growth, beliefs which are inconsistent with such aims, and beliefs that may play no significant role. (Pajares, 1992, p. 327)

While the current study had its strengths, and yielded a wealth of valuable data in respect of effective teaching practices and the role that self-efficacy plays in the learning process, it was not without its limitations.

Though a variable control component of the design, one limitation of the study was that no Independent or Catholic schools were involved. The research was confined to Government secondary schools drawn from the Education Department's Melbourne metropolitan regions with like school groups selected to control school and student level variance.

It was anticipated in this study that the students' reasons for their differing perceptions and attitudes towards mathematics would become apparent in their responses to the questions related to the actual mathematics lesson. Though the open questions were designed so as not to restrict students' responses it is possible the Stage 2 questionnaire did not accurately reflect the expectations of a student-directed classroom. The researcher is aware that some mathematics classrooms in this study had features associated with student-directed learning, for example, independence in learning, students exploring and attempting to solve problems, yet this did not appear in the students' perceptions and attitudes as expressed in their open responses. In a student-directed classroom students are constructing their understanding through experiences, thus, learning may often be hidden with students remaining completely unaware that learning has actually taken place. It may therefore be that the questionnaires employed in this study, which were developed for traditional classrooms, were not ideally suited to the modern classroom where constructivist approaches and curricula are employed.

While in this study purely quantitative data were collected, the researcher now recognises that qualitative data would have provided valuable additional insights into the physical activity of the mathematics lesson. And it is clear that a more intensive focus on individual students through

interviews and the actual observation of them whilst engaged in mathematics lessons, would have provided this. Ultimately, however, whether one opts for a quantitative, qualitative or combined approach to one's research depends on the objectives of the study and it is much easier to appear wise with the benefit of hindsight.

It is clear from the results of this study that when students are taught explicit strategies and teachers are given clear directions, all students learn. With the regular classroom in mind, it is this writer's submission that future research should examine different instructional sequences for teaching explicit strategies to ascertain which are the best at equipping our students with the broadest and best possible range of skills and knowledge.

Concluding words

From the research and literature reviewed in this study it is apparent that the domain of mathematics teaching in the western world is in a sorry state and there is no consensus on how this should be addressed. Almost all of the problems associated with student learning relate back to curriculum and teaching method and unfortunately, these problems are compounded in the transition years. After examining the literature on student-directed approaches to learning alongside that relating to Direct Instruction it was concluded that the empirical data heavily favoured the latter as being the more effective, yet almost every teacher education program in Australian universities is based on the student-directed approach. One has to ask why this is so, and it is perhaps appropriate here to consider Carnine (2000), when he makes the point that 'the best way for a profession to

ensure its continued autonomy is to adopt methods that ensure the safety and efficacy of its practices' (p. 13). And it is difficult not to sympathise with Finn's concern that 'so much of what passes for education research serves to confuse at least as much as it clarifies. The education field tends to rely heavily on qualitative studies, sometimes proclaiming open hostility towards modern statistical research methods' (2000, Forward). While it is obvious from the Follow Through aftermath that scientific research and the classroom are yet to come together, this writer strongly argues that until they do the schoolchildren of the western world will continue to receive an inferior education.

As stated in chapter 3 the notion that students act on their perceived capability has important implications for classroom practice and this is especially so with transition students. Given this important role it is crucial educators look to teaching methods that will increase students' self-efficacy and it is hoped the measured outcomes of this study can make some contribution towards this end. By determining how students come to estimate their ability at this critical juncture, we will be better able to work on improving their self-efficacy beliefs. While the current study clearly showed that self-efficacy beliefs play a powerful role in the learning process other studies have gone even further placing it on the same level as general mental ability, long recognised as the most powerful factor in the prediction of academic performance.

From the research and literature reviewed herein it is evident that in the field of mathematics there is a place for both teacher-directed and student-directed learning and that a balance should be struck between the two, acknowledging some skills are better acquired through one approach and

some through the other. When it comes to the employment and cultivation of higher order skills where reasoning and reflection are required it seems a constructivist approach would be more suitable. However, with the acquisition of basic mathematical skills it is argued teacher-directed learning is better suited, and it was with a view to this prospect that the experimental instrument employed in this study was designed. It was stated in chapter 3 that if children still lack basic mathematical skills by the time they reach the transition years then howsoever they have been taught hasn't worked. It is to be hoped the conclusions drawn in this study will go some way towards showing how well suited the Direct Instruction approach is for this crucial mathematical domain. The researcher is further hopeful these findings will help promote a better understanding of the respective circumstances in which both student-directed and teacher-directed approaches to learning can be best employed in the education of our children.

In conclusion it seems to this writer there is a dire urgency for the academics of the education world to put less emphasis on the ideology they feel most comfortable with and have a long hard look at the evidence. In the light of the research reviewed in this thesis it is impossible to deny the need for structured teaching in certain important circumstances just as it is impossible to deny the potential benefits to be had from student-directed learning in appropriate circumstances. If we are to provide the children of this nation with the best possible education, clearly, a balance must be achieved between teacher-directed learning and student-directed constructivist approaches — and for the children's sake it must be achieved soon. Further, it is submitted that the *fitness for purpose* principle

enunciated earlier in chapter 2 should be the guiding light when it comes to setting that balance. Unfortunately there is no *one stop shop* — no panacea when it comes to education — it just isn't that simple.

Appendices

- Appendix 1 Approval: Standing Committee on Ethics in Research Involving Humans
- Appendix 2 Response to the Standing Committee
- Appendix 3 Response to Standing Committee Approval, 10 December
- Appendix 4 Permission from the Department of Education, Employment and Training (DEET) to conduct research in Government schools
- Appendix 5 Explanatory statements and permission slips
- Appendix 6 Optical mark reader answer sheet
- Appendix 7 Teacher survey
- Appendix 8 Stage 1 student questionnaire
- Appendix 9 Stage 2 student questionnaire
- Appendix 10 Intervention sample lesson

Appendix 1

Approval: Standing Committee on Ethics in Research Involving Humans



5 December 2001

Dr. Barbara Clarke
Education
Peninsula Campus

Ms. Rhonda Farkota
ACER
Private Bag 55
Camberwell 3124

Re: Project 2001/577 - Effects of direct instruction on self-efficacy and achievement in mathematics

The above submission was approved by the Standing Committee on Ethics in Research Involving Humans at meeting A8/2001 on 4 December 2001 provided that the following matters are satisfactorily addressed:

- If a parent does not consent to their child's results being used in the research, how will the researcher ensure that results are not included.

The project is approved as submitted for a three year period and this approval is only valid whilst you hold a position at Monash University. You should notify the Committee immediately of any serious or unexpected adverse effects on participants or unforeseen events that might affect continued ethical acceptability of the project. Changes to the existing protocol require the submission and approval of an amendment. Substantial variations may require a new application. Please quote the project number above in any further correspondence and include it in the complaints clause which may be expressed more formally if appropriate:

You can complain about the study if you don't like something about it. To complain about the study, you need to phone 9905 2052. You can then ask to speak to the secretary of the Human Ethics Committee and tell him or her that the number of the project is _____. You could also write to the secretary. That person's address is:

*The Secretary
The Standing Committee on Ethics in Research Involving Humans
PO Box No 3A
Monash University
Victoria 3800
Telephone (03) 9905 2052 Fax (03) 9905 1420
Email: SCERH@adm.monash.edu.au*

Continued approval of this project is dependent on the submission of annual progress reports and a termination report. Please ensure that the Committee is provided with a report annually, at the conclusion of the project and if the project is discontinued before the expected date of completion. The report form is available at <http://www.monash.edu.au/resgrant/human-ethics/forms-reports/index.html>.

The Chief Investigators of approved projects are responsible for the storage and retention of original data pertaining to a project for a minimum period of five years. You are requested to comply with this requirement.


Ann Michael
Human Ethics Officer
Standing Committee on Ethics
In Research Involving Humans


RESEARCH GRANTS AND
ETHICS BRANCH
PO Box 3A
Monash University
Victoria 3800, Australia
Telephone: +61 3 9905 3012
Facsimile: +61 3 9905 3851
Email: offres@adm.monash.edu.au
www.monash.edu.au
ABN: 12 377 614 012

Appendix 2

Response to the Standing Committee



ACN 004 398 145
ABN 19 004 398 145

Australian Council for Educational Research
19 Prospect Hill Road (Private Bag 55), Camberwell, Vic 3124
Phone: (03) 9277 5555 Fax: (03) 9277 5500
+61 3 9277 5555 +61 3 9277 5500 (International)
Web site: <http://www.acer.edu.au>



FACSIMILE

3 6 3

TO: Ann Michael

FAX NO. 9905 1420

FROM: Rhonda Farkota

DATE: 7 December 2001

Number of pages to follow:

Ref:

Dear Ms Michael,

RE: Project 2001/577 – Effects of Direct Instruction on self-efficacy and achievement in mathematics.

Thank you for your letter dated 5 December.

In reference to your point regarding:

- If a parent does not consent to their child's results being used in the research, how will the researcher ensure that results are not included.

I wish to advise that cross-checking will be carried out by the researcher to remove the data for students who do not have consent to participate. An independent observer will verify this process.

