

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

NOTE TO USERS

This reproduction is the best copy available

UMI



Université d'Ottawa • University of Ottawa

Cross Age Learning in Primary and Junior Grades and the Self-Concept

© by Magda Stryk Therrien
1998

A Thesis submitted to the Faculty of Education in partial fulfilment of
the requirements for the degree of
Master of Arts

Faculty of Education
University of Ottawa
Ottawa, Ontario



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-36742-8

The paper is dedicated to my father,
Vladimir Karel Stryk
(1924-1980).

His example and inspiration have led me here.

Acknowledgements

I would like to thank all the numerous people who have inspired me and contributed to the process that culminated in this paper. First, I would like to thank my advisor, Dr. Raymond Leblanc. He was always encouraging, always supportive and always helpful. He gave me the confidence to make this thesis truly my own. I would also like to thank my committee members, Dr. Jan Ahola-Sidaway and Dr. Colla Jean MacDonald. Their example inspired me and their feedback guided me. Michel Brabant and Dr. Penny Guerney were invaluable in my quantitative analysis. Thank you to Dr. Jean Paul Dionne for his inspiration and his insights. I also extend a heartfelt gratitude to Gerry Denis and Bill Dolan who made my fieldwork a possibility; to Paula Dugdale and Dianna Shaffer who provided support, feedback and an academic sounding board; to my incredibly supportive husband, Paul Therrien who calmed my anxieties and knew all along that I would do it; to my family and my friends who had to put up with an ongoing dialogue about my thesis, but supported me and encouraged me throughout; to the staff, parents and students at Hillside school who allowed me to come into their lives and watch; and to the library staff at Morissette Library and the faculty and staff of the Education graduate programs who were ever helpful. Thank you all — I could never have done it without you!

Abstract

This study explored the effects on self-concept of a Cross Age Learning Program (CALP) where students in grades four to six taught mathematics to students in grades one to three. All 27 participating students were judged, by their teachers, to be having difficulty in mathematics but were capable of achieving in a regular environment. These children were divided into four groups, Older Learning Partners (OLPs), Younger Learning Partners (YLPs), Older Non Learning Partners (ONLPs), and Younger Non Learning Partners (YNLPs). The OLPs and the YLPs participated in learning sessions where each OLP was trained and then taught basic math to a YLP for approximately four months. The ONLPs and the YNLPs did not participate in the program. A self-concept measure, the Self Description Questionnaire-I (SDQ-I, Marsh, 1990) was administered three times to all students: before the program began, four and a half weeks into the learning sessions and then four weeks after that. The SDQ-I is based on a hierarchical model of self-concept which includes eight first order scales (Mathematics, Reading, General-school, Physical Ability, Physical Appearance, Peer Relations, Parent Relations and General-self), three second-order scales (Mathematics Academic, Reading Academic and Total Nonacademic) and one third-order scale (Total Self) (Marsh & Shavelson, 1985). The scale scores were compared for the two older groups (OLPs and ONLPs), the two younger groups (YLPs and YNLPs) and for the Learning Partners versus the Non Learning Partners (LPs and NLPs). A qualitative approach including document analysis, observation and interviews augmented the quantitative SDQ-I. The findings showed little conclusive data: there were no meaningful significant differences for the quantitative results. There were patterns within the interaction plots that suggested a non significant increase in mathematics self-concept over time. Qualitative results were more revealing, but showed no real pattern of improvement of self-concept, not even of mathematics related self-concept. Interviews with the learning partners did show that there were some improvements in how students felt about mathematics and about their contribution to the program.

Table of Contents

Acknowledgements	3
Abstract	4
Chapter 1 — Introduction	1
Chapter 2 — Literature Review	4
General Definitions	4
The Self-concept and Self-esteem Debate	7
Literature Review	16
Chapter 3 — Paradigmatic Positioning	24
Definitions	25
The Elements of Qualitative Research	28
Bringing Discordant Paradigms and Methodologies Together	34
Chapter 4 — Process	42
Participants	42
Setting	43
Data Collection	46
Quantitative	46
The first-order scales	48
Higher-order scales	54
Development and design of the Self Description Questionnaire - I	55
Administrating the Self Description Questionnaire I	57
Qualitative	59
Observation	61
Interviews	62
Document analysis	64
Sequence	65
Phase 1 — Preparation	66
School orientation	66
Consent and confidentiality	67
Preprogram SDQ-I administration	69
Grouping and pairing	70
Phase 2 — Training	73
Phase 3 — Learning Sessions And Concurrent Activities	76
Getting to know you	76

Learning sessions	77
Collaboration sessions	78
Within-program and final SDQ-I administrations	78
Phase 4 — Non Learning Partner Group Participation and Conclusion .	79
Analysis	80
 Chapter 5 — Quantitative Findings and Discussion	 82
Examining the Questionnaire Responses	83
Data screening	85
Assumptions	86
Normality	87
Homogeneity of Variance	92
Independence of scores	93
Sphericity	93
Transformations	96
Repeated Measures	97
Outliers removed	99
Nonparametric Tests	100
Interaction Plots Analysis	102
Summary of Quantitative Results	106
Examining the Instrument	107
The Manual	107
Questionnaire Responses	113
The Questionnaire and the Scales	114
Reliability Analysis	116
 Chapter 6 — Qualitative Findings	 121
Document Analysis	125
Observation and Interviews	129
The Older Learning Partners and the Younger Learning Partners	129
Individual Students	136
Steven	136
Kelly	140
Hailey	144
Aaron	147
Judy	152
Ryan	155
Seth	161
Sally	163
Cassie	166
Dan	168

Timothy	170
Joe	171
Individual Interviews	172
Sandra	172
Molly	175
Summary of Qualitative Results	180
Chapter 7 — Concluding Comments	183
References	194
Appendix A — Time Line for CALP Implementation	206
Appendix B — SDQ-I	208
Appendix C — Learning Logs	213
Appendix D — Sample Activities	216
Appendix E — Teacher Information Package	220
Appendix F — Consent Letter	230
Appendix G — Older Partner Training Handout	233
Appendix H — Getting to Know You Sheet	236
Appendix I — Permission to Use Figures	238

List of Figures

Figure 1 — Marsh and Shavelson's (1985) revised model of Self-Concept	22
Figure 2 — The Region of Congruence for a mix of paradigms and methods.	35
Figure 3 — The Reading, Mathematics and General-School Scales	49
Figure 4 — The Physical Abilities Scale.	50
Figure 5 — The Physical Appearance Scale.	51
Figure 6 — The Peer Relations Scale.	52
Figure 7 — The Parent Relations Scale.	52
Figure 8 — The General-self Scale.	53
Figure 9 — Scale box plots for all scales for all administrations comparing Learning Partners and Non Learning Partners.	88
Figure 10 — Scale box plots for all scales for all administrations comparing the four groups.	89
Figure 11 — Mathematic scale interaction plot for YLPs and YNLPs — unmodified. . .	103
Figure 12 — Mathematics scale interaction plot — modified to a common scale.	103
Figure 13 — Physical Appearance scale interaction plot — unmodified.	104
Figure 14 — Physical Appearance scale interaction plot — modified to a common scale.	104
Figure 15 — Cronbach's alpha score differences between older and younger participants for each scale of administrations one, two, and three.	119
Figure 16 — Comparing scale Cronbach's alpha for older and younger participants for each scale of administrations one, two and three.	120

List of Tables

Table 1 — Original and Final Composition of OLP and ONLP Groups.	72
Table 2 — Original and Final Composition of YLP and YNLP Groups.	72

Chapter 1 — Introduction

This study examined the effects of the Cross Age Learning Partners (CALP) program on the self-concept of low achieving children in primary and in junior grades. CALP was a custom designed program for students who were having difficulty in mathematics. In the program, children in junior grades tutored children in primary grades in basic math. The program was based on cross age and peer tutoring, which involve the "use of one student as 'expert' of the content area" (Scruggs & Osguthorpe, 1986, p. 187).

The target population of this study was children in primary and junior grades who were having difficulty in math. This target population was selected because resources to help these students are limited. Programs offering individual assistance to students by teachers or paraprofessionals are very expensive and many eligible students are not able to access this extra assistance (Jenkins, Mayhall, Peschka & Jenkins, 1974). As well, there are many students who are, using current standards, ineligible for extra help but would still benefit greatly from extra help if it were available.

Literature shows that cross age and peer tutoring are academically beneficial to both the tutors and the tutees (Byrd, 1990; Cohen, Kulik & Kulik, 1982; Scruggs & Osguthorpe, 1986), but the evidence is inconclusive regarding its effects on self-concept (DeRita & Weaver, 1991; Gautry, 1990; Gorrell & Keel, 1986; Labbo & Teale, 1990; Morgan & Toy, 1970). It is possible that these inconclusive findings are a result of only looking at changes to a fairly stable general self-concept. When only one small part of a

person's life is being targeted, it may be asking a lot of any program to have an impact on the general, overall self-concept of a person. It may make more sense to look at those aspects of a multidimensional self-concept (Shavelson, Hubner & Stanton, 1976) that are being addressed by an intervention. If we are implementing a cross age learning program in mathematics, then perhaps we should be examining its effects on mathematics self-concept specifically. This study sought to assess the effects of CALP on the self-concepts of the participants, particularly on the mathematics self-concept.

A combination of quantitative and qualitative approaches was used. The Self Description Questionnaire - I (SDQ-I) was used to assess students' self-concepts on three occasions, before the program, partway through, and after the program. Observation, interviews and document analysis were also used in trying to understand the relationship between CALP and self-concept.

Results were inconclusive trivial, with no meaningful significant results in the quantitative analysis. The qualitative findings fared no better. For various reasons the research failed to provide any insight into the linkage between cross age learning and self-concept.

The literature review that follows gives a brief historical perspective and describes studies on the possible benefits of cross age learning. The section gives evidence that cross age tutoring enhances academic achievement. It gives an account of the effects of tutoring on self-concept, current issues with self-concept and self-esteem terminology, and

introduces the theory of the multidimensional self-concept. Chapter three, the positioning chapter, explains the rationale and theory behind my use of a combined qualitative and quantitative approach. The process chapter, chapter four, describes the participants, the program and the data collection. Chapters five and six report the findings of the study; they detail the analyses that were carried out and the findings for the quantitative and the qualitative portions of the study respectively. Chapter five also includes discussion of the limitations of the study. In the final chapter I tie everything together and end with some concluding comments.

Chapter 2 — Literature Review

This chapter will define and discuss the main concepts associated with this thesis, specifically, terms related to cross age learning and terms and debate over self-concept and self-esteem. This will be followed by discussion of background and previous research relevant to cross age learning and the self-concept.

General Definitions

Cross age learning (usually referred to as tutoring) has been used for nearly 2000 years: Quintillian (cited in Osguthorpe & Scruggs, 1986) described cross age learning implemented in the first century; during the seventeenth century, Comenius advocated, "Qui docet, discet" — one who teaches, learns (Keatinge, 1967, p. 156); and throughout the last two hundred years, others (for example, Bell, 1817; DeRita & Weaver, 1992; Fleming, 1969; Gautrey, 1990; Lancaster, 1806) have also reported effective cross age and peer learning strategies. All these approaches to tutoring, however, differ vastly; and clustering these different uses of the term can only occur with the broadest definition of cross age learning or tutoring, that is, students teaching students. This definition is succinct, but it certainly is not descriptive enough to be a useful operational definition.

Though there are many studies that have looked at cross age tutoring, they usually have not formally defined the term. Tutoring is often used without explanation or definition, the definition apparently being self-evident. This term is then modified by an explanation of who is doing the tutoring and who is being tutored. Further description of the strategy

being addressed in a given paper often allows for a more comprehensive definition to be formulated by the reader, but rarely are formal definitions of any kind given. Again, this lack points to the need for a clear operational definition.

Hedin (1987) gives a rare and comprehensive definition of cross age and peer tutoring: "Cross-age and peer teaching and tutoring refer to a method of instruction in which students in elementary and secondary schools help each other learn. Teaching and tutoring will be used interchangeably, though tutoring usually refers to one-to-one or very small group instruction, while teaching encompasses work with larger numbers of pupils in a wide variety of subjects. Peer teaching refers to students helping classmates of their own age, while cross age teaching refers to older students helping younger ones. Tutor is the generic term for the person giving the instruction; the tutee is the one receiving the instruction" (p. 42). While this definition is quite comprehensive and general, the literature expands this definition slightly to include any learner, not just those in elementary and secondary schools¹ (Franklin, Griffin & Perry, 1994-1995; Goldstein, English, Shafer & Kaczmarek, 1997; McBride, 1995; Merrill, Reiser, Merrill & Landes, 1995; Zumwalt, 1994) and the term teaching is used in the literature to refer to tutoring (in fact, cross age teaching is the term that ERIC uses for cross age tutoring). Other definitions of tutoring include the qualification that the assistance is provided to those who are underachieving (Fresko, 1988) or needing extra assistance (Shafritz, Keoppe & Soper, 1988). An

¹Throughout this paper, the term "tutoring" will refer only to peer and cross age tutoring programs and not to programs using adults or paraprofessionals, except as noted.

extension of these definitions includes that the student in the tutoring role has relatively little specialized training for that role and that the tutoring occurs under the guidance of a teacher or supervisor (Bloom, 1976).

While none of these definitions excludes the possibility of both the tutor and the tutee learning from the experience, the term tutoring may have a stereotypical concept attached to it of only the tutee learning. There may be a stigma attached to this, that of tutees being needy of instructional assistance and the tutors providing this assistance. To combat these possible stereotypes and stigmas, I will use the term *cross age learning* instead of *cross age tutoring* to demonstrate that both the younger and the older students may learn from such a program. The term *tutor* will be replaced by *older learning partner* and the term *tutee* will be replaced by *younger learning partner*, reinforcing the concept of these students working together to learn.

Thus, for the purposes of this study, *cross age learning* is defined as older students, who have received minimal training, helping younger students who may benefit from supplemental aid, to learn a specific task, under the guidance of a teacher or supervisor. All the students in this study were in elementary school and the tutoring was one-to-one. For this study in particular, the older students who were chosen to participate were believed to be able to also benefit from supplemental aid.

Three other terms need to be defined before moving on: low achieving students, self-concept and self-esteem. *Low achieving students* are those who are having difficulty with

their academic work and whose academic performance is below the expected level for their grade, but who are capable of learning in a regular environment (Jenkins, Jewell, Leicester, Jenkins & Troutner, 1991). These are the students who were referred to above as underachieving and in need of extra assistance. While they can function without the extra assistance, they would most likely benefit from the experience in an attempt to have them perform more closely to the expected level for their grade.

The definition of self-concept and self-esteem are treated in detail in the section below since in the last three decades, there has been much controversy over the terms. For now, simple definitions will suffice. *Self-concept* is defined as "the ideas and feelings that we have about ourselves" (Johnson, 1995, p. 249) or "the individual's perceptions of himself² as a person, which includes his abilities, appearance, performance in his job, and other phases of daily living" (Good, 1973, p. 524). In this study the primary job of an elementary school student is identified as academics. *Self-esteem* is defined as "the level of positive or negative regard we have for ourselves" (Johnson, 1995, p. 250) and is strongly affected by the self-concept. In reality, it is difficult to separate these two concepts.

The Self-concept and Self-esteem Debate

In looking at the effects that the CALP program has on self-concept, there is an implication that we are looking at the effects on improved self-concept or self-esteem. Many programs and articles have attempted to improve self-concept and self-esteem and

²In quotations in this paper, original gendered wording will be retained. Otherwise, gender neutral or inclusive language will be used.

have claimed that such an improvement is a panacea or at least a prerequisite for improving children's academics and lives. This has been referred to as "the self-esteem movement" (Neuman, 1992). Over the years, this viewpoint has gained the support of many, but also the opposition of many. This has led to an inevitable debate: whether or not promoting increased self-esteem or increased self-concept³ is a good idea.

The key point around which much of the self-esteem controversy revolves is the definition of self-esteem. The self-esteem movement, in its popular incarnation, seems to advocate making children (and adults) feel good about themselves, regardless of whether or not there is a basis for such a positive evaluation. Understandably, critics of the movement have difficulty with this concept and can and have shown evidence that this kind of irresponsible self-esteem enhancement can be detrimental to the individual and society as a whole (Baumeister, Smart & Boden, 1996; Burr & Christensen, 1992; Kohn, 1994). Often the controversy seems to yield two camps: one in favour of building self-esteem, the other in opposition to building self-esteem as a goal within a classroom (or perhaps even within a society).