I hope this is satisfactory. I would be most grateful if you could let me know by telephone or email as soon as possible as I need to contact schools and deliver materials to them before the Christmas break.

Yours sincerely

Rhonda Farkota

Tel 9277 5627

Email farkota@acer.edu.au

Appendix 3

Response to Standing Committee Approval, 10 December



10 December 2001

Dr. Barbara Clarke
Education
Peninsula Campus

Ms. Rhonda Farkota
ACER
Private Bag 55
Camberwell 3124

Re: Project 2001/577 - Effects of direct instruction on self-efficacy and achievement in mathematics

Thank you for the information provided relating to the changes as requested by the Standing Committee on Ethics in Research Involving Humans.

This is to advise that the amendments have been approved and the project may proceed according to the approval as given on 4 December 2001.

2 Ann Michael
Human Ethics Officer
Standing Committee on Ethics
in Research Involving Humans



RESEARCH GRANTS AND
ETHICS BRANCH
PO Box 3A
Monash University
Victoria 3800, Australia
Telephone: +61 3 9905 3012
Facsimile: +61 3 9905 3831
Email: offres@adm.monash.edu.au
www.monash.edu.au
ABN: 12 377 614 012

Appendix 4

Permission from the Department of Education, Employment and Training (DEET) to conduct research in Government schools



Department of Education, Employment and Training

Office of Schools

2 Treasury Place
East Melbourne, Victoria 3002
Australia
GPO Box 4367
Melbourne, Victoria 3001
Australia
Telephone +61 3 9637 2000
DX 210083

SOS001991

4 December 2001

Ms Rhonda Farkota
Private Bag 55
Camberwell 3124

Dear Ms Farkota

Thank you for your application of 22 October 2001 in which you request permission to conduct a research study in government schools titled: *The effects of a Direct Instruction intervention in the regular mathematics class on student's mathematical self-efficacy and achievement.*

I am pleased to advise that on the basis of the information you have provided your research proposal is approved in principle subject to the conditions detailed below.

1. Should your institution's ethics committee require changes or you decide to make changes, these changes must be submitted to the Department of Education, Employment and Training for its consideration before you proceed.
2. You obtain approval for the research to be conducted in each school directly from the principal. Details of your research, copies of this letter of approval and the letter of approval from the relevant ethics committee are to be provided to the principal. The final decision as to whether or not your research can proceed in a school rests with the principal.
3. No student is to participate in this research study unless they are willing to do so and parental permission is received. Sufficient information must be provided to enable parents to make an informed decision and their consent must be obtained in writing.



Appendix 4 (cont'd)

4. As a matter of courtesy, you should advise the relevant Regional Director of the schools you intend to approach. An outline of your research and a copy of this letter should be provided to the Regional Director.
5. Any extensions or variations to the research proposal, additional research involving use of the data collected, or publication of the data beyond that normally associated with academic studies will require a further research approval submission.
6. At the conclusion of your study, a copy or summary of the research findings should be forwarded to me at the above address.

I wish you well with your research study. Should you have further enquiries on this matter, please contact Craig Irvin, Project Manager, School and Community Development Division, on 9637 2358.

Yours sincerely



Karen Moore
Acting Manager
School Community Links & Networks

encl.

Appendix 5

Explanatory statements and permission slips

Date:

Dear Principal

RE RESEARCH PROJECT: The effects of a Direct Instruction intervention in the regular mathematics class on students' mathematical self-efficacy and achievement.

My name is Rhonda Farkota and I am currently studying for a Doctor of Education at Monash University under the supervision of Dr Barbara Clarke, a senior lecturer in the Faculty of Education.

My research project is an important component of the course I am undertaking and in my research I aim to investigate:

- Student outcomes in mathematics at year 7; and
- The relationship of students' mathematics self-efficacy to their mathematics performance.

[PLEASE NOTE: This study has been approved by the *Department of Education, Employment and training Victoria* and *The Standing Committee on Ethics in Research on Humans Monash University* and no school or individual will be identified in any report from the study.]

Your school currently implements the Elementary Math Mastery program as part of its mathematics curriculum. I would like to examine the effects the program has on your students' self-efficacy and performance in mathematics.

The study involves a student questionnaire and mathematics assessment (ACER PAT maths). I am hoping that you will give approval for your year 7 classes to participate in the study.

The assessment and questionnaires would be taken during the usual mathematics lesson in the first week of school and during 4th term, or at a date that is convenient for your school. The questionnaire and assessment time will be approximately a total of 90 minutes.

The design of the project is such that any disruption to normal school program should be minimal, but please do not hesitate to contact me (03) 9277 5627 fax 9277 5500 or email farkota@acer.edu.au if you have any questions. I look forward to your support and would appreciate if you could return the attached Facsimile Response Sheet.

Yours sincerely

Rhonda Farkota

Appendix 5 (cont'd)

PARENT/GUARDIAN INFORMATION SHEET

Dear Parent/Guardian,

Rhonda Farkota is a Doctoral student at Monash University and is conducting a research regarding mathematics in year 7. The title of the research is: *The effects of a Direct Instruction intervention in the regular mathematics class on students' mathematical self-efficacy and achievement.*

The research project invites the participation of your son/daughter, who will be asked to complete an assessment of their achievement in mathematics and a short questionnaire on their feelings about mathematics. The object of the study is to gain a better understanding about what helps students learn mathematics in year 7.

Below I have answered some of the questions you may raise:

Does my child have to prepare for the assessment?

No, the assessment is designed to be a general test of achievement in mathematics at this level across different states and schools.

Will my child receive results of the assessment?

Your school will receive a summary of the research findings but not individual student results.

Will assessment be used by the school?

No, the content of the test is not directly aimed at any particular school or state curriculum.

Will my child's responses to the assessment be confidential?

Yes, the Researcher follows well established procedures to guarantee the anonymity of those who participate.

Can my child withdraw from the project?

Yes, your child can withdraw from the project at any time without prejudice.

If you are prepared for your child to take part, a Consent Form is attached for you to sign. Should you require additional information regarding this research or would like to be informed of the aggregate research finding, please contact Rhonda Farkota on (03) 9277 5627 fax 9277 5500 or email farkota@accr.edu.au.

Thank you for considering this request.

Signed _____ (Researcher's name) Date _____

You can complain about the study if you don't like something about it. To complain about the study, you need to phone 9905 2052. You can then ask to speak to the secretary of the Human Ethics Committee and tell him or her that the number of the project is 2001/577.

You could also write to the secretary. That person's address is:

The Secretary

The Standing Committee on Ethics in Research Involving Humans

PO Box No 3A

Monash University

Victoria 3800

Telephone (03) 9905 2052 Fax (03) 9905 1420

Email: SCERH@adm.monash.edu.au

Appendix 5 (cont'd)

PARENT/GUARDIAN CONSENT FORM

I(name) hereby consent to my child's involvement in the research project entitled:
The effects of a Direct Instruction intervention in the regular mathematics class on students' mathematical self-efficacy and achievement.

I have read and understood the information sheet on the above project and understand that my child is being asked to complete an assessment of their achievement in mathematics and a short questionnaire on their feelings about mathematics.

I understand that my child may not directly benefit by taking part in this research.

I understand that while information gained in the study may be published, my child will not be identified and all individual information will remain confidential.

I understand that I can withdraw my child from the study at any stage up until the end of the collection of data.

I understand that there will be no payment for my child taking part in the study.

I am aware that I should retain a copy of the information sheet and consent form for future reference.

I consent to my child being involved in this project.

Signed _____ Date _____

Relationship to child _____

Name of child _____

If you have any queries or would like to be informed of the aggregate research finding, please contact Rhonda Farkota telephone 9277 5627 fax 9277 5500

You can complain about the study if you don't like something about it. To complain about the study, you need to phone 9905 2052. You can then ask to speak to the secretary of the Human Ethics Committee and tell him or her that the number of the project is 2001/577. You could also write to the secretary. That person's address is:

The Secretary
The Standing Committee on Ethics in Research Involving Humans
PO Box No 3A
Monash University
Victoria 3800
Telephone (03) 9905 2052 Fax (03) 9905 1420
Email: SCERH@adm.monash.edu.au

Appendix 7

Teacher survey

Teacher survey to be completed by the mathematics teacher.

This survey is designed to give a brief picture of the EMM program in your school. Your responses will be kept completely confidential. Completed survey should be returned to your maths coordinator in the envelope provided.

1. The number of EMM lessons completed to date _____
2. The average number of minutes each lesson (do not include corrections or Debugging) _____ min
3. The average number of minutes each lesson for Corrections _____ min
4. The average number of minutes each lesson for Debugging _____ min
5. What have you found positive about the EMM program?

6. What are your concerns about the program?

7. Would you be prepared to trial the program another year? Please circle Yes No

8. Other comments

The space below may be used for any further comments you may wish to make.

9. Would you please list the names of students who on average are absent more than once a week.

Your name is not required. Please place in envelope provided.