However, careful reading of many of the challengers of the self-esteem movement indicates that they may have nothing against helping children and adults to build genuine self-esteem, but do have difficulty with building unrealistic, inflated self-esteem which

³ Many authors do not differentiate these two terms, self-concept and self-esteem. The term self-concept as used in this thesis encompasses both terms. This will be addressed in more detail later in the paper. In this and other sections of the paper self-esteem and self-concept refer to the same concept except where a clear distinction is made and a clear contrast is necessary.

should more accurately be called egotism or narcissism (Baumeister, Smart & Boden, 1996; Burr & Christensen, 1992; Kohn, 1994; Neuman, 1992). It may also be that many of those who defend enhanced self-esteem as a desirable goal — theorists, researchers and teachers alike — imply a genuine self-esteem, based on actual achievement and accomplishment of reasonable goals; but, unfortunately, they do not articulate this qualification.

The problem may be one of the interpretation of self-esteem promotion from theory into practice. Perhaps the idea that genuine self-esteem is what we need to enhance was never clearly articulated to practitioners and curriculum designers, or worse yet, was never clearly formulated in the minds of the researchers. Clear definitions and distinctions need to be made.

Perhaps, as with other well-intentioned, well-thought out innovations, the self-esteem movement was not well implemented. Teachers, self-esteem curriculum writers and others who try to implement self-esteem enhancing measures may not have been appropriately informed or trained. The changes that occur may be implemented in a somewhat superficial and impatient manner. This kind of implementation causes the original intentions of reformers to fail in the attempt at true reform (Fullan and Miles, 1992).

Self-concept and self-esteem are such commonly used terms in psychopedagogy, yet they are often not clearly defined in the literature. When they are defined, they are defined variously and often the two terms are used interchangeably. This seems to be one of the

causes of many of the points of the controversy surrounding self-esteem.

Rosenberg, one of the authorities and pioneers of research in self-esteem and self-concept pointed out that “investigators are as far as ever from agreeing on what it [self-concept] is or what it includes” (1979, p. 3). Wylie (1968, 1974, 1989) also expressed this concern, “It is still true that part of the problem . . . lies in the vague state of theorizing in the self-concept domain” (1989, p. 2), and from reviewing more recent work on the subject, I would have to conclude that things have not changed much as far as finding universally accepted definitions of the two terms and others related to them. Many others have also identified this problem (Beane, 1991; Demo, 1985; Gecas, 1982; Scheff, Retzinger & Ryan, 1989; Shavelson, Hubner & Stanton, 1976; Smelser, 1989; Wells and Marwell, 1976).

Of the many definitions of self-concept and self-esteem some are more cited and used than others. Rosenberg has been cited by various authors in their attempts to define self-esteem and self-concept. He defines self-concept as “the totality of the individual’s thoughts and feelings having reference to himself as an object” (Rosenberg, 1979, p. 7). He goes on to say, “The ego also works to protect and enhance the self-concept, but it does not constitute the self-concept . . . [it is] the picture of the self” (p. 7). According to Rosenberg there are three broad regions of the self-concept: how an individual sees himself or herself (extant self), how that person would like to see himself or herself (desired self), and how that person shows himself or herself to others (presenting self). He defines self-

esteem as “a positive or negative orientation towards an object” (p. 53). Here he goes on to say that self-esteem does not refer to “feelings of superiority, in the sense of arrogance, conceit, contempt for others, overweening pride; we mean, rather, that he has self-respect, considers himself a person of worth. Appreciating his merits, he nonetheless recognizes his faults, faults that he hopes and expects to overcome” (p. 54). Perhaps if researchers used the full definition that Rosenberg offers, including this latter part, some of the confusion would be abated.

In *The Antecedents of Self-esteem*, Stanley Coopersmith (1967) defined self-esteem as the evaluation which the individual makes and customarily maintains with regard to himself: it expresses an attitude of approval or disapproval, and indicates the extent to which the individual believes himself to be capable, significant, successful, and worthy. In short, self-esteem is a personal judgment of worthiness that is expressed in the attitudes the individual holds toward himself. It is a subjective experience which the individual conveys to others by verbal reports and other overt expressive behavior. (p. 4-5)

Shavelson, Hubner and Stanton (1976) give a very broad definition of self-concept as “a person’s perception of himself” (p. 411). They do, however, give a much more detailed description of features critical to the construct definition, including that it is organized, multifaceted, hierarchical, stable at the general level/unstable at the situational level, developmental, descriptive and evaluative, and differentiable from other constructs.

Campbell and Lavalley (1989) define self-concept as only “the knowledge aspects of the self-schema—that is, the beliefs that an individual holds about his or her attributes. The evaluative component of the self-schema is conceptualized here as self-esteem; a self-reflective attitude that is the product of viewing the self as an object of evaluation” (p. 4). Burns (1979) defines self-esteem as “the making of a conscious judgement regarding the significance, and importance of oneself or of facets of oneself” (p. 55). According to Hills (1982) self-esteem can be considered an aspect of self-concept. Brisset (1972) says that self-esteem is made up of the process of self-evaluation and the process of self-worth. Vasconcellos (1989) says that the term self-esteem “implies a deeply felt appreciation of oneself and one's natural being, a trust of one's instincts and ability” (p. xii).

The definitions of many of these authors are so broad and generic that they certainly leave room for narcissism to be included. It seems fairly clear from these definitions that self-concept is perceptions of self, while self-esteem is the evaluative component, yet it also seems clear that in practice, this definitional distinction gets lost when talking about the two concepts. It seems reasonable then that the self-esteem controversy exists given this diversity of meanings. After Burns gives his definition of self-esteem, he goes on to say that several terms have been used interchangeably by authors, and that “the terms ‘self concept’, ‘self attitudes’, and ‘self esteem’ will be regarded henceforth as synonymous in this book” (1979, p. 57). It is important however, to make these distinctions, to clearly state definitions and to come to a consensus as to definitions for these terms. It is also

essential that we are very clear and precise when we talk about what it means to enhance people's self-concept or self-esteem.

Perhaps one way around this issue would be to use and imply the concept of genuine (Neuman, 1992), healthy, authentic or optimal self-esteem (Fitts, 1972, quoted in Garazelli, Everhart and Lester, 1993; Katz, 1993) rather than the term high self-esteem.

Genuine or healthy self-esteem is one that is based on an accurate and informed self-concept (Bednar, Wells and Perterson, 1989; Holly, 1987). It is a self-esteem trained to see and acknowledge the positive aspects of self and at the same time to see and deal with the negative aspects (Katz, no date). Whereas the narcissistic, inflated self-esteem is based on an erroneous self-concept (Holly, 1987).

Instead of aiming for higher positive self-esteem, we need to aim for healthy, genuine, optimal self-esteem for people, and we need to find ways, in the words of Stanley Coopersmith (1967), to have a "concern with developing and altering an individual's self-esteem" rather than to always have a concern with enhancing it. Developing means moving along toward a mature, healthy state, or as the Oxford Dictionary of Current English defines it, "to bring or to come to an active, visible or mature state" (Thompson, 1996, p. 235). Mature has the meaning of sensible, wise and considered (Thompson, 1996). We need to help children and adults to bring their self-esteem to such a sensible, wise, considered, and healthy self-esteem. Once researchers have agreed on a definition of self-esteem and what it means to enhance or improve self-esteem then we can have more

focussed research.

Perhaps the aim of the self-esteem movement and of trying to improve self-esteem and self-concept needs to be redefined. Part of that redefinition could consist of helping people to realize and recognize their positive qualities, abilities, and achievements without disregarding the less positive, or negative ones. There could be several related aims: to provide opportunities and encouragement for finding activities and environments where children can learn about themselves — the good and the bad; to encourage the good elements; to develop strategies for coping with and improving on the not so good (Katz, no date); and to set attainable but challenging goals so that participants could see for themselves that they are able, perhaps in areas in which they had not previously seen themselves as able. Attention must also be paid to the those whose self-esteem is inflated, to the narcissists and the egotists, to redefining their self-concepts and self-evaluations to be more realistic, genuine and healthy.

By 1970 there were an estimated 350 self-esteem curriculum programs available which directly taught affective skills within organized programs (Weinstein & Fantini, 1970). These programs are not likely to be able to make real changes in the genuine self-esteem of students because they are generally taught in a separate course focussing solely on self-esteem. Similar programs in moral development have shown that results from such programs are generally not strong nor long lasting (Lockwood, 1978; Leming, 1981). Again, any changes that might occur as a result of such programs might have a tendency to

not be genuine as the learning about and enhancing of the self-esteem is occurring in a fairly contrived setting. Holly (1987) found several programs that did seem to focus on areas which would build genuine self-esteem, though there were some that were lacking even among the ones he identified. These programs are probably somewhat useful, but not sufficient.

Genuine self-esteem needs to be nurtured within an environment (Shavelson, Hubner and Stanton, 1976), not detached from it. A person's environment certainly is an essential component in determining that person's self-esteem and must be taken into account. The whole environment of the home, school, community and world, affects a person's self-esteem and as such, the most effective self-esteem enhancement programs would encompass as much of this environment as possible.

Based on this discussion, I would argue that to develop healthy self-esteem and self-concept, people need to experience meaningful learning, understanding and appreciation of self with a natural or at least meaningful setting. Arguably, a program that is academically driven and within the context of the regular curriculum, such as the cross age learning partners, would be fit these criteria.

The purpose of this paper is not to resolve the confusion of the terms and of their meanings. Instead of trying to clean up this conceptual mess, I only present this discussion to note that the controversies exist and allow me to deal with the varying terms presented within a context. As the terms are so often used interchangeably, unless specified, in the

remainder of this paper discussion of self-esteem and self-concept are discussed as self-concept to enhance readability and enhance understanding.

Literature Review

Although a vast majority of research findings confirm that cross age and peer tutoring are academically effective (Bar & Raviv, 1982; Cohen, Kulik & Kulik, 1982; DeRita & Weaver, 1992; Fleming, 1969; Gorrell & Keel, 1986; Jenkins, Mayhall, Peschka & Jenkins, 1974; Labbo & Teale, 1990; Scruggs & Osguthorpe, 1986), and cost efficient (Fleming, 1969; Hedin, 1987), they are not used as extensively as might be anticipated. The number of these tutoring programs has, in fact, declined. In 1816, at least 100,000 children in Britain were involved in cross age tutoring but those numbers have decreased since then (Osguthorpe & Scruggs, 1986). In North America, the popularity of cross age and peer learning has dwindled since the 1960's (Hedin, 1987). However, the reasons for this decline have not been thoroughly investigated.

Despite this decline, research suggests that both learning partners, at any level, benefit from cross-age and peer tutoring programs (DePaulo, Tang, Webb, Hoover, Marsh & Litowitz, 1989; DeRita & Weaver, 1991; Fleming, 1969; Gautrey, 1990; Gorrell & Keel, 1986; Hedin, 1987). Cross age and peer tutoring have been used at all levels, from preschool (Goldstein, English, Shafer & Kaczmarek, 1997), through elementary and secondary years (Gautrey, 1990; Haluska and Gillen, 1995; Maher, 1984; Schrader and Valus, 1990) to college and university (McBride, 1995; Merrill, Reiser, Merrill & Landes,

1995; Zumwalt, 1994), and even in the workforce, such as the Navy (Franklin, Griffin & Perry, 1994-1995).

Academic gains for tutors and tutees have been found in a variety of articles:

- in the majority of descriptions of peer and cross age tutoring programs (DeRita & Weaver, 1991; Fleming, 1969; Harris, 1971),
- in qualitative studies (Gorrell & Keel, 1986),
- in experimental studies (Bar & Raviv, 1982; Cloward, 1967; Fogarty & Wang, 1982; Heward, Heron & Cooke, 1982; Jenkins, Mayhall, Peschka & Jenkins, 1974; Johnson & Bailey, 1974; Lane, Pollack & Sher, 1972; Maher, 1984; Mollod, 1970; Morgan & Toy, 1970; Robertson, 1972),
- in studies which have combined qualitative and quantitative methods (Labbo & Teale, 1990, Maher, 1986).

In certain other studies, only tutors (Allen & Feldman, 1974; Erikson & Cromack, 1972; Haggerty, 1971; Rust, 1970) or only tutees made gains (Klosterman, 1970; Shaver & Nuhn, 1971). Most of these latter studies, however, had been designed to measure the effect of the program on only the one group. Although some studies indicated no gains for children in a tutoring situation over those in a control group (Jenkins, Jewell, Leicester, Jenkins & Troutner, 1991; Kelly, 1972; Marston, Deno, Kim, Diment & Rogers, 1995), an overwhelming majority of studies have supported the concept of cross age tutoring as a worthwhile academic approach.

The academic gains have been attributed to several factors. Byrd (1990) contended that the gains in achievement "occur largely because of the individualization of instruction and the increased chances to respond" (p. 115). With individualization, the instructional pace was modified (Jenkins, Mayhall, Peschka & Jenkins, 1974) while time on task tended to increase as a result of tutoring (Delquardi, Greenwood, Whorton, Carta & Hall, 1986). The tutors were able to review and master basic material that eluded them in previous years. As students, they may have been embarrassed to study material several grade levels below their own; in the role of tutor, they had a legitimate and respectable goal in "relearning" these basics: to teach the material to their tutees (Lane, Pollack & Sher, 1972; Morgan & Toy, 1970).

Thus, it seems that such programs are ideal for the low achieving students, including the low achieving Older Learning Partners, because of the remediation benefits that have been demonstrated for the student in the tutor role. Greenwood, Dinwiddle, Terry, Wade, Stanley, Thibadeau and Delquardi (1984) found that the lowest achievers gain the most in tutoring situations, and DePaulo, Tang, Webb, Hoover, Marsh and Litowitz (1989) found that tutees did better when taught by low achieving tutors. Even if high achieving students were shown to be more effective tutors, it may have been difficult to justify using students who are already high achievers as tutors and who do not need remediation (Gerber & Kauffman, 1981).

In many studies, psycho social benefits have been attributed to cross age learning

programs. Tutors said they felt more comfortable in their own abilities, less angry and more responsible (Lane, Pollack & Sher, 1972). There was an increase in request compliance and a decrease in deviant behaviours (Martella, Marchand-Martella, Young & MacFarlane, 1995), in truancy and in tardiness (Lazerson, Foster, Brown & Hummel, 1988). At-risk young people involved in well-planned and supervised peer-tutoring situations improved their grade point averages, literacy and study skills, reading comprehension, ability to identify long-range goals, and self-confidence (Martino, 1994). Some participants had improved attitudes toward school (Roswal, Mims, Evans, Smith, Young, Burch, Croce, Horvat & Block, 1995). Children could identify with their teachers better as they acted out the adult role of being a teacher's assistant (Paolitto, 1976).

Studies of the effects of peer and cross age tutoring on self-concept, however, have yielded contradictory evidence; and it has been impossible to determine whether or not growth in self-concept and self-esteem can be anticipated from such programs. Shavelson, Hubner & Stanton (1976) and others (Burns, 1979; Wylie, 1989) do not differentiate between the self-concept and self-esteem. It seems obvious (from examining the items in the Self Description Questionnaire I [Marsh, 1990]) that what Marsh refers to as self-concept is actually a combination of self-concept and self-esteem. Therefore, in the remainder of this review of the literature, studies of both self-esteem and self-concept are discussed as self-concept.

In program descriptions, authors have appeared convinced that cross age learning

programs result in gains in general self-concept (DeRita & Weaver, 1991; Gautry, 1990; Harris, 1971; Hedin, 1987; Thelen, 1967). Qualitative studies (Gorrell & Keel, 1986) or qualitative aspects of quantitative studies (Labbo & Teale, 1990; Lane, Pollack & Sher, 1972; Maher, 1986) have found that self-concept improved for participants in cross age tutoring programs. Quantitative studies have also shown increases in self-concept for participants (Heward, Heron & Cooke, 1982; Maher, 1986; Roswal, Mims, Evans, Smith, Young, Burch, Croce, Horvat & Block, 1995; Winson, 1994). However, in a review by Cohen, Kulik and Kulik (1982), the authors found that the average effect size of such gains ranged from only very small (0.09) to small (0.18). Other studies have shown no gain in self concept as a result of participation in cross age tutoring (Morgan & Toy, 1970).

The present study sought to explore and begin to reconcile some of these contradictions. It seems that the inconclusive results regarding growth in self-concept have not necessarily been due to an absence of effect. Instead, it is possible that one of the reasons the results have been indeterminate is due to the use of inappropriate self-concept measures. Until now, the instruments used to measure self-concept in tutoring studies have only looked at general self-concept.

According to Shavelson, Hubner and Stanton (1976),

self-concept is a person's perception of him [or her] self. These perceptions are formed through his [or her] experiences with his [or her] environment,

. . . and are influenced especially by environmental reinforcements and significant others. . . . Seven features can be identified as critical to the construct definition. Self-concept may be described as: organized, multifaceted, hierarchical, stable, developmental, evaluative, differentiable. (p. 411)

The key factor for this study is self-concept as multifaceted because this factor encompasses the other six. That is, based on Shavelson, Hubner and Stanton, a multifaceted self-concept is implied when a person's information about him or herself is *organized* in a *hierarchical* fashion. In order to be hierarchical, it must be multidimensional. In order to be hierarchical it must be organized. The general self-concept is *stable* but as one moves down the hierarchy, self-concept becomes more susceptible to change because it becomes more situation specific. Self-concept is *developmental* in that, as a child matures into adulthood, the multifaceted elements of self-concept become increasingly differentiated. Self-concept encompasses both descriptive and *evaluative* elements of a person's information about her or himself. Finally, self-concept can be *differentiated* from other constructs.

The general self-concept is only one global aspect of a self-concept hierarchy, and is divided into a reading-academic self-concept, a math-academic self-concept and a non-academic self-concept (see Figure 1). Interestingly, reading and math self-concepts (subsets of reading-academic self-concept and math-academic self-concept respectively)

have been found to be uncorrelated (Marsh, Barnes, Cairns & Tidman, 1984). Based on this finding, Marsh and Shavelson (1985) tested several models for the higher-order factor structure of the SDQ-I (using a chi-square fit analysis) and found that a model that had three second-order factors, which took into account this discrepancy between reading and math self-concepts (see Figure 1), best described the hierarchal structure of the instrument (Marsh & Shavelson, 1985). This hierarchal, differentiated model of self-concept brings into question measuring differences on a theoretically stable general self-concept or even an academic self-concept, when tutoring occurs in only one subject area. By measuring the self-concept related to the subject area being taught, it is possible that a more accurate portrayal of the effects of the CALP program on self-concept can be determined.

Therefore, the present research sought to answer the question **did students in the CALP program have greater gains in subject specific self-concept than children in the**

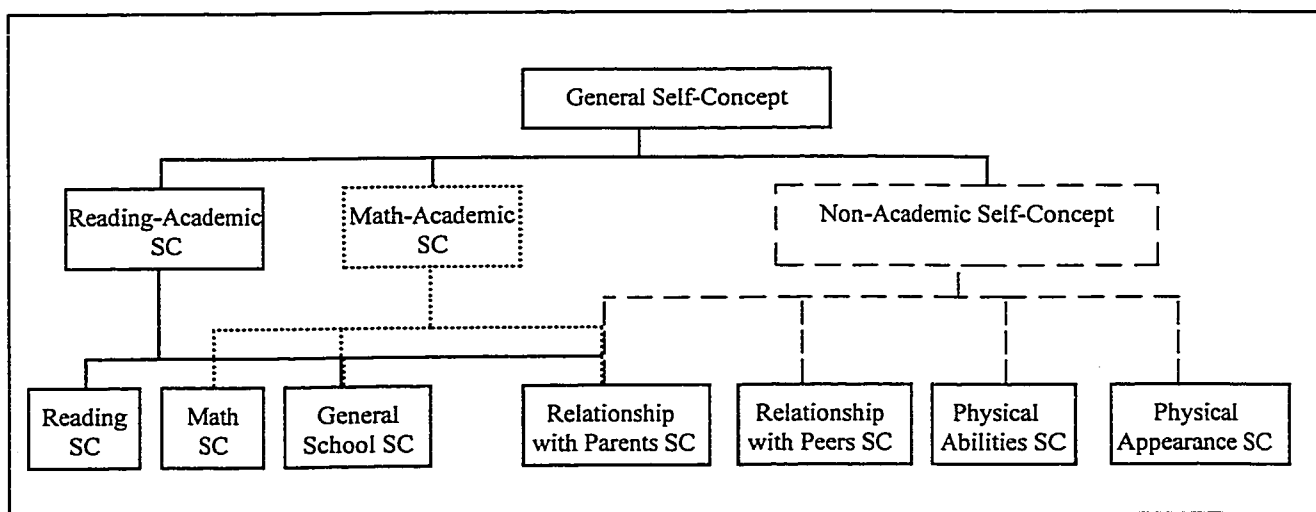


Figure 1 — Marsh and Shavelson's (1985) revised model of Self-Concept. Used with permission.

control groups? (Q_1). Further, would gains in general self-concept be less differentiated between the learning partners and the control groups than the subject specific self-concepts (Q_2). That is, did the students in the CALP program have greater gains in subject specific self-concepts (math self-concept and math-academic self-concept) than in the theoretically more stable general self-concept? If significant, did the difference in gains in general self-concept between the two groups have a relatively small effect size when compared to the difference in gains in subject specific self-concept?

The qualitative aspects of this research were designed to gain an understanding of process: **In what ways did the qualitative data help to understand, expand or refute the quantitative findings?** (Q_3). This question was designed to possibly help resolve some of the discrepancies found in the literature regarding the effect of cross age and peer tutoring on self-concept.

In designing and modifying the CALP program, careful attention was given to the factors that have been successful in previous research. Elements from some of the more successful studies and programs were combined (Gautrey, 1990; Johnson & Bailey, 1974; Maher, 1986; Morgan & Toy, 1970).

Chapter 3 — Paradigmatic Positioning

When a researcher embarks on a quantitative study, there is rarely either need or desire for an explanation as to why the researcher has chosen to work within the confines of the methods that such research entails. Methodology must be presented and explained of course, but often once there is a satisfaction that the methodology is clear and appropriate and that results are valid, the research is then usually accepted to some degree.

With research conducted with more qualitative methods, however, this appears not to be the case, at least at the present time. In general, qualitative researchers seem to account for their choice of methodological paradigm much more so than quantitative researchers. What is even more curious is the tendency of qualitative researchers to detail some of the historical, philosophical and practical reasons for qualitative research. These address the rationale for the research and methodology based on the paradigm of the researcher. Paradigm is defined here, as a basic belief system, a world view, especially regarding the nature of truth and humanity's relationship to truth. This world view guides the researcher, and is expressed in the ontology, epistemology and methodology underlying the research (Guba and Lincoln, 1994; Guba, 1990). These paradigms are what allow us to communicate meaningfully among those who share the same paradigms (Kuhn, 1962). Perhaps this is why it is so imperative for qualitative researchers to elaborate on paradigm issues — this elaboration makes communication possible with those who do not share the paradigms. Although there is no question that qualitative research is becoming

increasingly more accepted for areas focussing on the human condition, it may need more clarification to be accepted because its use is still quite new to many disciplines and such clarifications may give a common base for understanding and dialogue.

At this time it is unclear whether the apparent need for framing the researcher's paradigmatic positioning in such detail stems from researchers themselves, from outside pressure, or from both. Any of these is possible and I will not endeavour to explore this issue here. I will however, endeavour to make clear my reasons for selecting the processes that I have chosen and I will also attempt to position myself along the paradigm continuum from purely positivist to constructivist and along the methodological continuum from quantitative to qualitative research. This is necessary as, according to Guba and Lincoln (1994), "questions of method are secondary to questions of paradigm" (p.105).

Definitions

Here it is essential to define the major concepts used in qualitative research and to make distinctions between some of them. The term *paradigm* can be briefly defined as a "basic set of beliefs that guides actions" (Guba, 1990, p.17), especially in regards to truth and reality. Even with variations of this fairly simple definition, there is obviously some confusion as distinctions are sometimes made between paradigms and *perspectives*, and at other times, even within the same work, these distinctions are blurred to the point where it is implied that they are synonymous (Denzin & Lincoln, 1994). In some definitions, the term perspective is used to define the word paradigm, though usually in a more narrow

context, for example paradigm as theoretical perspective (Johnson, 1995). The word tradition (sometimes theoretical tradition) has also been used as a synonym for paradigm (Bradley & Sutton, 1993; Jacob, 1987; Kuhn, 1970; Glass, 1992; Miles & Huberman, 1994).

Paradigm has sometimes also been used synonymously with methodology, confusing or at least blurring the lines between the two. Terms such as qualitative paradigm (Bradley & Sutton, 1993) have been used with no distinction between paradigm and methods. While distinctions between paradigm and method are often confused or not sufficiently articulated, there are those who discriminate between them (Guba & Lincoln, 1994; Jacob, 1987). These authors often argue for a clarification between the two in other literature.

Methodology refers to the strategies and techniques used to collect, and analyse information or data, forming it into knowledge (Johnson, 1995; Denzin & Lincoln, 1994; Schwandt, 1994). These data can be verbal (qualitative) or numerical (quantitative). Whatever methods are employed, they need to stem from and be interlinked with the ontological and epistemological foundations of the researcher's paradigm.

There are various paradigms and these can generally be put along a continuum from positivist to constructivist. Each of these paradigms has a certain ontology, epistemology, and body of methodologies associated with it. Rather than refer to these individual paradigms, in this thesis, I will refer to those paradigms which are closer to the positivist pole on the continuum — positivist, postpositivist and the like — as quantitative

paradigms because they are usually associated with quantitative methodologies. I will refer to those paradigms which are closer to the constructivist pole on the continuum — phenomenological, constructivist, critical theory, interactionist, feminist, and the like — as qualitative paradigms because they are usually (but not always) associated with qualitative methodologies. I find this a useful separation as much of the literature treats more than one paradigm at a time giving them some collective name which breaks the paradigms along these lines. The term is used to prevent confusion of terminology and is based on examples of similar terms, such as Jacob's (1987) "qualitative traditions" where she equates tradition with paradigms and gives Kuhn's (1970) definitions for paradigm as the definitions for tradition.

When I refer to qualitative research, I will be referring to research conducted using generally qualitative methods and within the qualitative paradigms. Qualitative research has been defined as that which uses "interpretive procedures, relativistic assumptions, and verbally rather than numerically based representations" (Sutton, 1993, p. 411). Other terms that fit many of the qualitative paradigms include holistic or contextualization, examining data within a natural context; an emic perspective, seeing from the inside; understanding action and interaction; immersion, a long-term, connected field-study; flexibility; context sensitivity; and human instrumentation (Adler & Adler, 1994; Denzin & Lincoln, 1994; Miles & Huberman, 1994; Sutton, 1993; Whitt, 1991). All of the qualitative paradigms have some or all of these features, though the way these features are

interpreted for each paradigm is not the same.

The Elements of Qualitative Research

Qualitative research deals with, not only the what of human behaviour on an in-depth level, but also with the why of human behaviour, the "thick description" of what is going on. This approach is seen as necessary to truly understand human behaviour. Qualitative research gives participants credit for being human and for understanding their humanity from a very valid perspective, their own.

As our lives are constantly shaped, not just by what occurs but also by the reactions that are evoked, the nature of qualitative research is or can be shaped by the reactions that the research evokes. These shaping elements enrich, inform and become a part of both qualitative data and life, "in one sense, then, ordinary events become data when approached from a particular frame of mind - that of a researcher." (Bogdan & Biklen, 1992, p.106).

Just as history can be whatever just happened, research can concentrate on whatever is going on in the world — ordinary events, not just extraordinary events. It is this fact that makes qualitative research as rich, human and useful as it is. Ordinary events are what define each of us. It is our participation in these ordinary events that makes us human and makes us who we are. Extraordinary events and our reactions to them are interesting, but our essence, as human beings, is in day to day life. Qualitative research allows us to explore this within a natural setting.

The humanizing aspect of qualitative research is the key. We are, after all, humans.

We live within a culture and are all affected by our environment, our history, our culture. Taking these into account seems like a natural, essential aspect of research. Even if there is only one truth out there, how can that truth be found if we do not look at the human condition from within and take into account that human condition? How can we account for humanity in a dry, sparse, unconnected, abstract set of numbers? Even if cause and effect can be simplified to the positivist reality, the human element requires more than mechanistic results that have lost their humanity. We must bring the psyche, the body and the soul of the human experience to the existence represented by the radical positivist, quantitative research paradigm: "What were once abstractions became flesh and blood" (Bogdan & Biklen, 1992, p.6).

The whole mind set of the qualitative researcher must be quite radically different from the mind set of the more quantitative researcher. In fact, viewed from the more traditional, more positivist paradigms, qualitative research findings have been questioned (Kirk & Miller, 1986; LeCompte & Goetz, 1982; Huberman & Miles, 1994). Key elements include such abominations to the scientific method as flexibility, rapport and connectedness with subjects or participants, and personal involvement of the researcher (Bogdan & Biklen, 1992; Denzin & Lincoln, 1994; Miles & Huberman, 1994).

Flexibility is a key component in qualitative research (Bogdan & Biklen, 1992; Hamilton, 1994; Wolcott, 1992). Qualitative researchers have preferences or guidelines rather than a single-minded adherence to inflexible design. The open mind may be the

greatest and most essential tool in qualitative research. This non-single-mindedness may lead to wondrous, deliberate, accidental discoveries and research projects that can make real differences in the fields studied and the world at large.

In fact, flexibility is so integral to qualitative research that perhaps there cannot be a finite, contained list of qualitative research designs, methods and processes. Though there are some basic types (see Wolcott, 1992; Tesch, 1990), each study's design is unique, if not at the start of the study, then certainly after the study has influenced the design in the cyclical qualitative design process. This is quite in opposition to the positivist paradigm where the design and detailed methodology is, at least theoretically, etched in stone before research can begin.

Flexibility is not only essential with methodology, but even with theory (Strauss & Corbin, 1990). With this flexibility it is possible to "explain phenomena in light of the theoretical framework that evolves during research itself; thus you do not want to be constrained by having to adhere to a previously developed theory that may or may not apply to the area under investigation" (Strauss and Corbin, 1990, p.49).

Rapport and comfort with subjects/participants is another key element which differs radically from the positivist perspective (Fine, 1984; Janesick, 1994). Qualitative research makes no attempt to pretend that interaction does not exist; in fact, interaction is deemed essential. The comfort level of the participants determines the quality and authenticity of the observed behaviours and situation and these are the cornerstones of qualitative

paradigms. Rapport determines the quality of both the participants' solicited contributions of information as well as of observations. Though rapport can and should certainly exist in quantitative studies, because the end product sought is a numerical distillation, the process will be different and the depth of rapport necessary will be less than for a qualitative study where the data are thick and dripping with the participants' contributions, often in the form of quotes and anecdotes. To glean such useful data, rapport must be established. To glean numerical data the participant may only have to establish rapport with a consent form and a questionnaire.

In either type of research, an investigator's participation in and contribution to the study is at least partially determined by personal characteristics, attitudes and goals. Therefore, the investigator will undoubtedly influence the research. In qualitative research, the research itself is not only an investigation of whatever phenomenon or person is under study, but it is a professional, analytical and personal journey for her or him. It is possible for the investigator to discover the data, the research questions, the goal and the personal involvement as she or he travels through the project. Because the research is such a discovery process, there can be no strictly delineated maps telling her or him exactly where to go. Instead, there can be a framework, with designated areas of exploration, various methods laid out and even a recommended route, but within this map the researcher must be free to explore and allow the journey to determine some of the route, make side trips, take detours or sometimes, even follow a completely different route if the journey presents

sufficiently noteworthy landmarks and features. There are, of course, practical and ethical limits to this freedom, but the freedom must nevertheless be available in contrast to quantitative research where such freedom is proscribed.

The researcher's involvement within the context of the study is inherent to qualitative research. This involvement invites and requires some reflective practice in addition to the descriptive practices that the researcher is engaged in. This joining of the descriptive and the reflective aspects of field notes and in research in general, is one of the elements that makes qualitative distinct from quantitative research on a fundamental level. Though the actual kind of observation that takes place is different for the two types of observation, the real distinction lies in the addition of the subjective, reflective elements.