Thank you for your responses

Rhonda Farkota
Research Fellow
ACER

Appendix 8

Stage 1 student questionnaire

Student Questionnaire Form 1 – 2002

Student's Name : _____

Office use only

| | | |
|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|

Maths Teacher's Name : _____

| | | |
|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|

General Instructions for completing this questionnaire

This questionnaire is designed to measure your attitudes and beliefs about mathematics. Please answer the questions honestly; there are no right or wrong answers. It is more important to find out how you really are, not how you would like to be or how you think you should be. Remember that other people will answer these questions differently to you, and that all the information you give is private.

Make sure you answer all the questions, but don't spend too much time thinking about your answers - the first answer that comes into your head is what is needed. Make sure that you read the instructions for each of the different sections as they may vary.

YOUR VIEWS ABOUT MATHEMATICS

Read each of the following statements and decide how you feel about it. Mark the corresponding letter on the blue answer sheet that best reflects how true you feel that the statement is for you.

| | Very True | Sort of true | Not very true | Not at all true |
|--|-----------|--------------|---------------|-----------------|
| 1. Maths is one of my favourite subjects | A | B | C | D |
| 2. When a maths problem comes up that I can't solve immediately, I keep trying until I work it out | A | B | C | D |
| 3. I usually understand the work we do in maths | A | B | C | D |
| 4. I like maths more than I like most other subjects | A | B | C | D |
| 5. It is not important to me to do well in maths | A | B | C | D |
| 6. Some people are good at maths and some just aren't | A | B | C | D |
| 7. I give up working on maths problems when I can't understand them | A | B | C | D |
| 8. Doing well in maths is important to me | A | B | C | D |

Appendix 8 (cont'd)

The next group of questions refers to maths and how you feel when doing maths at school. Mark the letter on your answer sheet that you feel best reflects how true you feel that the statement is for you.

| | Very True | Sort of true | Not very true | Not at all true |
|---|-----------|--------------|---------------|-----------------|
| 9. I can do well in maths if I want to | A | B | C | D |
| 10. I don't know what it takes to do well in maths | A | B | C | D |
| 11. I am bored in maths | A | B | C | D |
| 12. Trying hard is the best way for me to do well in maths | A | B | C | D |
| 13. If I'm not smart, I won't do well at maths | A | B | C | D |
| 14. I don't know how to keep myself from doing badly in maths | A | B | C | D |
| 15. If I don't do well in maths, it's because I didn't try hard enough | A | B | C | D |
| 16. Maths doesn't scare me at all | A | B | C | D |
| 17. I have to be smart to do well in maths (If I want to do well in maths, being smart is what counts the most) | A | B | C | D |
| 18. I can't do well in maths | A | B | C | D |
| 19. When I'm in maths classes I usually feel uncomfortable and nervous | A | B | C | D |
| 20. I can work really hard in maths | A | B | C | D |
| 21. I dread having to do maths | A | B | C | D |
| 22. When I'm in maths classes I usually feel happy | A | B | C | D |
| 23. My mind goes blank and I am unable to think clearly when working mathematics | A | B | C | D |
| 24. I try and learn as much as I can about the maths we do | A | B | C | D |
| 25. When I'm in maths classes, I try very hard | A | B | C | D |
| 26. I pay attention in maths | A | B | C | D |
| 27. I enjoy doing maths | A | B | C | D |
| 28. When I'm in maths, I usually just pretend that I'm working | A | B | C | D |
| 29. I don't try very hard in maths | A | B | C | D |
| 30. I find maths interesting | A | B | C | D |

Appendix 8 (cont'd)

Think of the last time something bad happened in maths (like not doing well on a test, or not understanding what the teacher was explaining). Here are some things other students have said that they think and do after things like this happen. How true are these for YOU?

| When something bad happens to me in maths | Very true | Sort of true | Not very true | Not at all true |
|---|-----------|--------------|---------------|-----------------|
| 31. I tell myself it didn't matter anyway | A | B | C | D |
| 32. I feel really stupid | A | B | C | D |
| 33. I worry that the other students will think I'm dumb | A | B | C | D |
| 34. I tell myself I'll do better next time | A | B | C | D |
| 35. I say I didn't care about it anyway | A | B | C | D |
| 36. I try to see where I went wrong | A | B | C | D |

Suppose you are asked the following maths questions in a test. Please indicate how confident you are that you will give the correct answer to each question without using a calculator. PLEASE DO NOT ATTEMPT TO SOLVE THE QUESTIONS.

| | Very confident | Sort of confident | Not very confident | Not at all confident |
|--|----------------|-------------------|--------------------|----------------------|
| 37. Calculate $86,204 + 7,699$ | A | B | C | D |
| 38. Calculate $986,204 - 17,699$ | A | B | C | D |
| 39. Multiply 853 by 45 | A | B | C | D |
| 40. Divide 4518 by 9 | A | B | C | D |
| 41. Find the missing number in the pattern 18 27 ? 45 54 | A | B | C | D |
| 42. Express 0.65 as a percentage | A | B | C | D |
| 43. Draw an angle of 135° | A | B | C | D |
| 44. Express 0.001 as a fraction | A | B | C | D |
| 45. Evaluate $3 + 2m$ if $m = 6$ | A | B | C | D |
| 46. Complete $1.86 \text{ m} = ? \text{ cm}$ | A | B | C | D |
| 47. Express $\frac{12}{5}$ as a mixed number | A | B | C | D |
| 48. Find the product of the first two multiples of 4 | A | B | C | D |
| 49. Divide 7.5 by 6 | A | B | C | D |
| 50. How many vertices has a triangular prism? | A | B | C | D |
| 51. What is the surface area of a cube 2 cm in height? | A | B | C | D |
| 52. Express 2:45 pm in 24-hour time | A | B | C | D |
| 53. How many lines of symmetry does a rectangle have? | A | B | C | D |

Appendix 8 (cont'd)

For each of the following questions, choose one of the letters to indicate how you feel.

A means 'excellent', C is 'average', and E means 'weak'.

| | Excellent | | Average | | Weak | |
|---|-----------|---|---------|---|------|--|
| | A | B | C | D | E | |
| 54. How good are you at maths? | A | B | C | D | E | |
| 55. How good would you like to be at maths? | A | B | C | D | E | |
| 56. Where would your teacher put you on this scale? | A | B | C | D | E | |
| 57. Where would your parents put you on this scale? | A | B | C | D | E | |
| 58. How good do you think your parents would like you to be at maths? | A | B | C | D | E | |
| 59. Where would your classmates put you on this scale? | A | B | C | D | E | |

60. What language does your family mostly speak at home?

| | |
|---------|-------|
| English | Other |
| A | B |

61. Do you like maths? Explain why or why not.

62. What do you think will be the best things about secondary school?

63. Is there anything about secondary school that you aren't looking forward to?

64. What will you miss about primary school?

65. Do you think that maths classes will be different in secondary school than they are in primary school?

If so, how do you think things will be different?

☺ Thank you for completing this questionnaire! ☺

Appendix 9

Stage 2 student questionnaire

Student Questionnaire Form 2B – 2002

Student's Name : _____

Maths Teacher's Name : _____

General Instructions for completing this questionnaire

This questionnaire is designed to measure your attitudes and beliefs about mathematics. Please answer the questions honestly; there are no right or wrong answers. It is more important to find out how you really are, not how you would like to be or how you think you should be. Remember that other people will answer these questions differently to you, and that all the information you give is private.

Make sure you answer all the questions, but don't spend too much time thinking about your answers - the first answer that comes into your head is what is needed. Make sure that you read the instructions for each of the different sections as they may vary.

YOUR VIEWS ABOUT MATHEMATICS

Read each of the following statements and decide how you feel about it. Mark the corresponding letter on the blue answer sheet that best reflects how true you feel that the statement is for you.

| | Very true | Sort of true | Not very true | Not at all true |
|--|-----------|--------------|---------------|-----------------|
| 1. Maths is one of my favourite subjects | A | B | C | D |
| 2. When a maths problem comes up that I can't solve immediately, I keep trying until I work it out | A | B | C | D |
| 3. I usually understand the work we do in maths | A | B | C | D |
| 4. I like maths more than I like most other subjects | A | B | C | D |
| 5. It is not important to me to do well in maths | A | B | C | D |
| 6. Some people are good at maths and some just aren't | A | B | C | D |
| 7. I give up working on maths problems when I can't understand them | A | B | C | D |
| 8. Doing well in maths is important to me | A | B | C | D |

Appendix 9 (cont'd)

The next group of questions refers to maths and how you feel when doing maths at school. Mark the letter on your answer sheet that you feel best reflects how true you feel that the statement is for you.