The descriptive elements really do give the reader (the researcher in many cases) a true "word-picture" (Bogdan & Biklen, 1992, p. 108) of the situation. Reading field notes can make the reader almost a part of the scene, or at least an observer to it. The narrative character of such a script can be interesting, light and enlightening. The addition of the observer's comments should add understanding and depth to the "observations" made by reading the descriptive part of the text. It can also be beneficial to see the perspective of the observer when judging the integrity of the observations.

Perhaps the greatest advantages of the qualitative research approach are the recognition of the human mind as a valid means of analysis in an investigation and of cognition, emotion and judgement as having merit. Often, in more quantitative research,

the mind of the researcher is rejected as an analytic tool; sometimes, it is even considered a hindrance to the analytic process. In the more quantitative research the human mind is certainly used in making decisions, especially in the planning stages of the study and in the interpretation of the results. The human mind is often rejected as a primary tool in the data collection and the finding of results. In qualitative research, as in life, the mind of the researcher is used as the primary analytic tool. It is used to plan, to collect data, to find results as well as to interpret them.

The task of the qualitative researcher is to convince the reader that the results of the research are plausible or credible given the data (Adler & Adler, 1994; Bogdan & Biklen, 1992; Guba & Lincoln, 1994; Miles & Huberman, 1994). Plausibility is possibility, and lends itself to reflection. Laws lend themselves to routine. Through the reflection fostered by plausibility, we can empower people to scrutinize and discover, to seek, not the one universal truth or reality, but the realities which emerge from within a context.

Credibility becomes paramount, not provability. Proof implies that what is proven is true and not just at the instant it was proven. Even the "proven" laws of Newtonian physics have been challenged by relativism. If proof of such inanimate phenomena as the "laws" of physics can be challenged and shown to be invalid under certain conditions, then what worth does proof really have? What good is proof if it is transient? And if proof is transient in the relatively systematic inorganic discipline of physics, how much more transient, even unstable, must it be in the highly dynamic study of humans, their

behaviours and their societies? Can there really ever be proof? Is credibility not more important, more appropriate, especially in such disciplines? Credibility stands on its own.

The world is 'naturalistic,' full of unpredictable (at least at this time) creatures and phenomena. Is it conceivable that in some distant future, we will have developed enough of a database on the world and all that is in it and advanced enough scientific methods, that we will be able to predict what will happen? Perhaps, however it is unlikely because humans have a will and regardless of a situation, they act unpredictably. They do not easily fit into strict predictions and stereotypes. Regardless of the possibility of precise prediction, at present we really have relatively very little such knowledge and it is naive to believe that we can really understand the world from such a perspective alone.

Therefore another, additional perspective is necessary. As the world is 'naturalistic,' should not the study of it be naturalistic as well? Life certainly seems based more on beliefs, observations, narratives, etc., than it is on things that are proven to us through positivist, scientific methods. We can ask: How can anything in a natural world be 'proven'? Why do we test phenomena through experimental means, when these phenomena occur within a natural framework? Does the supposed 'need' for experimental control justify the distortion caused by the experimental method? Can the two perspectives coexist and work to support each other?

Bringing Discordant Paradigms and Methodologies Together

While both qualitative and quantitative methods can be used by any paradigm, in order

for there to be a congruence between the paradigm and the method, the two would seemingly have to have some commonality. We can view paradigms and methodologies as two continua crossing in a matrix of possible combinations (see Figure 2).

It would seem reasonable that someone very near the positivist pole (on the positivist/constructivist continuum) would not use qualitative observational techniques (she would likely not see their value) though that investigator may choose quantitative observational techniques. Someone very near the

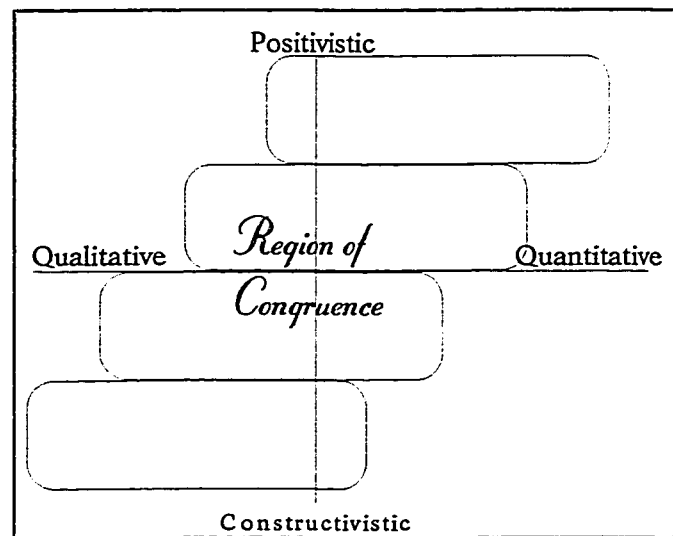


Figure 2 — The Region of Congruence for a mix of paradigms and methods.

constructivist pole would also be unlikely to use scientific experimentation (she would likely not see its value). Neither of these combinations (positivist/qualitative observation and constructivist/scientific experimentation) makes paradigmatic sense.

I postulate that there is some dynamic "region of congruence" and most researchers would fall within this region with their mix of paradigm and method. Outside of this "region of logical congruence" it may be likely that an investigator's stated paradigm or methodology is not that investigator's true paradigm or methodology and has been mislabeled or misidentified.

Die hard positivists and zealous interpretivists⁴ would both agree that various participants in an activity often have different perspectives on an activity (for example the studies of crime scenes where every witness testifies to something quite different having occurred although all of them saw the same thing happen at the same time). Perhaps the question is not whether positivists believe in one reality and interpretivists believe in multiple realities, but rather the question may be whether different perspectives on something constitute "multiple realities" or whether they are just multiple perspectives of the same reality. We need to investigate whether perspective and reality are the same concept, different facets of the same concept or entirely different concepts. Perhaps the raging "paradigm wars" of the past few decades have been less the result of different views of reality and perspectives, and more the result of a semantic misunderstanding. Increasingly, discussions of paradigm appear to reconcile some of these issues, although beginning researchers are still counselled to choose one or the other (Bogdan & Biklen, 1992, 1998).

It is possible that there is an overriding paradigm which can accommodate the perspectives of many of the existing paradigms in a multi-dimensional approach.

I propose that there are layers of reality:

- At one level, there is a positivist dimension where something has some global

⁴The term constructivist and interpretivist are interchangeable here. Both acknowledge that observations and data need to be interpreted and constructed by the social actors within the situation, including the researcher; there is no one reality, but rather multiple realities determined by context (Johnson, 1995; Guba & Lincoln, 1994; Miles & Huberman, 1995)

meaning. The meaning may be very trivial, a child's grades. I posit that everyone would agree that the child has, somehow, attained a certain grade (though not all would agree how it was obtained, whether it is important, or what it means). At some generic level it exists in a positivist sense, what I will call a *generic reality*.

However, the reality of the grade does not end there.

- At another level, some subset of the world population would agree that this grade is a good grade or a poor grade, that it is important or totally irrelevant, that the child earned the grade or the grade was given to the child as a result of systematic biases. Some subsets of people might agree on the grade's description, its purpose, its meaning. These are the general socially or perhaps politically constructed or directed meanings, those of the social reconstructionist and related paradigms, what I will call *socio-political realities*.
- And at yet another level, for me alone, and only a moment in time, in a specific situation, the grade will have a personal and specific meaning and purpose for me; the reality I have constructed myself. And every other individual, including the child, at each moment, would have a somewhat different to a radically different reality of this interaction. These would be the interpretivist realities, what I will call *individual-episodic realities*.
- There are many intermediate layers between these three, where larger or smaller groups of people would agree on a reality. These intermediate layers would fit

within the broadly-defined socio-political realities above.

Each of these levels — from general realities to individual-momentary realities — tends to increase in the number of possible realities, from very few realities for the general level to an infinite number for the individual-momentary level.

To continue with the example, without the human perspective, the socio-political realities and the individual-momentary realities of this grade or any entity, do not exist. Humans created it and continue to create it. These realities do not exist independently of people.

Perhaps then, the paradigm wars (Gage, 1989) are not a question of what is reality, but rather of which layer of reality we are looking at. With this perspective, it is possible to combine methodologies and even paradigms to look at the realities of various layers. In fact, if we want to study an entity in its entirety, it is *necessary* to use methods both qualitative and quantitative. Using the perspective of multiple layers of reality, it would be impossible to study the whole of the entity using only one paradigm because different paradigms study different layers of the reality of the entity.

Meaning, truth and reality are all along that paradigmatic continuum. I position myself within this *combined multi-level* paradigm — consisting of levels of truth and reality, “a paradigm of choices” (Patton, 1990, p.38). While I do not think there is any one truth or set of truths for all time, I do believe there are generic truths that are quite stable and a structure of more specific and less widely held truths and realities for each of these generic

truths. This view seems close to the paradigmatic positioning of Huberman and Miles who consider themselves as "transcendental realists" or those who "think that social phenomena exist not only in the objective world as well, and that there are some lawful, reasonably stable relationships to be found among them. The lawfulness comes from the sequences and the regularities that link social phenomena together; it is from these that we derive the constructs that account for individual and social life" (Huberman and Miles, 1994, p.429).

This view is considered post-positivist by some theorists such as Denzin and Lincoln (1994) who argue that "the framework remains basically unresponsive to the poststructural, constructionist, cultural studies, feminist and critical theory perspectives. These models argue against a transcendental realism. They see the world of data (empirical materialists) as one that is created in and through the interactions that occur between the observer and the observed" (p.357). I am confident that Denzin and Lincoln would give a similar argument concerning my multi-level paradigm. I anticipate that any person who holds a view that radically espouses one of the standard paradigms that are the topic of discussion in the paradigm wars, would reject the approach, each probably calling it further along the continuum heading in the opposite direction.

In Multimethod research: A synthesis of styles, Brewer and Hunter (1989) suggest that many different kinds of research be combined to explore phenomena. They say that empirical measurement is essential but incomplete, and is really "only one step, or stage, in the research process" (p. 21). While they argue for the need for the "integration of

different and often conflicting views and empirical findings” (p. 24) they only refer to differing methods and do not address the paradigmatic issues underlying the methodologies. While they do address theories, concepts and propositions, they do not go to the level of the view of reality and its relation to research and the researcher. The authors never delve into the paradigmatic underpinnings of the multimethod approach.

Many authors (Bradley & Sutton, 1993; Brown, 1992; Firestone, 1987) state that no single paradigm is appropriate for all questions, which at some level implies that all or at least many questions of different sorts might be included within one paradigm, but they do not explore the possibility further. They cite the incompatibility of various paradigms and the methods associated with them, without exploring the possibility of examining the various levels of reality using levels of a paradigm.

Using the descriptive strengths of each of the existing paradigms at different levels of the multi-level paradigm, we can capitalize on examining size and extent, of showing patterns extending beyond the sites, of finding abstracts of a study of the quantitative methods and paradigms at one level and we can capitalize on the depth and understanding of the phenomena behind the abstractions using the qualitative methods and paradigms at another level (Firestone, 1987).

The key to using qualitative and quantitative research in a complementary fashion is to ask the right questions. No paradigm, not even a paradigm such as this one, with its multiple layers of questions and understandings, is appropriate for all questions. However,

many questions asked in research can be fitted into different levels. Willingness to see these as levels and not as incompatible separates is what defines this paradigm.

“Debates and even controversies are the very process by which problems, methods, analyses, and theories become sharpened and refined. The public clash of ideas is essential to the collective scientific process” (Brewer & Hunter, 1989).

Chapter 4 — Process

This chapter provides a description of the participants, the setting, the data collection strategies and the procedure followed during the six month field study. I describe the self-concept instrument, the SDQ-I in detail and I discuss the qualitative approach that was planned.

Participants

Primary (grades 1 to 3) and junior (grades 4 to 6) children who were identified by their teacher as low achieving in mathematics were invited to participate in this study. Teachers nominated those students who were experiencing difficulty and were performing below the expected level for their grade in mathematics but who were capable of learning in a regular environment (Jenkins, Jewell, Leicester, Jenkins & Troutner, 1991). Teachers regularly identify such students in their classrooms, when putting children into groups, when writing report cards, or when designing activities. As such, they were in a good position to make the selection decisions for this study. Teacher selection was also used to prevent the disappointment that could foreseeably occur if all children in the class had been tested, but only some were invited to participate in the program and others were invited to stay in their own classroom. It also eased the burden and the requirements on the school that would have come from selection derived from class-wide testing.

Children in both primary and junior grades were chosen as participants in this study for two main reasons:

- these grades are usually found within one school and therefore children would not have to travel between schools, making the program more feasible in terms of time and money both in this inquiry and in future applications
- this age range maintains a balance between evidence that has suggested that tutees perform better academically when they are several years younger than their tutors, while children close in age may have more pleasant interactions (Devin-Sheehan, Feldman & Allen, 1979).

Twenty-seven children were identified, agreed to participate and received permission from their parents: 13 primary children and 14 junior children. All were white (this reflected the school makeup very well; there were only a handful of non-whites at the school). Sixteen were male, 11 female. Teachers reported that most participants were from middle income families with parents whose occupations ranged from farmers to professionals and that most parents had a post secondary education. There seemed to be relatively few children from single parent families in the school as a whole and this was reflected in the small number of single parent families of the participants.

Setting

Beginning my research at Hillside School⁵ was very exciting for me. I found the school quaint and the staff and students friendly. The students come from a rural setting and the large majority of them are bussed in each morning and bussed home again each

⁵All identifying names and features of the school, the board and the participants have been changed to protect the confidentiality of everyone involved.

evening.

The students are fairly well behaved and have relatively few academic problems. In one teacher's opinion, there are few of the difficulties experienced in inner city schools. There are few ethnic minorities at this school. Teachers and staff characterized the school as having few problems with theft, drugs and the like. In fact, it is common practice for a teacher to leave her purse on her desk and so far there have been no reported thefts from staff that I heard about. Even when I asked where I should leave my things, including my notebook computer, the secretary told me that things here are safe and that I should not worry about it.

This is not to say that there are no instances of these types of problems. One day in March, a teacher mentioned that one of his students had stolen \$5.00 from another student. There were fights in the playground, breaking of class and school rules and many instances of the regular mayhem that is evident in any place that 200 young people gather. However, observation of the school leads me to believe that there are fewer of the severe problems that many inner city schools are experiencing.

Several of the teachers said that they really have nothing to complain about in this school. The children are great. The only thing that they have to complain about is the building. The teachers feel that it is old and dirty.

This is also an unusual school because it is an English Language school in Quebec. The board is reportedly poor, and fairly small. The school itself is small.

It is situated in an older building that is a misshapen three stories high. It is on a plateau about half way up a large hill overlooking the village, a river and the hills surrounding it. This a lovely, peaceful and almost fictional setting.

The room where the CALP activities took place is on the lowest level of the building. It is a fairly long and narrow conference room with two tables standing side by side with a small space between them. At one end of the room there is a window, and the door is at the end of the room.

This is a fairly progressive school; the principal is a man of action. When I came, he had been at the school for three years and since then had cleaned and fixed up the school building, introduced new programs, invited prominent education speakers, applied for ministry grants, and pushed for reform and involvement in the school. He actively sought my program out after one of the teachers at the school had heard about my program and told him about it. He not only asked the principal of another school who was originally interested in the project if she would allow his school to have it if her school for some reason did not want to pursue it, but he also applied to the Quebec Ministry of Education for a grant and for more attention as a promising project. Out of "many" projects that this board put forward, the ministry only chose the CALP to pursue.

Many of the teachers here are fairly young. They seem fairly broad minded and the principal agrees that they are. After the initial meeting and with few questions the staff enthusiastically accepted my project.

This study involved approximately 6 months of field work in the school, beginning in January 1995, with half-hour learning sessions twice a week running from February to the beginning of June 1995, as outlined in the sequence section of this chapter. A timetable appears in Appendix A.

Data Collection

Data collection and analysis combined both qualitative and quantitative strategies. These are described in detail below.

Quantitative

The Self-Description Questionnaire (SDQ-I) was administered at three different times to examine the self-concept of all children participating in the study. This instrument (which is usually a paper and pencil questionnaire) was administered in the standard written form for children in grades 3 to 6. For children in grades 1 and 2, the instrument was modified into an individual, verbal questionnaire as described in a study by the developer of the SDQ and his colleagues (Marsh, Craven & Debus, 1991). They found that administering the SDQ-I as an individual interview to children 5 to 8 years of age allowed for the identification of each of the SDQ-I factors as well as giving a reliable and reasonably well-defined General-self scale.

The standard SDQ-I is a 76 item instrument, the modified SDQ-I has 64 items. The difference between the two is the twelve items that are worded negatively to control for response bias, but are not included in the scoring (other than in the calculation of control

scores, which are distinct from the self-concept scores). Previous research indicated that preadolescents' responses to negatively worded items are not valid. These items are not asked in the verbally administered version of the questionnaire (Marsh, Craven & Debus, 1991). Each item has five possible responses: 1) False, 2) Mostly False, 3) Sometime False/Sometimes True, 4) Mostly True and 5) True.

The instrument gives eight first-order scales (Mathematics, Reading, General-school, Physical Ability, Physical Appearance, Peer Relations, Parent Relations and General-self). According to the scoring calculation and summary sheet, the first seven of these combine into two second-order scales (the Total Academic and the Total Nonacademic) and then into one third-order scale (Total Self). These (except for the first-order scale, General-self) directly relate to the Shavelson, Hubner and Stanton (1976) model of self-concept.

Unfortunately, this scoring calculation and summary sheet that came with it and is reprinted in Chapter 2 of the manual does not conform to the reformulated model that research and the manual show to have "provided the most accurate description of the hierarchical structure" (Marsh, 1990, p. 39). This model is a reformulation (Marsh & Shavelson, 1985) of the model originally designed by Shavelson, Hubner and Stanton (1976). According to the revised model, the first seven first-order scales combine into three second-order scales (Reading Academic, Math Academic and Nonacademic), and one third-order scale (Total Self).

The manual does not explain this discrepancy, but, as the revised model uses the same

first-order scales, and I am not interested in looking at the children's scores in relation to norms, and it is simple to use the revised second-order scales in looking at the SDQ-I, I feel justified in having applied the revised model. The rest of the discussion will assume the revised model is being used.

The following two sections describe the first-order and higher-order scales. As the higher-order scales are merely composites of the first-order scores, I will go into some detail in describing the first-order scales but will only briefly touch on the higher-order scales.

The first-order scales

The seven first-order factors that are derived from the Marsh and Shavelson Model (1985) each have 8 items. These items appear to refer to perceived fact or to affect.

For the three academic factors, Reading, Mathematics and General-school (see Figure 3)⁶, the items for each scale are identical except for identifying what area is being addressed. For example, one item from the Reading scale is "I enjoy doing work in READING," which has a parallel item in the Mathematics scale, "I enjoy doing work in MATHEMATICS," and in the General-school scale, "I enjoy doing work in all SCHOOL SUBJECTS," though the items appear in a different order in each scale. The other seven items also have parallel wording in these three scales.

⁶Note: All figures and items used with permission. No SDQ items may be reproduced in any form without permission of the author (H. W. Marsh).

The eight items (and their item numbers) in the Reading Scale are

- 4. I get good marks in READING.
- 11. I like READING.
- 18. I am good at READING.
- 25. I am interested in READING.
- 26. I enjoy doing work in READING.
- 27. Work in READING is easy for me.
- 28. I look forward to READING.
- 29. I learn things quickly in READING.

The eight items (and their item numbers) in the Mathematics Scale are

- 13. Work in MATHEMATICS is easy for me.
- 20. I look forward to MATHEMATICS.
- 27. I get good marks in MATHEMATICS.
- 35. I am interested in MATHEMATICS.
- 43. I learn things quickly in MATHEMATICS.
- 51. I like MATHEMATICS.
- 59. I am good at MATHEMATICS.
- 68. I enjoy doing work in MATHEMATICS.

The eight items (and their item numbers) in the General-School Scale are

- 2. I am good at all SCHOOL SUBJECTS.
- 9. I enjoy doing work in all SCHOOL SUBJECTS.
- 16. I get good marks in all SCHOOL SUBJECTS.
- 31. I learn things quickly in all SCHOOL SUBJECTS.
- 39. I am interested in all SCHOOL SUBJECTS.
- 55. I look forward to all SCHOOL SUBJECTS.
- 63. Work in all SCHOOL SUBJECTS is easy for me.
- 71. I like all SCHOOL SUBJECTS.

Figure 3 — The Reading, Mathematics and General-School Scales — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 6-7) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

For these three scales, half the items referred to perceived fact or ability with words like "good," "easy," and "quickly," and half refer specifically to affect with words like "interested," "like," "enjoy," and "look forward to."

This pattern does not seem to carry over into the other scales where items tend to overwhelmingly refer to perceived fact as opposed to affect. In fact, in one scale (the peer relations scale), I would argue that all the items refer to perceived fact and that in no other scale do more than two items refer specifically to the affect rather than to perceived fact. This is somewhat accounted for in the brief descriptions of each scale provided on pages 5-7 of the manual (Marsh, 1990). Regardless, the inconsistency leads me to question the effectiveness and usefulness of such diversely defined scales. At the very least any declarations about meanings or comparisons regarding first-order self-concept scales need to include what aspects of that self-concept are being addressed.

In the Physical Abilities scale (see Figure 4) there are 2 items that specifically refer only to "run" and one item that specifically refers only to "throwing a ball," an additional item refers to "running and playing hard." This leaves only 4 items which refer to "sporting and games," "muscles," "sports," and "athlete."

The eight items (and their item numbers) in the Physical Abilities Scale are

- 3. I can run fast.
- 10. I like to run and play hard.
- 24. I enjoy sports and games.
- 32. I have good muscles.
- 40. I am good at sports.
- 48. I can run a long way without stopping.
- 56. I am a good athlete.
- 64. I am good at throwing a ball.

Figure 4 — The Physical Abilities Scale — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 5) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

The Physical Appearance scale (see Figure 5) seems to be better at addressing its stated scope. One item refers to “face,” one to “body,” and one to “features.” The rest of the items are quite generic, referring to a positive term such as “good,” “pleasant,” or “nice” with “looking.” One item specifically refers to how the child feels about how he or she looks.

The eight items (and their item numbers) in the Physical Appearance Scale are

1. I am good looking.
8. I like the way I look.
15. I have a pleasant looking face.
22. I am a nice looking person.
38. Other kids think I am good looking.
46. I have a good looking body.
54. I am better looking than most of my friends.
62. I have nice features like nose, and eyes, and hair.

Figure 5 — The Physical Appearance Scale — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 5) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

The Peer Relations (see Figure 6) scale refers to how many friends the respondent has and to how likeable she or he is. It does not address how this affects the person, nor how or whether the person likes others, nor how she or he treats or is treated by peers except where they are the child’s friends.

The Parent Relations (see Figure 7) scale seems to do a much better job of expressing various elements of the parent/child relationship, including liking, understanding, communication, fun, time, ease and desire to emulate. This scale gives a much deeper

The eight items (and their item numbers) in the Peer Relations Scale are

- 7. I have lots of friends.
- 14. I make friends easily.
- 28. I get along with easily.
- 36. I am easy to like.
- 44. Other kids want me to be their friend.
- 52. I have more friends than most other kids.
- 60. I am popular with kids my own age.
- 69. Most other kids like me.

Figure 6 — The Peer Relations Scale — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 5) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

exploration of the concept of relationship than does the Peer Relations scale, making it virtually impossible to make any statements comparing or contrasting these two relationship scales.

The eight items (and their item numbers) in the Parent Relations Scale are

- 5. My parents understand me.
- 19. I like my parents.
- 26. My parents like me.
- 34. If I have children of my own, I want to bring them up like my parents raised me.
- 42. My parents and I spend a lot of time together.
- 50. My parents are easy to talk to.
- 58. I get along with my parents.
- 66. My parents and I have a lot of fun together.

Figure 7 — The Parent Relations Scale — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 6) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

The eighth first-order factor, the General-self scale (see Figure 8), is based on Rosenberg (1965, 1979) and refers to what is usually the operational definition of self-concept, that is, a person's perceptions of herself or himself, without qualifiers that restrict the range of what areas of the self the person is having these perceptions about. This first-order score is not included in any of the second nor in the third-order scores.

The eight items (and their item numbers) in the General-self Scale are

- 29. I do lots of important things.
- 45. In general, I like being the way I am.
- 53. Overall I have a lot to be proud of.
- 67. I can do things as well as most other people.
- 70. Other people think I am a good person.
- 72. A lot of things about me are good.
- 74. I'm as good as most other people.
- 76. When I do something, I do it well.

Figure 8 — The General-self Scale — adapted from Self Description Questionnaire — I: SDQ I Manual (p. 7) by H. W. Marsh, 1990, Campbelltown, Australia: University of Western Sydney. Copyright 1990 by H. W. Marsh. Used with permission.

The General-self scale is a fairly new addition to the SDQ-I and, as such, does not have the empirical and theoretical depth of the other seven first-order factors. However, in research done using another version of the instrument, SDQ-III (for use with late adolescents), the General-self and the Total-self consistently had correlations close to .90 (Marsh, 1987). While there is no evidence to show how high this correlation is for the SDQ-I, it has been shown to be a distinct and identifiable separate factor from the other seven first-order factors.

This scale has items which refer to “overall,” “in general,” and undefined “things,” when referring to the what the perceptions are about, with descriptors such as “good,” “like,” “proud,” and “important.”

Emotional self-concept is not examined in the SDQ-I (though it is included in the SDQ-II and the SDQ-III) because Marsh found it was not possible to construct a scale due to the apparent difficulty that children in the age range had with such items (Marsh, 1990). Because of this and perhaps other unforeseen gaps with the first-order factors, I would suggest that caution be used when interpreting any of the higher-order factors as it is impossible to say what impact these unmeasured factors might have, if any.

Higher-order scales

The Reading Academic scale is the mean of the Reading, General-school and Parent Relations scales and therefore the mean of the 24 items in those scales. The Mathematics Academic scale is the mean of the Mathematics, General-school and Parent Relations scales. The Nonacademic scale is the mean of the Physical Abilities, Appearance, Peer Relations, and Parent Relations scales. The Total Self is the mean of the Reading, Mathematics, General-school, Parent Relations, Peer Relations, Physical Abilities and Physical Appearances scales. As stated earlier, the General-self is not included in any of the higher order scales.

Development and design of the Self Description Questionnaire - I

Over the years, the development of the Self Description Questionnaire and the development of the revised model of self-concept have influenced each other. In fact, one of the original purposes of the SDQ was to test specific hypotheses put forth by the original Shavelson model (Shavelson, Hubner & Stanton, 1976) and the results of research using the SDQ have effected changes in the model itself (Marsh, Byrne & Shavelson, 1988; Marsh & Shavelson, 1985; Shavelson & Marsh, 1986). As a result, Marsh claims the SDQ-I is very closely connected to its theoretical basis; though we have seen that there are some problems with this cyclical revision.

The instrument was originally designed for children in grades 4 to 6. Results of further research have shown that the SDQ-I is suitable for younger children as well, even as young as those in kindergarten, with proper modifications (Marsh, Craven & Debus, 1991; Marsh, Barnes, Cairns & Tidman, 1984, Marsh and Hocevar, 1985). Norms provided in the manual only provide data for children in grades 2 to 6, but as I am not looking at these students in relation to the norms, this does not cause a problem.

The SDQ originates in Australia and the vast majority of validation and norming has been done there. It is dangerous to use any psychometric measure indiscriminately between countries: different cultures and nationalities may have varying definitions of, structures of, attitudes about, and manifestations of self-concept. Some groups may value humility, thereby suppressing or concealing positive self-concept whereas others may value pride

and self-promotion, thereby encouraging and expressing positive self-concept. Clearly the same instrument can not be used in the same way with these different groups. While it may seem obvious which cultures and nationalities differ in which ways, in reality, we can not really know without research. Even within a fairly small geographical area there may be vast differences. Even when the country of origin is an English speaking, Commonwealth country like our own, it is important to make sure that the instrument can be used here with confidence. Therefore, it is essential to know that an instrument has validity in the country or area in which it is being used. Several studies have shown that the SDQ is valid for use in England (Marsh & Smith, 1987) and Canada (Byrne & Shavelson, 1986; Marsh, Byrne & Shavelson, 1988).

The instrument has been validated by many studies, unfortunately, the overwhelming majority of these has involved the author and originator of the instrument, Herbert W. Marsh.

The construct validity of the instrument is implied by the relationship of responses to various other predictors including demographics (sex, age, socioeconomic status), academic achievement and teacher ratings of achievement, inferred self-concept (by teachers and peers), student self-attributions, other self-concept instruments, and to self-concept enhancement interventions (Marsh, 1990).

Internal consistency for the instrument ranges from .80 to .90 for the subscales and is .94 for the total score. Of course, internal consistency information is not available for the

Reading Academic and Mathematics Academic scales because they were not addressed in the manual.

Stability over time data is only available for children in grades 4 to 6. The data is based on a reanalysis of two studies involving 671 children from grades 4 to 6 (Marsh, Smith, Barnes & Butler, 1983). Except for one correlation of .27 (the Parent Relations scale for grade 4) the correlations ranged from .51 to .74. This demonstrates that there were changes in the SDQ over time, further analysis showed that these changes were systematic rather than random.

In addition to the scoring sheet and instructions not corresponding to the best-fit model, and the concerns over individual first-order scales, there are other difficulties with this instrument. These issues will be discussed in the Results section (Chapter 5).

Administering the Self Description Questionnaire I

The students in grades 3 to 6 completed the questionnaire in groups of up to 6 children. Each child was positioned so that she or he would not be able to see other students questionnaire. The children were told that no one else, other than myself, would see their answers and that no one else would ever know what their scores had been. With questionnaires closed, I read the instructions and sample items on the front aloud and then had the children do the example item (see Appendix B — The Self Description Questionnaire I). Then I went around and ensured that everyone understood and had no questions, after which I gave the verbal instructions on pace, skipped questions and silence.

Once all these procedures had been followed, as outlined in the SDQ-I manual, the students open their questionnaires and I read each item aloud twice. Almost all students followed along after a few reminders to keep pace with my reading of items.

After having gone through the questionnaire, I gave students time to complete any skipped items and to check that each item had one and only one response beside it. When I picked up the papers, as each student indicated being done, I scanned them to ensure that each item was responded to exactly once. If not, I gave the paper back to the student to rectify.

There were some groups where several children had difficulty with a particular item and I paraphrased the item to the whole group rather than to each child individually, just to ensure that everyone understood the item.

The students in grades 1 and 2 were each given the questionnaire individually using the modified SDQ-1 interview approach. Before each child came in I filled out the identifying information at the top of the questionnaire form. Each child was assured confidentiality in the same way as the older students, followed by a simplified set of instructions, and the sample and example items. I ascertained that the student understood the different responses of true and not true, and then of always true, sometimes true and always not true and sometimes not true. Occasionally throughout the interview, and for any items where I had any doubt that the student understood the item, I asked the child if he or she understood, periodically asking the child to demonstrate the understanding by

putting the item into his or her own words. After each item, I asked the student to respond with true or not true, and then to refine that by asking “always true or sometimes true,” or “always not true or sometimes not true.” I then marked the student’s paper with the response. The students did not need to look at the questionnaire unless they were curious and asked to see it.

In a few cases where it seemed that a student was answering in a way that was not directly related to the item content, I went back and reread the questionable items after the questionnaire had been completed. I noted discrepancies, but used the original responses in calculations.

Qualitative

In addition to the collection and analysis of the quantitative data addressed above, a qualitative research approach was used in an attempt to provide the participants’ perspectives of the CALP experience and to further understand the reasons for the results obtained. That is, to help understand why what happened did happen. Three qualitative techniques were used in this study: observation, interviewing and document analysis. These are considered the central techniques in qualitative research (Greene, 1994; Punch, 1994).

These three techniques were to be framed within a modification of Miles’ (1990) prestructured case. The prestructured case is appropriate to use when the research questions being investigated are clear, as in the case of a hypothesis testing study (such as

this one) and when the time available to do the study is limited (here, because of the variety of techniques used, there is a time limitation placed on all parts of the study). The method "is aimed at solving the problems of researcher overload and labour-intensiveness, while retaining a rigorous, systematic approach to analysis" (p. 43).