| | Very True | Sort of true | Not very true | Not at all true |
|---|-----------|--------------|---------------|-----------------|
| 9. I can do well in maths if I want to | A | B | C | D |
| 10. I don't know what it takes to do well in maths | A | B | C | D |
| 11. I am bored in maths | A | B | C | D |
| 12. Trying hard is the best way for me to do well in maths | A | B | C | D |
| 13. If I'm not smart, I won't do well at maths | A | B | C | D |
| 14. I don't know how to keep myself from doing badly in maths | A | B | C | D |
| 15. If I don't do well in maths, it's because I didn't try hard enough | A | B | C | D |
| 16. Maths doesn't scare me at all | A | B | C | D |
| 17. I have to be smart to do well in maths (If I want to do well in maths, being smart is what counts the most) | A | B | C | D |
| 18. I can't do well in maths | A | B | C | D |
| 19. When I'm in maths classes I usually feel uncomfortable and nervous | A | B | C | D |
| 20. I can work really hard in maths | A | B | C | D |
| 21. I dread having to do maths | A | B | C | D |
| 22. When I'm in maths classes I usually feel happy | A | B | C | D |
| 23. My mind goes blank and I am unable to think clearly when working mathematics | A | B | C | D |
| 24. I try and learn as much as I can about the maths we do | A | B | C | D |
| 25. When I'm in maths classes, I try very hard | A | B | C | D |
| 26. I pay attention in maths | A | B | C | D |
| 27. I enjoy doing maths | A | B | C | D |
| 28. When I'm in maths, I usually just pretend that I'm working | A | B | C | D |
| 29. I don't try very hard in maths | A | B | C | D |
| 30. I find maths interesting | A | B | C | D |

Appendix 9 (cont'd)

Think of the last time something bad happened in maths (like not doing well on a test, or not understanding what the teacher was explaining). Here are some things other students have said that they think and do after things like this happen. How true are these for YOU?

| When something bad happens to me in maths | Very true | Sort of true | Not very true | Not at all true |
|---|-----------|--------------|---------------|-----------------|
| 31. I tell myself it didn't matter anyway | A | B | C | D |
| 32. I feel really stupid | A | B | C | D |
| 33. I worry that the other students will think I'm dumb | A | B | C | D |
| 34. I tell myself I'll do better next time | A | B | C | D |
| 35. I say I didn't care about it anyway | A | B | C | D |
| 36. I try to see where I went wrong | A | B | C | D |

Suppose you are asked the following maths questions in a test. Please indicate how confident you are that you will give the correct answer to each question without using a calculator. PLEASE DO NOT ATTEMPT TO SOLVE THE QUESTIONS.

| | Very confident | Sort of confident | Not very confident | Not at all confident |
|--|----------------|-------------------|--------------------|----------------------|
| 37. Calculate $86,204 + 7,699$ | A | B | C | D |
| 38. Calculate $986,204 - 17,699$ | A | B | C | D |
| 39. Multiply 853 by 45 | A | B | C | D |
| 40. Divide 4518 by 9 | A | B | C | D |
| 41. Find the missing number in the pattern 18 27 ? 45 54 | A | B | C | D |
| 42. Express 0.65 as a percentage | A | B | C | D |
| 43. Draw an angle of 135° | A | B | C | D |
| 44. Express 0.001 as a fraction | A | B | C | D |
| 45. Evaluate $3 + 2m$ if $m = 6$ | A | B | C | D |
| 46. Complete $1.86 \text{ m} = ? \text{ cm}$ | A | B | C | D |
| 47. Express $\frac{12}{5}$ as a mixed number | A | B | C | D |
| 48. Find the product of the first two multiples of 4 | A | B | C | D |
| 49. Divide 7.5 by 6 | A | B | C | D |
| 50. How many vertices has a triangular prism? | A | B | C | D |
| 51. What is the surface area of a cube 2 cm in height? | A | B | C | D |
| 52. Express 2:45 pm in 24-hour time | A | B | C | D |
| 53. How many lines of symmetry does a rectangle have? | A | B | C | D |

Appendix 9 (cont'd)

For each of the following questions choose one of the letters to indicate how you feel. Mark the corresponding letter on the blue answer sheet and give written response here.

54. Do you like maths?

| | | | | |
|---------------|----------|----------|---------------|----------|
| Yes very much | Yes | Sort of | Not very much | No |
| A | B | C | D | E |

55. Which part of the maths lesson do you like the most?

| | | | | |
|--------------------------------|--|------------|------------------------|----------|
| EMM Elementary Math Mastery | CLASS TEXT BOOK (eg Heinemann, Jacaranda, Nelson) | WORKSHEETS | CORRECTING HOMEWORK | OTHER |
| A | B | C | D | E |

If you chose OTHER (E) please describe it. _____

56. How good are you at maths?

| | | | | |
|-----------|---------------|----------|---------------|----------|
| Excellent | Above average | Average | Below average | Weak |
| A | B | C | D | E |

57. Where would your classmates put you on this scale?

| | | | | |
|-----------|---------------|----------|---------------|----------|
| Excellent | Above average | Average | Below average | Weak |
| A | B | C | D | E |

58. From which part of the maths lesson do you think you have learned the most?

| | | | | |
|--------------------------------|--|------------|------------------------|----------|
| EMM Elementary Math Mastery | CLASS TEXT BOOK (eg Heinemann, Jacaranda, Nelson) | WORKSHEETS | CORRECTING HOMEWORK | OTHER |
| A | B | C | D | E |

Why do you think this is? _____

If you chose OTHER (E) please describe it. _____

59. From which part of the maths lesson do you think you have learned the least?

| | | | | |
|--------------------------------|--|------------|------------------------|----------|
| EMM Elementary Math Mastery | CLASS TEXT BOOK (eg Heinemann, Jacaranda, Nelson) | WORKSHEETS | CORRECTING HOMEWORK | OTHER |
| A | B | C | D | E |

Why do you think this is? _____

If you chose OTHER (E) please describe it. _____

60. Compared to primary school how do you think you are doing in maths?

| | | | | |
|--------------|-----------------|----------------|----------------|-------------|
| A lot better | A little better | About the same | A little worse | A lot worse |
| A | B | C | D | E |

Why do you think this is? _____

☺ Thank you for completing this questionnaire! ☺

Appendix 10 Intervention sample lesson

THE EMM STRANDS

- | | |
|------------------------------------|---------------------------------------|
| 1 Addition | 11 Space |
| 2 Subtraction | 12 Geometry |
| 3 Multiplication | 13 Average, percentage, ratio, chance |
| 4 Division | 14 Math language |
| 5 Number patterns | 15 Money |
| 6 Equations and inverse operations | 16 Time |
| 7 Whole number properties | 17 Algebra |
| 8 Fractions | 18 Visual perception |
| 9 Decimals | 19 Data analysis |
| 10 Measurement | 20 Problem solving |

USING THE SCRIPT

Lesson 1

Display Blackline Master 1.

◀ Prepare and display Blackline Master indicated.

$$\begin{array}{r} 564 \\ + 212 \\ \hline 776 \end{array}$$

◀ What you write on the board appears in a display box.

With column problems like this, digits must be lined up.

◀ What you say appears in this type.

Start with ones column, **FOUR** add **TWO** equals **SIX**.

◀ What you say and simultaneously point to on the board appears in CAPITAL LETTERS.

Add the tens column, **SIX** add **ONE** equals **SEVEN**.

Finally add the **HUNDREDS COLUMN**.

The sum is the total after you add.

Question 1. Find the sum of 564 and 212.

◀ What you repeat is underlined.

Write your answer in today's column next to question 1.

Amend board display from previous question.

◀ What you do appears in *italics*.

$$\begin{array}{r} 564 \\ - 212 \\ \hline 352 \end{array}$$

Lesson 18

Display Blackline Master 1.

6 2 3

Question 1. Find the sum of SIX HUNDRED AND TWENTY-THREE and 423.

Use board display from previous question.

6 2 3

Question 2. SIX HUNDRED AND TWENTY-THREE minus 423.

Question 3. Multiply 15 by 4.

$$\begin{array}{r} 40 \\ 2 \overline{) 802} \end{array}$$

EIGHT hundreds divided by TWO. That's FOUR. ZERO divided by TWO. That's ZERO. Now work the problem for the last digit: TWO divided by TWO.

Question 4. 802 divided by 2.

If 5 goes into a number exactly, the number is divisible by 5.

Question 5. True or false: 542 is divisible by 5.

Question 6. I thought of a number, multiplied it by 2 and got 18. Find the number.

Question 7. Multiply 125 by 10 and circle the digit in the tens place.

$$\frac{2}{2} = 1$$

A fraction equals 1 whole when the TOP NUMBER and BOTTOM NUMBER of the fraction are the same.

Question 8. Write the fraction that equals 1 whole when the denominator, that's the bottom number, is 4.

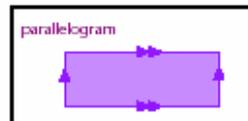
4 and 1-tenth

Question 9. Write the decimal for this STATEMENT.

3 m 300 cm 300 mm

THREE METRES, THREE HUNDRED CENTIMETRES, THREE HUNDRED MILLIMETRES.

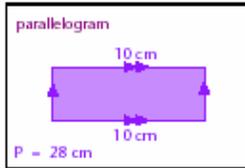
Question 10. What is the shortest measurement?



A 4-sided shape with opposite sides parallel is called a PARALLELOGRAM. Sides marked with an EQUAL NUMBER OF ARROWS are parallel.

Question 11. True or false: a rectangle is a parallelogram.

Add to board display from previous question.



The perimeter of a parallelogram is **TWENTY-EIGHT CENTIMETRES**.
To find the lengths of the missing sides, subtract the lengths of the 2 known sides from the perimeter and divide by 2.

Question 12. Find the lengths of the missing sides of this parallelogram.