The modifications included using complete write-ups of the data for observations and of transcription of recordings of meetings and interviews and having all aspects of the modified prestructured case scrutinized and audited by sceptical colleagues to combat the problems of only using very private field notes which would not be open for verification. This is a true drawback of the prestructured case, especially for a researcher who does not have extensive experience in qualitative research (Miles & Huberman, 1994). More flexible use of displays and other elements was also anticipated. In the end, the modifications that did take place were so extensive that the actual process could not be called a prestructured case at all.

The qualitative process took place throughout the study. Preliminary conclusions were explored before any quantitative analysis was undertaken. This reduced the possibility of biasing the qualitative findings with the more "hard" conclusions on the basis of the quantitative data.

Mini-cases were written up describing each of the learning partners. Group cases were also written up describing characteristics of each of the four groups. These cases formed the basis for the qualitative findings and were based on the notes, transcripts and documents collected while in the field.

Observation

In the field, I took on a "peripheral membership role" (Adler & Adler, 1994, p. 380). This role allows an insider's perspective, without the researcher becoming too involved with the activities that determine membership in the group being studied. That is, as supervisor, I participated in activities, but not in the same way as the children I was observing. This role is somewhere in the middle on Gold's (1958) naturalistic researcher continuum of complete participant, participant as observer, observer as participant and complete observer. It combines the subjective point of view of an insider with the somewhat more objective view of someone who does not completely belong to the group.

I made observations throughout the training sessions, the collaboration groups, the learning sessions and testing sessions.

The records of the observations included many of the essential elements as outlined by Denzin (1989) as was appropriate for the setting and for the particular day: specifics regarding who is involved, the interactions, the routines and rituals involved in the interactions and the setting, the social organization, the temporal and spatial characteristics of the interactions, and my interpretations of these. For example, some learning session interaction notes included the arrangement of pairs in the work space, the interaction within and between partner pairs, routines and rituals followed as well as changes in them, and the timing and duration of observed events. It was intended that all these elements, when put together, would provide insights into how the learning partners saw themselves,

each other and the situation.

Interviews

To give the participants a voice and to get elaborations and explanations on observations and on the CALP, both formal and informal interviews were included. Informal interviews were carried out as appropriate and were usually quite spontaneous and even in the case of the planned collaboration groups were indeed informal. They consisted of asking children and their teachers what they thought of the program, of any progress and of how they felt about it.

Following the program, one student was selected from each of the two learning partner groups for individual interviews. These children were selected on the basis of information that emerged from the study. However, the selection was partially limited because of logistics: The grade fours were away on a school trip and I was only informed of the planned trip one week prior to the interview week. Of the four remaining OLPs, one girl was an exceptionally good older learning partner and seemed totally confident in the role right from the beginning and I felt that it would not be appropriate to interview her because I was looking for someone more representative. Another older learning partner had such a difficult younger learning partner because of the YLP's low level of academic achievement and understanding, that I was reluctant to use her for the interview, again because I was looking for someone less of an unusual case or anomaly. That left two Older Learning Partners, both of whom were fairly competent. The deciding factor was that there had

been a change in one of the OLPs, Sandra, in relation to her participation in CALP and therefore, I decided to interview her.

For the Younger Learning Partners, I did not want to choose the learning partner that had been so particularly difficult to work with for the same reason I did not choose her older learning partner: she was just too difficult a partner. Another younger learning partner had great difficulty grasping most concepts and I was unsure whether he would be able to comprehend what was being asked of him without expending a disproportional effort. I did not want to choose the child that had been the partner of the particularly exceptional OLP, again because I did not want to choose someone who was in an extreme case in the CALP. That left three YLPs. As she was one of these three, I felt that it would be beneficial to interview the partner of the older learning partner that I was interviewing.

The teacher of each of these two children was also interviewed.

Children's interviews were used to obtain information about how they perceived the program, themselves and mathematics, and how those perceptions had changed. The interviews had been intended to garner data that were more specific than those obtained through observations, as well as to supplement and validate observational data and to build stronger data through a one-on-one approach (Miles & Huberman, 1994). Teacher interviews were intended to validate and enhance understanding of student responses, observations, test and classroom results as well as to obtain different perspectives on the effectiveness of CALP. The confirmation of observation data turned out to not be a

particularly strong function of the interview due to the limited data that were garnered through the observation portions of the study.

Two of the formal interviews were tape recorded and the tapes were transcribed. The other two interviews were not tape recorded, one because the participant did not want to be recorded and the other because the participant accidentally broke the tape recorder right at the start of the interview. Tapes from meetings and sessions were reviewed, and it was planned that those that were judged to yield data pertinent to the study were to be transcribed and any others were to be stored with a detailed contact summary sheets (Miles & Huberman, 1994) to outline what each tape contained. However, because of the sparse data throughout, all tapes were transcribed. For all transcriptions the names of the participants were changed during transcription, then the tapes were destroyed.

In all transcriptions and reports all names of participants, teachers, parents, others, the school and the school board were assigned a pseudonym and real names were never revealed outside of the school.

Document analysis

Throughout the study, learning logs (see Appendix C) and activity sheets (for some examples, see Appendix D) from the learning sessions were kept and then collected for document analysis to explore how the participants felt that the learning sessions were unfolding and to look at how things were progressing in the learning sessions.

Classroom work was to be gathered, but it became incredibly difficult to gather this

information. I had to ask several teachers many times for examples of classroom work, and in some cases still did not receive it. It seemed to be a great burden on the teachers and I decided that the information from such documents would be of limited usefulness and not worth aggravating the participating teachers over. Therefore I did not collect classroom work samples after my first attempts near the beginning of the program, and I did not use any of the samples that I did collect.

Sequence

The sequence of the project can be divided into four general steps: the *preparation phase* which included preliminary work to set the stage for the CALP program, including school orientation, selection and consent, the first administration of the SDQ-I and the grouping and pairing of participants; the *training phase*, during which the Older Learning Partners learned basics of the tutoring role; the *learning sessions and concurrent activities* during which the learning partners participated in one-to-one learning sessions and collaboration sessions and during which all participants completed the SDQ-I two more times; and finally the *Non learning partner group participation and conclusion phase* which included training for the older non learning partners, learning sessions for all non learning partners, continuing sessions for learning partners, a closing party and interviews with an older learning partner, a younger learning partner, and with each of their teachers.

Phase 1 — Preparation

The first phase of the study included familiarizing various school groups regarding the CALP program: teachers, staff, parents and participants. The preliminary steps of selecting, obtaining consent, administering the first SDQ-I, and then matching and grouping based on the Mathematics scale scores of the SDQ-I and pairing of learning partners were completed during this first phase as well.

School orientation

The program began with a teacher orientation session. This session lasted a half-hour and was scheduled adjacent to a planned staff meeting to maximize convenience for the teachers. All teachers at Hillside School were given a copy of the CALP program outline, a time line, objectives and other materials describing the program (see Appendix E). Topics discussed included the structure of the program, previous research on possible benefits and time considerations. Following the meeting, teachers had up to a week to decide whether they wanted students in their classes to participate. They then gave me the names of students who they thought would be ideal for the program using the criteria outlined above, specifically, those students who were capable of learning in a regular environment but who were experiencing difficulty and were performing below the expected level for their grades in mathematics.

All teachers of grades 1 through 6 participated in the program. Only one teacher that I know of had any kind of negative remarks: she had difficulty with the term “low achiever.”

I explained to her that the term was used as a defining term, a term had to be used that had an appropriate operational definition, and low achiever was such a term. I also explained that any correspondence with parents and any interactions with parents and children did not use that term, but referred to children who were “experiencing difficulty.”

I gave a presentation to the parent committee at one of their regular meetings. Here again, I gave a short presentation, slides and previous research findings as well as information on the requirements of the study. Afterward there was a question and answer period. The parents were inquisitive and supportive about the proposal.

I also explained the program to school staff (a secretary, a replacement secretary and a custodian) in a much more informal manner. They were supportive and very helpful for the duration of the program.

Throughout my time at the school I was available to answer any student, teacher, or parent questions or to discuss the CALP program. I was at the school two full days a week, on Mondays and Thursdays, from about an hour before classes started till after the school day ended.

Consent and confidentiality

A letter of permission was sent to the parents of all the children whose teachers had identified them as low achievers (see Appendix F). It was made clear that at any point throughout this study, if a child or his or her parents wanted that child to withdraw, he or she would be free to do so. One child did chose to leave CALP.

With a few children there was some delay in returning their consent forms and there were some apprehensions from parents. Children were to have returned the consent forms to their teachers by January 18th, the next day I came in to pick up the returned forms. There were 9 children who had not returned their forms or whose parents wanted to speak with me before making a decision. When I called the parents of these children, I was able to receive verbal permission (followed by a signed letter of consent) from 4 of the parents immediately. One parent wanted to give her child's participation more thought and another couple wanted more information about the program and to discuss it with me individually. One mother who was very wary of any kind of testing or research had a lot of questions for me. One participant's father was concerned about his child's participation in this project. Another mother wanted to discuss the program with her child's teacher. This mother made and cancelled several appointments with the teacher. Eventually, after speaking to the mother several times and being assured she would speak to the teacher and reach a decision very soon, I had to drop this child because the program was starting and no decision had been made.

I had to send more than one form home with some children, with one child I had to send three before the form actually made it home and back again.

Pseudonyms were given to all participants in the program. No real names were used in any notes, transcriptions, or reports, except in conversation about the project with school staff or parents, in correspondence and reporting to the parents of the participating

children, and in an electronically protected file where names and pseudonyms were matched.

Tape recordings were fully transcribed (using pseudonyms) and then destroyed. Field, research, and other notes generated or transcribed had pseudonyms substituted for real names as soon as was possible and before anyone else saw them. Documents that were collected from and about the participants (that is, learning logs and activity sheets) and any other materials with identifiers were kept in a safe, locked and private location to which only I have access, in my home. Any copies of such documents that were shown to others had all confidential identifiers removed.

Preprogram SDQ-I administration

Before any grouping decisions were made, or any training or learning sessions occurred, all children who agreed to and who received parental permission to participate in the study were asked to complete the Self Description Questionnaire (SDQ-I; Marsh, 1990) as a measure of self-concept (see Administrating the Self Description Questionnaire I for detailed administration procedures). The participants who were in grades three to six completed the questionnaire in small groups and those in grades one and two were administered the questionnaire individually. The administration was carried out as described in the Administrating the Self Description Questionnaire I section above.

Grouping and pairing

Children were placed into one of four groups:

- The Younger Learning Partners, the YLPs (7 children) — This group consisted of children in primary grades, 1 to 3. These children took on a tutee role throughout the program.
- The Older Learning Partners, the OLPs (7 children) — This group consisted of children in junior grades, 4 to 6. These children took on a tutor role throughout the program.
- The younger non learning partners, the YNLPs (6 children) — This group consisted of children in primary grades, 1 to 3. These children participated only in the SDQ-I during the initial portion of the program. They participated in the tutee role only after the third and final administration of the SDQ-I.
- The older non learning partners, the ONLPs (7 children) — This group consisted of children in junior grades, 4 to 6. These children participated only in the SDQ-I during the initial portion of the program. They participated in the tutor role only after the third and final administration of the SDQ-I.

Children at each grade level were placed into either the learning partner group or the non learning partner group, based on matching as closely as possible the pretest scores on the Mathematics scale of the SDQ-I for the two groups.

Because I was dealing with a quota in each grade and only a few children in each class,

to be fair, for classes with more than one participant, I generally used an ABABAB matching scheme instead of the traditional ABBAABBAA scheme to avoid putting the participants with the two lowest scores into the same group. This would have happened in all classes where there were an odd number of participants from a given grade. This would not present a significant issue if there were many children, but with only so few, it could have been problematic. In other cases matching was determined by the situation and by necessity.

In the ABABAB scheme, the scores of all participants within a grade were ranked from the highest to the lowest. Then, the highest score was put into group A, the second highest score into group B, the third highest score into A and so on. Group A was either assigned to the learning partner group or the non learning partner group, and group B was assigned to the other, depending on which (A or B) most balanced the scores of the learning partners and non learning partners.

The Mathematics scale averages for the originally formulated groups were 24.43 for the OLPs and 24.00 for the ONLPs. Later, with one student changing groups at his mother's request, and another dropping out of the program, the averages were 23.17 for the OLPs and 25.14 for the ONLPs. The unavoidable changes threw off the matching process and made the two groups quite different. The composition of the original and final OLP and ONLP groups is shown in Table 1.

	OLP		ONLP	
Original	Aaron Hailey Garry Kelly Ryan Sandra Steven	Mathematics Scale average = 24.43	Allan Judy Jacob Jeanne Jesse Krista Tara	Mathematics Scale average = 24.00
Final	Aaron Hailey Judy Kelly Sandra Steven	Mathematics Scale average = 23.17	Allan Garry Jacob Jeanne Jesse Krista Tara	Mathematics Scale average = 25.14

Table 1 — Original and Final Composition of OLP and ONLP Groups.

	YLP		YNLP	
Original	Dan John Kelsey Molly Sally Seth Timothy	Mathematics Scale average = 29.00	Darren Dorian Joe Kurt Milicent Shawn Warren	Mathematics Scale average = 27.29
Final	Dan Kelsey Molly Sally Seth Timothy <i>Joe (part way)*</i>	Mathematics Scale average = 25.43 (not including Joe's score)	Darren Dorian Kurt Milicent Shawn Warren <i>Joe (part way)*</i>	Mathematics Scale average = 31.83 (not including Joe's score)

* Joe's partner left the program on March 30th the day of the second administration of the SDQ-I and therefore Joe could not continue. He did continue with a new partner after the third SDQ-I administration. For this reason, Joe's scores are not included in the general analysis for either group.

Table 2 — Original and Final Composition of YLP and YNLP Groups.

For the younger participants, again, the groups were quite well matched originally, but because of unavoidable changes, this matching was severely affected through attrition. The original groups had similar scores of 29.00 (YLP) and 27.29 (YNLP) while the final groups had quite differing scores of 25.43 (YLP) and 31.83 (YNLP) as is shown in Table

2. Before the grouping was completed, I did have to administer the questionnaire to 2 students for whom I had only verbal permission, with the understanding that written permission would be forthcoming and that the students would not be able to actually participate until the consent form was returned. This allowed all the students to have an equal opportunity to be placed in the appropriate group.

Children placed in the learning partner groups (YLPs and OLPs) participated in the Cross Age Learning Partner program right from the beginning. These children participated in testing (three times), learning sessions and collaboration groups (after the OLPs are trained) from February until June. Older learning partners participated in training sessions immediately after the first SDQ-I administration.

Children in the non learning partner groups (YNLPs and ONLPs) participated in the three testing sessions. After the last administration of the SDQ-I in May, they too participated in CALP program as described below.

Phase 2 — Training

Junior children in both groups (OLPs in February, ONLPs in May) went through a four session training program which incorporated training techniques from successful

tutoring studies and programs. Those in the learning partner group were trained immediately following the first administration of the SDQ-I and those in the non learning partner group were trained immediately following the last administration of the SDQ-I.

The training sessions included several components, some of which are listed below:

- learning to work in a flexible manner with a learning partner (Morgan & Toy, 1970)
- learning to encourage participation from a learning partner (Morgan & Toy, 1970)
- discussions (Maher, 1986)
- modelling appropriate behaviour (Johnson & Bailey, 1974)
- role playing and simulations (Johnson & Bailey, 1974; Maher, 1986)
- learning to keep and use a planning and comment log (Gautrey, 1990; Johnson & Bailey, 1974).

The training sessions began on February 2nd. Only four students were present for the first session. One student was late for school, one was absent, and one had not brought in his signed permission slip. The children were given a handout describing the program (see Appendix G). In the first session we went through the handout quickly as I explained the program and its elements. I also explained that we would go through the elements in greater detail in the other three training sections.

In session two, six of the students were present, one was absent for the second time. I quickly reviewed what had been discussed in the first session to bring the two newcomers

up to speed and to remind those who had been at the first session, and then went on to brain storming about why children might have trouble learning math. First the children brain stormed about difficulties and then about possible solutions. I added a few items that I thought were relevant but had not been expressed by the participants.