Question 13. If 4 children have an average height of 121 centimetres, what is their total height?

$$2 \times 5 = 10$$

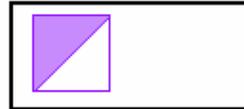
The product is the total after multiplication. **TEN** is the product of **TWO** and **FIVE**.

Question 14. Find the product of 4 and 5.

Question 15. Subtract \$2.18 from \$4.28.

Question 16. Write quarter-past 8 in the evening.

Question 17. Write the number sentence for this fact: the product of 75 and 4 is 300.



Question 18. What fraction of this square is shaded?

Refer to Blackline Master 1.

Look at the **TABLE**.

Question 19. Which grade has the least number of girls?

Question 20. A cork and bottle costs \$3. The bottle costs \$2 more than the cork. Find the cost of the cork.

Answer key

- 18.1 1046
- 18.2 200
- 18.3 60
- 18.4 401
- 18.5 False
- 18.6 9
- 18.7 12⁰⁰
- 18.8 $\frac{1}{4}$
- 18.9 4.1
- 18.10 300 mm
- 18.11 True
- 18.12 4 cm, 4 cm
- 18.13 484 cm
- 18.14 20
- 18.15 \$2.10
- 18.16 8:15 p.m.
- 18.17 $75 \times 4 = 300$
- 18.18 $\frac{1}{2}$
- 18.19 Grade two
- 18.20 50c

References

- Abt Associates. (1977). *Volume IV-B: Effects of Follow Through models* (Office of Education Series Vol. II-B AAI-76-196B under USOE Contract No. 300-75-0134). Cambridge MA: Author.
- Adams, G. L., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.
- Adams, R. J., & Khoo, S. T. (1993). *Quest: The interactive test analysis system*. Hawthorn: Australian Council for Educational Research.
- Ainley, J., Graetz, B., Long, M., & Batten, M. (1995). *Socioeconomic status and school education*. Canberra: Aust Govt Pub Service.
- Airasian, P., & Walsh, M. (1997). Constructivist cautions. *Phi Delta Kappan*, 78(6), 444-449.
- Alexander, R. J. (1995). *Versions of primary education*. London: Routledge.
- Andrich, D. (1997). Rating scale analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (2 ed., pp. 874-880). Oxford: Pergamon Press.
- Andrich, D., & Masters, G. N. (1988). Rating scale analysis. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp. 297-303). Oxford: Pergamon Press.
- AREA. (1999). *Teaching numeracy and literacy: Interventions and strategies for "at risk" students*. Australian Resource Educators' Association Inc. [AREA].
- Australian Bureau of Statistics. (1997). *1996 census of population and housing, basic community profile, Australia* (Catalogue no. 2020.0).
- Australian Council for Educational Research. (1997). *Progressive achievement tests in mathematics (PATMaths)* (Revised ed.). Camberwell, Vic: Australian Council for Educational Research (ACER).
- Australian Council for Educational Research. (1998). *Progressive achievement tests in mathematics (PATMaths) norming manual* (Revised ed.). Camberwell, Vic: Australian Council for Educational Research.
- Australian Education Council and Curriculum Corporation. (1991). *A national statement on mathematics for schools Australian schools*. Carlton, Victoria: Curriculum Corporation.

- Ayres, P. (2000). *Numeracy in Australia: Are standards falling?* (Vol. 7). North Sydney: Australian Education Network.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*, 117-148.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bandura, A. (2002). *The distinction between self-efficacy and confidence*, [internet]. Emory University. Available: <http://www.emory.edu/EDUCATION/mfp/banconfidence.html> [2002, 2/1/02].
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality & Social Psychology*, *41*, 586-598.
- Barnes, D. (1985). Why not direct instruction? *The Australian Educational & Developmental Psychologist*, *1*, 59-62.
- Battista, M. T. (1999). The mathematical miseducation of America's youth: Ignoring research and scientific study in education. *Phi Delta Kappan*, *80*(6), 425-433.
- Bay, J., Beem, J., Reys, R., Papick, I., & Barnes, D. (1999). Student reactions to standards-based mathematics curricula: The interplay between curriculum, teachers and students. *School Science and Mathematics*, *99*(4), 182-187.
- Bay, M., Staver, J. V., Bryan, T., & Hale, J. B. (1992). Science instruction for the mildly handicapped: Direct instruction versus discovery teaching. *Journal of Research in Science Teaching*, *29*, 555-570.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*, 506-521.
- Becker, W. C. (1977). Teaching reading and language to the disadvantaged — What we have learned from field research. *Harvard Educational Review*, *47*(4), 516-545.
- Becker, W. C. (1978). The national evaluation of Follow Through: Behaviour-based programmes come out on top. *Education and Urban Society*, *10*(4), 431-458.

- Becker, W. C., & Engelmann, S. (1996). Sponsor findings from project Follow Through. *Effective School Practices*(Winter).
- Becker, W. C. E. (1988). Direct instruction: A general case for teaching a general case. [Special issue]. *Education and Treatment of Children, 11*, 301-396.
- Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1992). Theory into practice: How do we link? In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and technology of instruction*. (pp. 17-34). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Bender, W. (1993). *Learning disabilities: Best practices for professionals*. Boston: Andover Medical Publishers.
- Benson, J. (1989). Structural components of statistical test anxiety in adults: A exploratory study. *Journal of Experimental Education, 57*, 247-261.
- Bereiter, C., & Kurland, M. (1981-1982). A constructive look at Follow Through results. *Interchange on Educational Policy, 12*(1), 1-22.
- Binder, C., & Watkins, C. L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly, 3*(4), 74-96.
- Bishop, A. J. (1999). Mathematics teaching and values education — An intersection in need of research. *International Reviews on Mathematical Education, 31*(1), 1-4.
- Bishop, W. (1999). The California mathematics standards. *Phi Delta Kappan, 80*(6), 439-440.
- Bjorklund, D. F. (1995). *Children's thinking: Developmental function and individual differences*. (2 ed.). Pacific Grove, CA: Brooks-Cole.
- Bloom, B. S. (1986). The hands and feet of genius. *Educational Leadership, 43*, 70-77.
- Boaler, J. (1997). *Experiencing school mathematics*. Buckingham, UK: Open University Press.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. New Jersey: Lawrence Erlbaum Associates, Inc.
- Bong, M., & Clark, R. (1999). Comparisons between self-concept and self-efficacy in academic motivation research. *Education Psychologist, 34*, 139-154.

- Boomer, G. (1988). *Standards & literacy. Two hundred years on the road to literacy: Where to from here? Directions: Literacy* (Supplement to *Education Victoria*). Melbourne, Australia: State Board of Education.
- Brammer, L. M. (1992). *Coping with life transitions*, [on-line]. ERIC Digest. Available: <http://136.242.172.58/db/riecij/ed350527.htm> [2001, July].
- Bransford, J. D., Brown, A. L., Cocking, R. R., Donovan, M. S., & Pellegrino, J. W. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (expanded ed.). Washington, DC: National Academy Press.
- Byrne, B. M., & Shavelson, R. J. (1987). Adolescent self-concept. Testing the assumption of equivalent structure across gender. *American Educational Research Journal*, 24, 365-385.
- Carnine, D. (1991). Curricular interventions for teaching higher order thinking to all students: Introduction to the special series. *Journal of Learning Disabilities*, 24(5), 261-269.
- Carnine, D. (1997). Instructional design in mathematics for students with learning disabilities. *Journal of Learning Disabilities*, 30(2), 130-141.
- Carnine, D. (2000). Why education experts resist effective practices (and what it would take to make education more like medicine). *Effective School Practices*, 18(3), 6-14.
- Carpenter, T. P., Fennema, E., Fuson, K., Hibert, J., Human, P., Murray, H., Olivier, A., & Wearne, D. (1999). Teaching and learning with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499-531.
- Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 18-32). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Casazza, M. E. (1993). Using a model of direct instruction to teach summary writing in a college reading class. *Journal of Reading*, 37(3), 202-209.
- Cassidy, S. (2000, 4 July 2000). Maths to meet the workers' needs. *Times*, pp. 7.

- Chambliss, C., & Murray, E. J. (1979). Cognitive procedures for smoking reduction: Symptom attribution versus efficacy attribution. *Cognitive Therapy & Research*, 3(1), 91-95.
- Chapman, J. W. (1988). Cognitive motivational characteristics and academic achievement of learning. *Journal of Educational Psychology*, 80, 357-365.
- Clarke, D. (1989). *Mathematical behaviour and the transition from primary to secondary school*. Unpublished PhD thesis, Monash University, Melbourne.
- Cockcroft, W. H. (1982). *Mathematics counts*. London: HMSO.
- Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks: Sage Publications.
- Cole, P., & Chan, L. (1990). *Methods and strategies for special education*. Sydney: Prentice Hall.
- Collins, J. L. (1982). *Self-efficacy and achievement behaviour* (Paper). New York: American Educational Research Association.
- Combs, A. 1912-1999, [online]. Available: <http://www.coe.ufl.edu/webtech/greatideas/pages/peoplepage/combs.htm> [2003, February].
- Commission of Inquiry into Poverty. (1976). *Poverty and education in Australia* (Fifth main report). Canberra: AGPS (R. T. Fitzgerald, Commissioner).
- Confrey, J. (1990). What constructivism implies for teaching. In R. B. Davis & C. A. Maher & N. N (Eds.), *Constructivist views on the teaching and learning of mathematics, JRME Monograph 4* (pp. 109).
- Cooper, S. E., & Robinson, D. A. G. (1991). The relationship of mathematics self-efficacy beliefs to mathematics anxiety and performance. *Measurement and Evaluation in Counselling and Development*, 24, 4-11.
- Cossey, R. (1999). Are California's math standards up to the challenge? *Phi Delta Kappan*(February), 441-443.
- Cotton, K., & Savard, W. (1982). *Direct instruction: Research on school effectiveness project*. Portland, OR: Northwest Regional Laboratory.
- Cumming, J. (2000). *Computational numeracy*. Paper presented at the ACER Research Conference 2000: Improving numeracy learning: What does the research tell us? Brisbane.
- Curriculum Corporation. (1994). *Mathematics: A curriculum profile for Australian schools*. Carlton, Vic: Curriculum Corporation.

- Darch, C., Carnine, D., & Gersten, R. (1984). Explicit instruction in mathematics problem solving. *Journal of Educational Research*, 77, 351-359.
- de Groot, C. (2000). *Three female voices: The transition to high school mathematics from a reform middle school mathematics program*. Unpublished PhD thesis, New York University, New York.
- Department of Education and Science. (1992). *Curriculum organisation and classroom practice in primary schools: A discussion paper*. London: Author.
- Department of Education Employment and Training. (2001). *Standards and accountability: Quality assurance in Victorian schools: Benchmarks 2000: VCE Victorian certificate of education*. Melbourne, Vic: Department of Education, Employment and Training.
- Dick, W. (1992). An instructional designer's view of constructivism. In T. M. Duffy & D. Jonassen (Eds.), *Constructivism and technology of instruction*. (pp. 91-98). Hillsdale NJ: Erlbaum.
- Dietz, F. (1998). Teaching the technology workforce. *Mechanical Engineering*, 120(5), 40.
- Dougherty, K. M., & Johnston, J. M. (1996). Overlearning, fluency and automaticity. *The Behavior Analyst*, 2(Fall), 289-292.
- Doyle, W., & Ponder, G. A. (1977). The practicality ethic in teacher decision-making. *Interchange*, 8(3), 1-12.
- Drecktrah, M. E., & Chiang, B. (1997). Instructional strategies used by general educators and teachers of students with learning disabilities: A survey. *Remedial and Special Education*, 18(3), 174-181.
- DuCharme, C., Earl, J., & Poplin, M. S. (1989). The author model: The constructivist view of the writing process. *Learning Disability Quarterly*, 12, 237-242.
- Edgar, D. (1981). Social class differences and the structure of education. In P. Hiller (Ed.), *Class and inequality in Australia: Sociological perspectives and research*. Sydney: Harcourt Brace Jovanovich.
- Ellerton, N. F., & Clements, M. A. (1988). The transition from year 6 to year 7: A problem or an opportunity? In D. Firth (Ed.), *Mathematics counts: Who cares?* (pp. 134-140). Melbourne: Mathematical Association of Victoria.
- Engelmann, S. (1980). Towards the design of faultless instruction: The theoretical basis of concept analysis. *Educational Technology*(February), 23-36.

- Engelmann, S., Becker, W., Carnine, D., & Gersten, R. (1988). The direct instruction Follow Through model: Design and outcomes. *Education and Treatment of Children, 11*(4), 303-317.
- Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications*. New York: Irvington.
- Engelmann, S., Carnine, D., & Steely, G. (1991). Making connections in mathematics. *Journal of Learning Disabilities, 24*(5).
- Englert, C. S. (1984). Effective direct instruction practices in special education settings. *Remedial and Special Education, 5*(2), 38-47.
- Erikson, E. H. (1959). *Identity and the life cycle*. New York: Norton.
- Erikson, E. H. (1963). *Childhood and society*. New York: Norton.
- Farkota, R. (2000). *Elementary math mastery*. Roseville: McGrawHill.
- Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitude towards learning of mathematics by females and males. *Psychological Documents, 6*,1(31).
- Fetherston, T. (1997). The derivation of a learning approach based on a personal construct psychology. *International Journal of Science Education, 19*(7), 801-819.
- Fielding, G., Kameenui, E., & Gersten, R. (1983). A comparison of inquiry and direct instruction approaches toward teaching legal concepts and applications to secondary students. *Journal of Educational Research, 76*, 287-293.
- Fields, B. A. (1986). Direct instruction and the practicality ethic: A case study and position in curriculum innovation. *Australian Journal of Remedial Education, 18*(1), 17-22.
- Finn, C. E. J. (2000, April). *Why education experts resist effective practices (and what it would take to make education more like medicine)*, [internet]. Thomas B. Fordham Foundation. Available: www.edexcellence.net [2002, 11 December].
- Fosnot, C. T. (1996). Constructivism: A psychological theory of learning. In C. T. Fosnot (Ed.), *Constructivism: Theory, perspectives and practice* (pp. 8-32). New York: Teachers College Press.
- Fulk, B. M., Brigham, F. J., & Lohman, D. A. (1998). Motivation and self regulation. *Remedial and Special Education, 19*(5), 300-309.

- Fullan, M. (1982). *The meaning of educational change*: Columbia University, N.Y.: Teachers College Press.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47(2), 335-397.
- Fullarton, S. (1993). *Confidence in mathematics: The effects of gender: A study of 450 students at 13 schools*. Geelong, Vic: Deakin University. National Centre for Research and Development in Mathematics Education.
- Fullarton, S. (1998). *Engagement or alienation: The transition to secondary school mathematics*. Unpublished PhD thesis, Monash University, Melbourne.
- Gagne, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning*. (2 ed.). New York: Harper Collins.
- Galton, M., Hargreaves, L., Comber, C., Wall, D., & Pell, A. (1999). *Inside the primary classroom: 20 years on*. London: Routledge.
- Gersten, R. (1985). Direct instruction with special education students: A review of evaluation research. *Journal of Special Education*, 19(1), 41-58.
- Gersten, R., & Carnine, D. (1984). Direct instruction mathematics: A longitudinal evaluation of low-income elementary school students. *The Elementary School Journal*, 84(4), 395-407.
- Gersten, R., & Carnine, D. (1986). Direct instruction in reading comprehension. *Educational Leadership*, 4, 70-78.
- Gersten, R., Carnine, D., & Woodward, J. (1987). Direct instruction research: The third decade. *Remedial and Special Education*, 8(6), 48-56.
- Gersten, R., & Guskey, T. (1985). Transforming teacher reluctance into a commitment to innovation. *Direct Instruction News*, Fall, 11-12.
- Gersten, R., & Maggs, A. (1982). Teaching the general case to moderately retarded children: Evaluation of a five-year project. *Analysis and Intervention in Developmental Disorders*, 2, 329-343.
- Gersten, R., Woodard, J., & Darch, C. (1986). Direct instruction: A research-based approach to curriculum design and teaching. *Exceptional Children*, 53(1), 17-31.
- Goldsmith, L. T., & Shifter, D. (1997). Understanding teachers in transition: Characteristics of a model for the development of mathematics teaching. In E. Fennema & B. S. Nelson (Eds.), *Mathematics teachers in transition* (pp. 19-54). Mahwah, New Jersey: Lawrence Erlbaum Associates.

- Gordon, E. E. (1997). The new knowledge worker. *Adult Learning*, 8(4), 14-18.
- Graham, S., & Harris, K. R. (1994). Implications of constructivism for teaching writing to students with special needs. *The Journal of Special Education*, 28(3), 275-289.
- Grant, S. G. (1998). *Reforming reading, writing and mathematics*. Mahwah, NJ: Erlbaum.
- Grossen, B. (1993). Child-directed teaching methods: A discriminatory practice of western education. *Effective School Practices*, Spring, 9-20.
- Grossen, B. (1995). The story behind Follow Through. *Effective School Practices*, 15(1).
- Grossen, B., & Ewing, S. (1994). *Raising mathematics problem-solving performance* (Technical 102): National Centre to Improve the Tools of Education, University of Oregon.
- Gruder, C. L. (1977). Choice of comparison persons in evaluating oneself. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes*. Washington, D C: Hemisphere publishing corporation.
- Hallam, S., & Toutounji, I. (1996). *What do we know about the grouping of pupils by ability?:* University of London. Institute of Education.
- Hanchon Graham, L. (2000). *Self-efficacy, motivation constructs, and mathematics performance of middle school students: A three-year longitudinal study*. Unpublished PhD thesis, Emory University, Emory.
- Hands, B. P. (1993). *Measurement of teacher attitudes to direct instruction*. Unpublished Master of Education thesis, Edith Cowan.
- Hara, K. (1995). Quantitative and qualitative research approaches in education. *Education*, 115(3), 351-354.
- Harris, K. R., & Graham, S. (1996). Memo to constructivists: Skills count, too. *Educational Leadership*, 53(5), 26-29.
- Hempenstall, K. (1996). The gulf between educational research and policy: The example of direct instruction and whole language. *Behaviour Change*, 13, 33-46.
- Hempenstall, K. (1997). *The effects on the phonological processing skills of disabled readers of participating in direct instruction reading programs*. Unpublished PhD thesis, RMIT, Melbourne.
- Hunter, J. E. (1986). Cognitive ability and aptitudes. *Journal of Vocational Behaviour*, 29, 340-362.