The third session had five students, the other two students were absent, one of them for all three sessions. This session consisted of role-playing correction scenarios. Afterwards, the students brain stormed ideas on how they could help their learning partners.

After this third session, I called the mother of the boy who had missed all three sessions. She asked that her son not participate in the program with the first group (the OLPs), but that he start with the second group (the ONLPs) since he had missed all the training so far. This left me with one OLP too few. I looked at the SDQ-I scores again, rearranged the scores to fit the ABA pattern and determined a replacement OLP. I worked individually with this girl to cover the material of the first two training sessions. Then, for the third session, I had her work with the other boy who had missed the session. This gave both of them the opportunity to role-play like the group did and prepared them for the fourth training session with the rest of the group.

During the fourth session I explained about the "Getting to know you" session and interview. We talked about how to show your partner you care and how to show your partner how he or she is doing. And then the OLPs had the opportunity to practice.

At the beginning of each of the last three sessions, I briefly reviewed what had been done up to that point in the training.

In addition to pre-service training, OLPs received in-service training in the form of collaboration groups (Labbo & Teale, 1990). In these group discussions learning partners discussed ideas, problems, and tutoring in general. Heward, Heron & Cooke (1982) used a similar concept ("Tutor Huddles") and described these as tutoring sessions for the children acting as tutors, as did Lane, Pollack and Sher (1972), using what they called, "rap sessions." Primary students met in separate collaboration groups to discuss their learning session experiences. I acted as facilitator during these discussions with both groups.

Phase 3 — Learning Sessions And Concurrent Activities

Following training, OLPs and YLPs (and later, ONLPs and YNLPs) were matched, based on grade level: grade 6s were paired with grade 3s, grade 5s with grade 2s and grade 4s with grade 1s. This maintained a three grade difference between the older and the younger learning partner, in congruity with previous research (Devin-Sheehan, Feldman & Allen, 1976).

Getting to know you

The first meeting of the partner pairs was an orientation meeting called "Getting to Know You." During this session, the Older Learning Partners used a "Getting to Know You" questionnaire (see Appendix H) to gain some background knowledge about their

Younger Learning Partners. With this information, the OLPs could be more comfortable with their YLPs and gear conversation and activities to the YLPs' interests and background.

Learning sessions

Learning sessions occurred two times a week for thirty minutes in an empty meeting room. Learning partners sat in pairs and worked on their daily activities. I supervised these sessions.

The learning partners left their classrooms during class time to engage in the learning partner sessions. The children in the non learning partner groups did not leave the classroom, but engaged in normal classroom activity until after the last administration of the SDQ-I. Learning sessions were held during class time to ensure all children who could benefit from the program had the opportunity to do so (Hedin, 1987) and also to show the children that the learning sessions were considered important enough to take up class time.

There were a total of 17 possible learning sessions before the final administration of the SDQ-I and another six possible learning sessions after for LPs. There were five possible learning sessions for NLPs. Of the total 23 sessions, the six LP pairs completed anywhere from 18 to 22 with an average of 19.33 sessions (four partner pairs completed 19 sessions) or 84.06%. The NLP pairs completed from three to five of the sessions, with an average of 4.00 sessions (or 80.00%).

Each day, after the session, all learning partners filled out a learning log (see Appendix

C). Originally, this process was taking far too long. The Older Learning Partners were filling out planning information and information on what they did that day. Then each partner was filling out his or her log on how things went that day, and in addition to not knowing what to say, the younger partners were taking a very long time to write out their answers to the question. The older partners helped their younger partners to complete their logs, sometimes prompting for answers and writing out what the younger learning partner said. These learning logs required answering a multiple choice question of how the session went and very short answers. After the first few sessions, and in seeing how long it was taking, I decided to fill out at least part of the log, the planning section, for the learning partners to speed up the process.

Collaboration sessions

Collaboration sessions occurred throughout the program. These were used to help the OLPs and to further their training, and to get the YLPs to discuss their learning experience. In addition to being helpful for the program, some of these collaboration sessions were used as informal group interviews.

Within-program and final SDQ-I administrations

The second administration of the SDQ-I occurred after nine learning sessions. Because of absences, the majority of learning partners had completed seven or eight learning sessions and the remaining pair had completed six (average was 7.14 sessions). This second SDQ-I is what I will call the within-program administration. Within-program

assessment is essential as Niedermeyer and Ellis (1970) found that variations existed between results from testing partway through and at the end of the Ontario-Montclair School District's tutoring program. Four weeks was chosen for the within-program testing because a meta-analysis of 65 studies done by Cohen, Kulik and Kulik (1982) showed that programs of up to 4 weeks show the highest average effect size for both tutor and tutee achievement. This seemed a reasonable time at which to look at the effects on self-concept changes as well. Because the present study was fairly long term, it was important to assess what was happening with the self-concept both on a short term and a long term basis.

In May, after each learning partner pair completed 14 to 16 learning sessions (14.67 average), the SDQ-I was administered for the third and last time. Right after this administration, the non learning partners participated in training and then in learning sessions. The learning partners continued in their learning sessions. The qualitative data collection continued for the learning partner groups.

Phase 4 — Non Learning Partner Group Participation and Conclusion

After the third and final administration of the SDQ-I in May, the junior students in the non learning partner groups were trained as Older Learning Partners. Following this training, all the students in the non learning partner groups engaged in CALP learning sessions. I supervised these students along with the learning partners, but qualitative data collection focussed on the students in the learning partner groups.

Analysis

Analysis utilized both the quantitative and the qualitative approaches. Except for the initial administration of the SDQ-I which was used to determine placement of students, quantitative analysis was left until after the preliminary qualitative analysis was done and preliminary conclusions were drawn from it without the influence of the quantitative results.

The observations, learning logs, discussions and interviews were looked at from a qualitative perspective. The analysis of qualitative data consisted of the following steps (Miles & Huberman, 1994):

- *Creating fully written-up data.* Observations were recorded during sessions on a notebook computer. These were checked, filled in and "cleaned up". Brief notes were taken during tape recorded interviews and meetings. Transcriptions of tapes was also part of this phase. Names were changed to pseudonyms to protect the identity of participants.
- *Coding.* The fully written-up data were coded for recurring themes and patterns. Reflective or other notes were added to the data when these are appropriate.
- *Data Displays.* Very few data displays and graphical summaries of the data were needed as not much useful qualitative data was gleaned from the study.

Consultation with critical friends was minimal as a result of the limited usefulness of the data.

- *Analytic Text.* Analytic text was written up, mostly in the form of mini case studies and summaries, including data gleaned from document analysis.

After initial qualitative conclusions were reached and reviewed, quantitative analysis was done. A repeated measures was the primary analysis. Before the repeated measures design could be used, the data had to be screened, then described and the assumptions for the analysis had to be tested. Nonparametric designs were also done as the assumptions of the repeated measures design were not met. The instrument was also examined for validity and reliability for this group of participants.

Chapter 5 — Quantitative Findings and Discussion

This chapter presents and discusses the quantitative findings of the Self Description Questionnaire-I (SDQ-I). The quantitative analysis portion of the findings covered many steps including examining the questionnaire responses, data screening, testing of the assumptions, attempts at transformations of data, the primary analysis — repeated measures, reanalysis with outliers removed, some nonparametric measures, post hoc analysis using interaction plots and a more in depth look at the instrument. In this quantitative analysis, I refer to:

- eight first-order scales (Physical Abilities, Physical Appearance, Peer Relations, Parent Relations, Reading, Mathematics, General-school, and General-self),
- four second-order scales (the Total Nonacademic, Reading Academic and Math Academic, which are the three second-order scales that best fit the research thus far (Marsh & Shavelson, 1985), as well as the Total Academic which is part of the scoring sheet for the SDQ-I, but is not part of the best-fit model for the scales), and
- one third-order scale (Total Self).

There were three administrations of the questionnaire, referred to as administration one, administration two and administration three. This gave a total of 39 scale scores for each participant.

The analysis looks at these scales in three sets: between the Learning Partners and the Non Learning Partners, between the Older Learning Partners and the Older Non Learning

Partners, and between Younger Learning Partners and the Younger Non Learning Partners. The original groups were well matched in terms of size and of scores on the Mathematics scale from administration one of the SDQ-I. Due to unavoidable circumstances, this changed. The final groups were not particularly well matched on the administration one Mathematics scores nor in size of groups. The OLPs had 6 participants due to the loss of one participant, Ryan, whose leaving was plainly associated with the program conditions. He left because he was not enjoying himself. The inherent loss of his partner, Joe, may also have been problematic in that his partner had to deal with Ryan's attitude most directly and either his loss or his retention in the program could have affected results. However, Ryan's partner was dropped after the second administration of the questionnaire (which was administered just before Ryan left the program). Joe later participated in the CALP with the NLPs, but none of his scores were included in analysis. One questionnaire of another YLP was lost. This left only 5 in the YLP group. The ONLPs had 7 and the YNLPs had 6 participants who completed all three questionnaires. Analyses were also done comparing all the Learning Partners and all the Non Learning Partners. These had 11 and 13 participants respectively.

Examining the Questionnaire Responses

At each administration of the SDQ-I, I checked to make sure that all statements had been responded to by each student and invited each student to respond to any missed items before leaving. Even so, there were a small number of items that had been missed. The

manual recommended that in such a case the item mean for the norm group be entered as the student's score for that item. This did not make any sense to me, as some of the children were scoring consistently below the means of the norm group and thus I decided to use the mean of the score on the remaining items in the first-order scale for that child to fill in the missing item score. This occurred for less than ten items for all the questionnaires for all three administrations (a total of 936 scores or 4608 items positively worded items, or 5076 items all together).

Then I examined the control scores of the questionnaire as put forth in the manual. The SDQ-I has six control scores to check for inappropriate responses (Marsh, 1990). The participants in the CALP program did not fare too well on these control scores and thus it seems that the instrument is not appropriate for this sample, even though it has been validated, to some extent, here in Canada (Byrne & Shavelson, 1986; Marsh, Byrne & Shavelson, 1988).

Thirteen of the 24 participants had at least one control score in the three administrations which indicated inappropriate responding. The Younger Learning Partner group had the most questionable scores (a mean of 2.6 per person over the three administrations, as opposed to a mean of 0.7 to 1.0 for the other three groups). This inappropriate responding of the YLPs is further supported by assumption testing and reliability analysis discussed below. This large discrepancy is particularly telling as the younger participants had only four control scores for each administration and the older

participants had six, because the younger participants did not answer the negatively worded items that are the basis of two of the control scores. So they had a total of 12 control scores each over the three administrations while the OLPs and the ONLPs had 18 each.

There were relatively few control scores indicating inappropriate responses for the third administration of the questionnaire (a total of 4 individual scores). But the first and the second administration had virtually the same number (a total of 13 individual scores on the first and 14 individual scores on the second) so the drop does not seem to be a result of comfort with or learning of the questionnaire, or one would have expected a drop from the first questionnaire to the second. However, some practice effect (Keppel, 1982) may be present for the third administration as it was given only four and a half weeks after the second administration, while the second was given nine weeks after the first.

Data screening

Before doing any analyses, preliminary data screening was done. This included screening the data for incorrectly entered data, for outliers, as well as checking the assumptions behind the main analysis used. In this case, the procedure also included transforming the data several ways to try to compensate for situations where the data did not conform to the assumptions (these transformations are discussed in a separated section below).

To check for accuracy of the data entry, I examined the maximums and minimums

and frequency counts for all items, as well as all the scales to ensure that no scores were out of the range of possible scores (1 to 5 for each item, and 8 to 40 for each scale).

Because this study looked at the results of the same questionnaire over three different administrations, a repeated measures design was the primary quantitative data analysis method used. A repeated measures design allows an investigation of a single measure administered more than one time. It allows for testing the null hypothesis both for between-subjects factors and for within-subjects factors, as well as for interactions between factors (Keppel, 1982; SPSS, 1997). The next chunk of data screening was used to examine the assumptions of the repeated measures design.

Assumptions

The repeated measures design is based on four assumptions: normal distribution (Howell, 1985; Keppel, 1982; May, Masson & Hunter, 1990), homogeneity of variances (or homoscedasticity (Darlington, 1990)) (Howell, 1985; Keppel, 1982; May, Masson & Hunter, 1990), independence of scores (Keppel, 1985, May, Masson & Hunter, 1990), and sphericity (Girden, 1992, Huynh and Feldt, 1970) or compound symmetry, depending on whether the design is univariate or multivariate (Darlington, 1990). Fortunately, however, the sampling distribution of F is very robust and therefore it is quite insensitive to “even flagrant violations of the assumptions” (Keppel, 1982, p. 85).

Normality

The assumption of normality states that the scores for each group come from normally distributed populations. However, it is not actually necessary for a population to be normally distributed, as long as it is either unimodal or somewhat symmetrical (May, Masson & Hunter, 1990). In fact, with equal sample sizes, even quite severe departures from the normal distribution apparently have very little effect on the distribution of F (Howell, 1985; Keppel, 1982). With significance of F set at $\alpha = .05$ and with the most deviant distributions, the actual level for alpha would be no greater than .08 (and this is with sample sizes of only 3 to 5) (Norton, 1952, cited in Lindquist, 1953). When very severe violations of the normality assumption are present in the sample, values of F which have a significance close to alpha may be suspect. So, in practice, the assumption of normality is extremely flexible because the consequences of departure from the assumption are negligible. As well, for small samples, like the ones in this study, assessment of normality is difficult and imprecise (Afifi & Clark, 1996). However, it is still important to check the assumption.

The box plots of the first-order and higher-order scales at all three administration times seem to show that for the Learning Partner and Non Learning Partner groups, for the most part, the assumption of normality has been met. This is shown in Figure 9, which shows the box plots both the LPs and the NLPs for each scale for each of the three administrations.

However, for the smaller groups this is not the case, most especially not for the YLPs, particularly at administration two. Administration two also gave somewhat non-normal distributions for the other groups as well, though the deviations did not seem as severe.

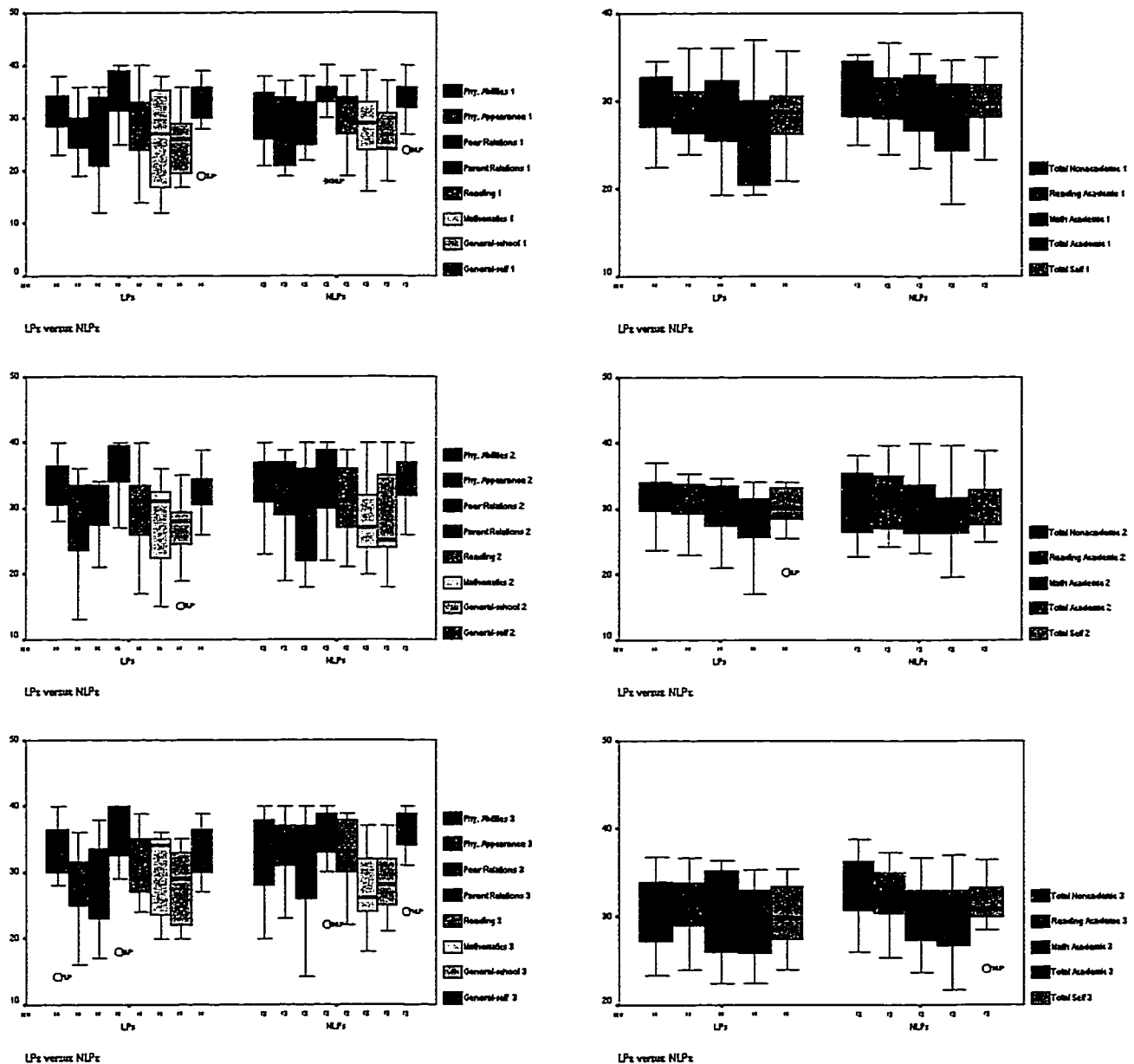


Figure 9— Scale box plots for all scales for all administrations comparing Learning Partners and Non Learning Partners.