- Inagaki, K., Hatano, G., & Morita, E. (1998). Construction of mathematical knowledge through whole-class discussion. *Learning and Instruction*, 8(6), 503-526.
- Institute for Research and Reform in Education. (1998). *Research assessment package for schools (RAPS): Student supports and opportunities in school: Engagement, beliefs about self, and experiences of interpersonal support: Manual for elementary and middle school assessments*. NJ: Institute for Research and Reform in Education.
- International Association for the Evaluation of Educational Achievement (IEA). (1997). *The third international mathematics and science study (TIMSS)*. Massachusetts: Boston College.
- Johnston, D., Proctor, W., & Corey, S. (1994). Not a way out: A way in. *Educational Leadership*, December/January 1995, 46-49.
- Jones, E. D., Wilson, R., & Bhojwani, S. (1997). Mathematics instruction for secondary students with learning disabilities. *Journal of Learning Disabilities*, 30(2 March/April), 151-163.
- Kavale, K. (1990). Variances & verities in learning disability interventions. In T. Scruggs & B. Wong (Eds.), *Intervention research in learning disabilities* (pp. 3-33). New York: Springer-Verlag.
- Keeves, J. P. (1997). Longitudinal research methods. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (2 ed., pp. 138-149). Oxford: Pergamon Press.
- Kelly, B., Carnine, D., Gersten, R., & Grossen, B. (1986). Effectiveness of videodisc instruction in teaching fractions to learning-disabled and remedial high school students. *Journal of Special Education Technology*, 8, 5-9.
- Kinder, D., & Carnine, D. (1991). Direct instruction: What it is and what it is becoming. *Journal of Behavioral Education*, 1(2), 193-213.
- Kuder, S. J. (1991). Language abilities and progress in a direct instruction reading program for students with learning disabilities. *Journal of Learning Disabilities*, 24, 124-127.
- Lamb, S. (1997). *School achievement and initial education and labour market outcomes* (Research 4). Melbourne: ACER.
- Lindsley, O. (1992). Why aren't effective teaching tools widely adopted? *Journal of Applied Behavior Analysis*, 25(1&2).
- Lochhead, J. (1991). Making math mean. In E. von Glaserfeld (Ed.), *Radical constructivism in mathematics education* (pp. 75-87). Dordrecht: Kluwer.

- Lockery, L., & Maggs, A. (1982). Direct instruction research in Australia: A ten year analysis. *Educational Psychology, 2*, 263-288.
- Lokan, J., Greenwood, L., & Creswell, J. (2001). *15-up and counting, reading, writing, reasoning: How literate are Australian students?: The PISA 2000 survey of students' reading, mathematical and scientific literacy skills*. Melbourne, Vic: ACER Press.
- Ludlow, L., & O'Leary, M. (2000). What to do about missing data? *Rasch Measurement Transactions, 14*(2), 751.
- Ma, L. (2000). *Knowing and teaching elementary mathematics*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Marks, G. N., & Fleming, N. (1998). *Factors influencing youth unemployment in Australia: 1980-1994* (Research 7). Melbourne: ACER.
- Marsh, H. W. (1989). Sex differences in the development of verbal and mathematics constructs: The high school and beyond study. *American Educational Research Journal, 26*, 191-225.
- Marsh, H. W. (1990). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology, 82*(1), 107-116.
- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*, 107-125.
- Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology, 16*, 331-345.
- Maslen, G. (2000, 11/24/2000). 'Give to the rich' plan sparks outrage. *Times Education Supplement*, pp. 14.
- Maslen, G. (2002, 1/11/2002). Australia's Aborigines struggle to keep gains and exhibit their value. *Times Higher Education Supplement*, pp. 10-11.
- Mastropieri, M., Scruggs, T., & Butcher, K. (1997). How effective is inquiry learning for students with mild disabilities? *Journal of Special Education, 31*(2), 199-211.
- Matthews, M. R. (1992). Old wine in new bottles: A problem with constructivist epistemology. *Philosophy of Education, 303-311*.
- Matthews, M. R. (2000). Appraising constructivism in science and mathematics education. In D. C. Phillips (Ed.), *Constructivism in education: Opinions and second opinions on controversial issues* (pp. 161-192). Chicago, Illinois: The National Society for the Study of Education.

- McCarthy, L. P., & Schwandt, T. A. (2000). Seductive illusions: Von Glasersfeld and Gergen on epistemology and education. In D. C. Phillips (Ed.), *Constructivism in education: Opinions and second opinions on controversial issues* (pp. 41-85). Chicago, Illinois: The National Society for the Study of Education.
- McLaughlin, M. W. (1976). Implementation as mutual adaptation: Change in classroom organization. *Teachers College Record*, 77(3) (February), 339-351.
- McLeod, D. B. (1992). Research on affect in mathematics education: a reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York: Macmillan.
- Meadowcroft, G. (2000). *The influence of tracked maths classes on mathematical knowledge, mathematical self-concept and goal orientation*. Unpublished Master Educational Psychology thesis, University of Melbourne, Melbourne.
- 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.
- Midgley, C., Feldauer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 82(2), 247-258.
- Miller, R., Mills, C., & Tangherlini, A. (1995). The Appalachia model mathematics program for gifted students. *Roeper Review*, 18(2), 138-141.
- Miller, R. L., Greene, B. A., Montavalo, G. P., Ravindran, B., & Nichols, J. E. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others and perceived ability. *Contemporary Educational Psychology*, 21, 388-422.
- Miller, S. P., & Mercer, C. D. (1997). Educational aspects of mathematical disabilities. *Journal of Learning Disabilities*, 30, 47-56.
- Mills, P. E., Cole, K. N., & Dale, P. S. (2002). Early exposure to direct instruction and subsequent juvenile delinquency: A prospect examination. *Council for Exceptional Children*, 69(1), 85-96.
- Ministerial Council on Education Employment Training and Youth Affairs. (1999). *The Adelaide declaration on national goals for schooling in the twenty-first century*. Melbourne: MCEETYA.
- Moore, L., & Carnine, D. (1989). Evaluating curriculum design in the context of active teaching. *Remedial and Special Education*, 10, 28-37.

- Moore, R. (1989). *The transition problem: A study of the mathematics curriculum in the primary — secondary transition in West Gippsland*. Paper presented at the Twenty-sixth annual conference: Everyone counts, Melbourne.
- National Association for the Education of Young Children. (1987). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8*. Washington DC: Author.
- Noddings, N. (1990). Constructivism in mathematics education. In R. B. Davis & C. A. Maher & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Norwich, B. (1987). Self-efficacy and mathematics achievement: A study of their relation. *Journal of Educational Psychology*, 79, 384-387.
- O'Reilly, R., & Poindexter, S. (1999, Dec. 12). TWICE FAILED: California's lost students. *LA Times*, pp. 1.
- Pajares, F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543-578.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.
- Pajares, F. (2002). *Social foundations of thought and action by Albert Bandura, selected passages*, [Online]. Available: <http://www.emory.edu/EDUCATION/mfp/effpassages.html> [2002, December].
- Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology*, 20, 426-443.
- Pajares, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology*, 86(2), 193-203.
- Pajares, F., & Miller, M. D. (1997). Mathematics self-efficacy and mathematical problem solving: Implications of using different forms of assessment. *Journal of Experimental Education*, 65(3), 213-229.
- Paul, D. A. (1977). Change processes at the elementary, secondary and post secondary levels of education. In N. Nash & J. Culbertson (Eds.), *Linking processes in educational improvement: Concepts and applications* (pp. 7-

- 73). Columbus, Ohio: The University Council for Educational Administration.
- Perkins, D. N. (1992). What constructivism demands of the learner. In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 161-165). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Piaget, J. (1952). *The child's conception of number*. London: Routledge & Kegan Paul.
- Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (3 ed., Vol. 1, pp. 703-732). New York: Wiley.
- 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.
- Plato. (1955). *Republic*. London: Penguin.
- Plowden. (1967). *Children and their primary schools: Report of the Central Advisory Council for Education*. London: HMSO.
- Popkewitz, T. S. (1984). *Paradigm and ideology in educational research: The social functions of the intellectual*. London: Falmer Press.
- Poplin, M. (1988a). Holistic/constructivist principles of the teaching/learning process: Implications for the field of learning disabilities. *Journal of Learning Disabilities*, 21(7), 401-416.
- Poplin, M. (1988b). The reductionistic fallacy in learning disabilities: Replicating the past by reducing the present. *Journal of Learning Disabilities*, 21(7), 389-400.
- Power, C., & Cotterell, J. (1981). *Changes in students in the transition between primary and secondary school*. Canberra: Australian Government Publishing Service.
- Rasch, G. (1960). *Probabilistics: Models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Rasch, G. (1980). *Probabilistical models for some intelligence and attainment tests* (Expanded ed.). Chicago, Ill: University of Chicago Press.
- Reetz, L. J., & Hoover, J. H. (1992). The acceptability and utility of five reading approaches as judged by middle school LD students. *Learning Disabilities, Research & Practice*, 7, 11-15.

- Resnick, L. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
- Rice, J. (1997). *Explaining the negative impact of the transition from middle school to high school on student performance in mathematics and science. An examination of school discontinuity and students background variables*. (ERIC document reproduction services ED 409 183): Paper presented at the annual meeting of the American Educational Research.
- Romberg, T. A., & Kaput, J. J. (1999). Mathematics worth teaching, mathematics worth understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Rosenshine, B., & Stevens, R. (1984). Classroom instruction in reading. In D. O. Pearson (Ed.), *Handbook of research in teaching*. New York: Longman.
- Rothman, S. (2002). *Achievement in literacy and numeracy by Australian 14-year-olds, 1975-1998* (Research Report Number 29). Camberwell: The Australian Council for Educational Research.
- Rothman, S. L. (1983). *The gifted child movement in New South Wales: Public schools and the new class*. Unpublished Master of Education, University of New England.
- Schraw, G., & Brooks, D. W. (2001). *Helping students self-regulate in math and sciences courses: Improving the will and the skill*. University of Nebraska-Lincoln. Available: http://www.cci.unl.edu/Chau/SR/Self_Reg.html [2001, 6/29/01].
- Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*, 73, 93-105.
- Schunk, D. H. (1982a). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology*, 74, 548-566.
- Schunk, D. H. (1982b). Verbal self-regulation as a facilitator of children's achievement and self-efficacy. *Human Learning*, 1, 265-277.
- Schunk, D. H. (1987). Peer models and children's behavioural change. *Review of Educational Research*, 57, 149-174.
- Schunk, D. H. (1989). Self-efficacy and cognitive achievement: Implications for students with learning problems. *Journal of Learning Disabilities*, 22(1), 14-22.

- Schunk, D. H. (1990). Introduction to the special section on motivation and efficacy. *Journal of Educational Psychology*, 82(1), 3-6.
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26, 207-231.
- Schunk, D. H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational appliances* (pp. 77-99). Hillsdale, NJ: Erlbaum.
- Schunk, D. H. (1998). Teaching elementary students to self-regulate practice of mathematical skills with modeling. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 137-159). New York: The Guilford Press.
- Schunk, D. H. (1999). Social-self interaction and achievement behavior. *Educational Psychologist*, 34(4), 219-227.
- Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 125-151). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schunk, D. H., & Gunn, T. P. (1986). Self-efficacy and skill development: Influence of task strategies and attributions. *Journal of Educational Psychology*, 79, 238-244.
- Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, 77, 313-322.
- Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, 32, 195-208.
- Schweinhart, L. J., Weikart, D. P., & Larner, W. B. (1986). Consequences of three pre-school curriculum models through age 15. *Early Childhood Research Quarterly*, 1(1), 15-45.
- Sexton, C. W. (1989). Effectiveness of the Distar Reading I program in developing first graders' language skills. *Journal of Educational Research*, 82, 289-293.
- Sharp, R., Green, A., & Lewis, J. (1975). *Education and social control: A study in progressive primary education*. London: Routledge & Kegan Paul.
- Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology*, 74, 3-17.

- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46, 407-441.
- Siemon, D. (2000). *Researching numeracy in the middle years: The experience of the middle years numeracy research project*. Paper presented at the ACER Research Conference: Improving numeracy learning: What does the research tell us? Brisbane.
- Simon, M. A. (1997). Developing new models of mathematics teaching: An imperative for research on mathematics teacher development. In E. Fennema & B. S. Nelson (Eds.), *Mathematics teachers in transition* (pp. 55-86). Mahwah, NJ: Erlbaum.
- Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: The role of perceived control in children's engagement and school achievement. *Journal of Educational Psychology*, 82, 22-32.
- Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 12(3), 6-13.
- Smith, M. R., & Kramer, G. A. (1992). A comparison of two methods of test equating in the Rasch model. *Educational and Psychological Measurement*, 52(4), 835-846.
- Sowder, J. T., & Philipp, R. (1999). Learning in middle grades. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 88-158). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Stacey, K. (1997). *Implications of TIMSS for mathematics education*. Paper presented at the Australian Council for Educational Research National Conference: Raising Australian standards in mathematics and science: Insights from TIMSS, Melbourne.
- Stebbins, L., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation: A planned variation model* (Vols IV-A). Cambridge, MA: Abt Associates.
- Stein, M., Carnine, D., & Dixon, R. (1998). Direct instruction: Integrating curriculum design and effective teaching practice. *Intervention in School & Clinic*, 33(4), 227-234.
- Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap*. New York: Simon & Schuster.
- Stodolsky, S. S. (1988). *The subject matters*. Chicago: University of Chicago Press.

- Suls, J. M., & Miller, R. L. (Eds.). (1997). *Social comparison processes: Theoretical and empirical perspectives*. Washington, DC: Hemisphere.
- Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, 68(3), 277-321.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3 ed.). New York: Harper Collins College.
- Tarver, S. C., & Jung, J. S. (1995). A comparison of mathematics achievement and mathematics attitudes of first and second graders instructed with either a discovery-learning mathematics curriculum or a direct instruction curriculum. *Effective School Practices*, 14, 49-57.
- Tashman, B. (1994, November 15). Our failure to follow through. *New York Newsday*, pp. A.36.
- Thorndike, R. L. (1986). The role of general ability in prediction. *Journal of Vocational Behaviour*, 29, 332-339.
- Townsend, M. A. R., Moore, D. W., Tuck, B. F., & Wilton, K. M. (1998). Self-concept and anxiety in university students studying social science statistics within a cooperative learning structure. *Educational Psychology*, 18(1), 41-54.
- Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334-341.
- UNICEF. (2002). *A league table of educational disadvantage in rich nations* (Issue No. 4 November 2002). Florence, Italy: The United Nations Children's Fund.
- United States Department of Education. (1992). *Hard work and high expectations: Motivating students to learn*. Washington DC: Author.
- Vail, C. O., & Huntington, D. G. (1993). Behavioral interventions. In W. M. Bender (Ed.), *Learning disabilities: Best practices for professionals* (pp. 163-176). Boston: Andover Medical Publishers.
- Vitale, M., & Romance, N. (1992). Using videodisc instruction in an elementary science methods course: Remediating science knowledge deficiencies and facilitating science teaching. *Journal of Research in Science Teaching*, 29, 915-928.
- von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. London: Falmer Press.

- von Glasterfeld, E. (1996). Introduction: Aspects of constructivism. In C. T. Fosnot (Ed.), *Constructivism: Theory, perspectives and practice* (pp. 3-7). New York: Teachers College Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of psychological process*. (J. Bruner, Trans.). Cambridge, MA: Harvard University Press.
- Wakefield, A. P. (1997). Supporting math thinking. *Phi Delta Kappan*, 79(3), 233-236.
- Watson, J. M. (1987). The attitudes of preservice primary teachers towards mathematics: Some observations. *Research in Mathematics Education in Australia*(August), 48-56.
- Weaver, C. (1991). Weighing the claims about "phonics first". *The Education Digest*, April, 19-22.
- Weir, R. (1990). Philosophy, cultural beliefs and literacy. *Interchange*, 21(4), 24-33.
- Wellborn, J. G., & Connell, J. P. (1987). *Manual for the Rochester assessment package for schools*. Rochester NY: University of Rochester.
- Westwood, P. (2000). *Numeracy and learning difficulties*. Camberwell: ACER Press.
- Westwood, P. S. (1999). Constructivist approaches to mathematics learning: A note of caution. In D. Barwood & D. Greaves & P. Jeffery (Eds.), *Teaching numeracy and literacy: Interventions and strategies for 'at-risk' students* (pp. 175-189). Melbourne:: AREA Press.
- White, W. A. T. (1988). A meta-analysis of the effects of direct instruction in special education. *Education and Treatment of Children*, 11, 364-367.
- Wiersma, W. (1991). *Research methods in education: An introduction*. Boston: Allyn and Bacon.
- Wild, R. A. (1978). *Social stratification in Australia*. Sydney: George Allen & Unwin.
- Wilhite, S. C. (1990). Self-efficacy, locus of control, self-assessment of memory ability, and study activities as predictors of college course achievement. *Journal of Educational Psychology*, 82, 696-700.
- Willett, J. B. (1997). Change, measurement of. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (2 ed., pp. 327-334). Oxford: Pergamon Press.

- Williams, J. E. (1994). Gender differences in high school students efficacy-expectations/performance discrepancies across four subject matter domains. *Psychology in the Schools, 31*, 232-237.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, Ill: MESA Press.
- Yates, G. C. R. (1988). Classroom research into effective teaching. *Australian Journal of Remedial Education, 20*(1), 4-9.
- Zammit, S. A., Routitsky, A., & Greenwood, L. (2002). *Mathematics and science achievement of junior secondary school students in Australia*. Melbourne, Vic: ACER Press.
- Zimmerman, B. J., Bandura, A., & 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., & Williams, R. H. (1982). Gain scores in research can be highly reliable. *Journal of Educational Measurement, 19*, 149-154.