For the YLPs, the other two administrations caused quite a bit less deviance but still non normal distributions. The other three groups had box plots for administrations one and three which appeared to be quite consistent with normal distributions (see Figure 10).

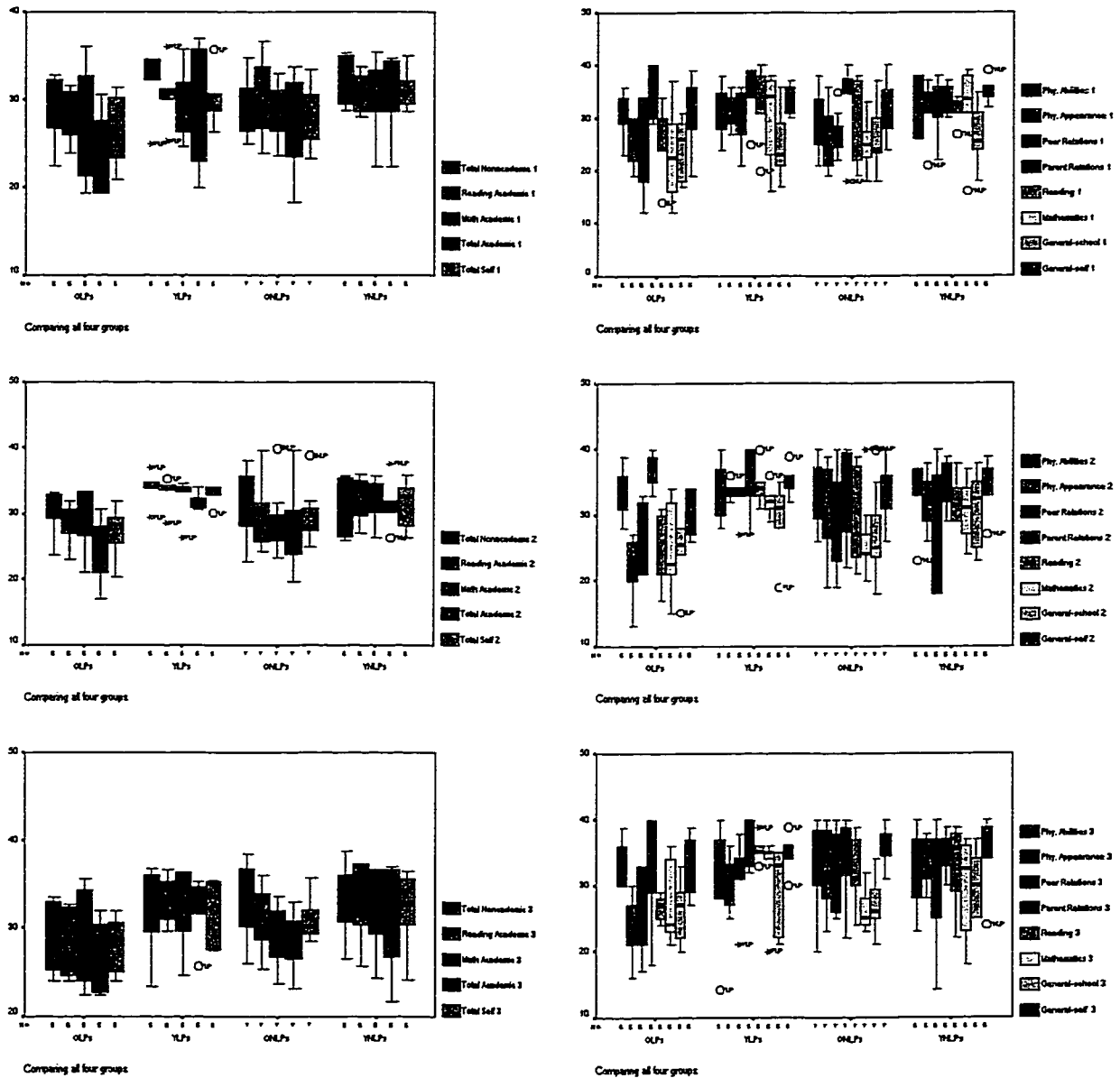


Figure 10— Scale box plots for all scales for all administrations comparing the four groups.

Further tests for normality were conducted as well. With the Peer Relations, Parent Relations, Mathematics, and General-self scales, for at least one of the groups and at least one of the three administrations, had significant Shapiro-Wilks statistics suggesting that their distributions significantly deviate from normal.

I also looked at skewness and kurtosis. Skewness refers to the asymmetry of the distribution, where a normal distribution has a skewness of zero (May, Masson & Hunter, 1990; Tabachnick & Fidell, 1983). Values of skewness of ± 2.58 constitute distributions that are significantly skewed (at $p \leq .01$) (Tabachnick & Fidell, 1983). I looked at the distributions for each group and none of the scales had skewness of that magnitude, though some of the skewness measures might have been significant at $p \leq .05$, but the value of skewness for a significance of .05 has not been discussed in the literature. There have been indications that a skewness value of over one is significantly different from a normal distribution though there were no indications of what the value of p would be (Afifi & Clark, 1996; Weber, 1968). Using this criteria, all the groups have some scores that differ from normal. Again the YLPs had the most scale distributions that were skewed more than ± 1.000 — 15 negatively and one positively. The absolute skewness went up to 2.175. The two Non Learning groups each had 10 scales with skewness of more than ± 1.000 . The YNLPs had 9 negatively skewed and one positively and the ONLPs had 3 negatively skewed and 7 positively skewed scales. The absolute skewness for the YNLPs went up to 1.827, and for the ONLPs, up to 2.202. The OLPs had 7 negatively skewed scales over

± 1.000 up to -1.873 . The larger groups fared somewhat better, the NLPs had only 3 such scores, all skewed negatively and going up to -2.107 . The LPs had 8 negatively skewed scores over ± 1.000 , up to -1.1673 . Thus with the more conservative criteria, skewness is an issue.

Kurtosis is a measure of how clustered around the centre a distribution is, that is whether it is fairly peaked, fairly flat or somewhere in between like in the normal distribution (where the value of kurtosis is 0). While it is not clear what a significant value for kurtosis is, in my sample there were a number of distributions that had a kurtosis of ± 3.000 . Not surprisingly, the YLPs had the highest number, with 8 scales, the YNLPs had 3 scales, the ONLPs and the LPs had two, and the OLPs and NLPs had one scale each. As a whole group, the participants had no kurtosis scores of more than ± 3.000 . Only the Parent Relations 1 scale had two groups for which the kurtosis was over ± 3.000 , and one of the groups (ONLPs) was a component of the other (OLPs). All other scales had such kurtosis for only one group.

Clearly, the assumptions for normality have not been met by all the scales for all administrations for all groups. This is probably partly due to the sample, especially the YLPs, but is also most likely partly due to the sample sizes, especially between the four groups, where the largest discrepancies occurred.

Homogeneity of Variance

A rule of thumb for the homoscedasticity assumption is that if sample sizes are unequal and the variance estimate of one of the samples is no more than four times that of the other sample, then it is permissible to violate the assumption (Boneau, 1960, Howell, 1985). In fact, based on findings of the robustness of F and how little it is affected by unequal variances, unless sample sizes differ considerably (by at least two to one), researchers rarely even test this assumption (Keppel, 1982). Unfortunately, this very loose interpretation of the assumption of homoscedasticity applies when sample sizes are not too small, specifically, greater than or equal to seven (Ramsey, 1980).

In the analysis comparing the two older groups and comparing the two younger groups, sample sizes are only 5 to 7, and so care must be taken. The comparison of the Learning Partners and the Non Learning Partners sample sizes are large enough and equal enough to not be a cause for concern. Even using just the rule of four times variance, of the 39 scales, none of the variance differences would have been too large in comparing the LPs and the YLPs. However, in comparing the two older groups, four scales had variance differences of over four times the smaller variance and when comparing the two younger groups seven scales had variance differences of over four times the smaller variance.

Levene's Test of Equality of Error Variances shows that only one scale has a significant F which indicates that the error variance of the dependent variable is not equal across groups for the Peer Relations scale during the second administration.

Independence of scores

The third assumption, independence of scores, was clearly not met. I used a matched groups design (May, Masson and Hunter, 1990) in which I attempted to match scores on the first administration of the Mathematics scale to determine which group (Learning Partner or Non Learning Partner) each child was placed in, in an attempt to equalize the beginning means for each group (see Tables 1 and 2 in chapter 4). As such, the repeated measures procedure results can only be seen as estimates of the randomized results (May, Masson & Hunter, 1990). Of course the normality and some of the homogeneity of variance results have already determined that results can be no more than an approximation.

Sphericity

A repeated measures design also must meet the assumption of sphericity (or in multivariate methods the assumption of compound symmetry). Sphericity assumes that for each pair of conditions, the variance of difference scores is the same, or put another way, that the covariance between the repeated treatments in the population are the same (Chen & Dunlap, 1994; Keselman & Keselman, 1988; Morris, 1981). Compound symmetry states that when between-subjects factors are held constant all the correlations between trials are equal (Darlington, 1990). Darlington goes on to explain that in practice, compound symmetry is what is being actually examined whether the univariate or the multivariate method is used and he goes on the deal with compound symmetry only.

In at least the univariate method, moderate deviations from compound symmetry have little effect and the fewer the repetitions of a repeated factor, the more minor the effects of the violation of compound symmetry. Darlington suggests that the assumption of compound symmetry seems to be unreasonable, as testing the null hypothesis is actually testing for a lack of differences in similarity between trials. If the null hypothesis is false, there will of course be some differences in similarity between trials.

Regardless of the effects on validity that violations of the compound symmetry assumption have, it is possible to automatically assume compound symmetry by using an extremely strong null hypothesis: no treatment has any effect on any between or within subjects condition and variations between instances are due to random fluctuations. If these variations are due to chance then all correlations would equal zero and thus be equal to one another. This would satisfy the assumption of compound symmetry. Because the assumption is correct when this very strong null hypothesis is true, then the test is valid and the assumption is met even when the null hypothesis is false (Darlington, 1990). Moreover, compound symmetry has been found to be a sufficient condition for the F test, but not a necessary one, while the less restrictive sphericity was found to be necessary (Girden, 1992).

None of the Box's M for the 13 variables measured in comparing the LPs and the NLPs were significant, indicating that the observed covariance matrices of the dependent variables are equal when comparing LPs and NLPs. However, Bartlett's Test of Sphericity

for each of these variables was highly significant going from an α of below .0005 to .004. This suggested that the residual covariance matrix for each of these variables was not proportional to an identity matrix. Mauchly's Test of Sphericity also showed that four subscales, Peer Relations, Reading, Total Nonacademic and Total Self had significant Mauchly's W. When comparing the two older groups, one Box's M was significant, all Bartlett's tests were significant and 3 Mauchly's W were significant also. When comparing the two younger groups, 2 Box's Ms, four of the 13 Bartlett's tests and none of Mauchly's Ws were significant. These mixed results can be confusing.

However, there is evidence that Mauchly's W may be the best measure to use. In a study examining eight different tests of sphericity, Cornell, Young, Seaman and Kirk (1992) found that the V test (based on Rao, 1947) was the best test of sphericity. However, Mauchly's W (Mauchly, 1940) came in a fairly close second when p/n ratio (the number of levels of the dependent variable over the number of subjects in a condition) was not approaching 1 (especially in cases when the p/n was 1/2 or smaller). As SPSS only performs Rao's V test for discriminant analysis and not for repeated measures, and as my p/n ratio is less than 1/2 in all but one case, I chose the Mauchly's W test of sphericity as the primary test of sphericity. There are also corrections for a lack of sphericity (in the form of epsilons) provided with the Mauchly's M. I conclude that while care must be taken as far as the assumption of sphericity is concerned, for the majority of cases, the assumption of sphericity has been met.

Transformations

I attempted to solve the problems with the assumptions by applying both the recommended power transformations that SPSS generates with its Spread-versus-Level plots and with other suggested and available transformations. None of these transformations helped fulfill the assumptions nor was the analysis of repeated measures helped by them. This is not surprising, however, considering the guidelines that Afifi and Clark (1996) put forth regarding transformations. They contend that when a scale has a natural zero (or when appropriate changes are made when there is no natural zero), if the maximum/minimum is less than two, or the standard deviation/mean is less than 0.25, then a transformation is likely to not be beneficial. In the case of my data there were only 4 scales, out of the 39, that met the criteria for all four small groups for all three administrations. These scales were Mathematics 1, General-school 1, General-school 2 and Total Academic 1.

Also, the recommended power transformations (generated by SPSS) were often vastly different for the three administrations of the same scale. Any transformation that could be beneficial to one administration of the given scale would most likely be detrimental to another administration of the same scale.

Thus, the problems with the assumptions could not be corrected using statistical means and subsequent analyses must proceed with caution and be viewed as estimations at best. This is especially true of any analyses pertaining to the YLPs, whose scores seemed

to be particularly problematic, especially at the second administration.

Repeated Measures

I chose to do the repeated measures analyses notwithstanding the difficulties with the assumptions, but with the understanding that the interpretation of the results would be interpreted with a good deal of caution. This was a fairly easy thing to do as there were virtually no significant results and those that were significant were difficult to interpret at best.

None of the between-subjects effects were significant for any of the scales for the three comparisons. Only one scale even approached significance, the Physical Appearance scale when comparing the OLPs and the ONLPs, with an F of 364.794 and an α of 0.067.

Only one interaction was significant in the within-subjects effects, the Physical Appearance by LPs vs NLPs. Because Mauchly's W for this comparison was not significant, I used the sphericity assumed interaction (without an epsilon adjustment). At $F=3.466$, the significance was $\alpha=.0040$ (significant at $p \leq 0.05$). The effect size for this interaction was 0.136. For the comparison of the two older groups there were two significant main effects in the between-subjects effects, Total Nonacademic and Total Self. With an F of 4.243, the Total Nonacademic scale had a significance of 0.037 for the Huynh-Feldt and a significance of 0.047 for the Greenhouse-Geisser epsilon adjustment for degrees of freedom. This scale had a significant Mauchly's W ($\alpha=.049$) and so the suggested epsilon adjustments were made to the degrees of freedom to compensate for the

nonsphericity. The Total Self needed no such adjustment and for an F of 3.720, it had a significance of 0.041.

Some authors argue that the tests which use adjusted degrees of freedom should not be used when other, equally efficient and powerful tests are available that do not require the assumption of sphericity (Jensen, 1987). These other tests include Pillai's Trace, Wilk's Lambda, Hotelling's Trace and Roy's Largest Root. In my study, all of these tests gave the same F values and significances. Again there were very few significant effects. The only significant interaction was in the comparison of the two older groups. The Physical Appearance interaction with the two older groups had a significant F ($\alpha=0.036$). The Physical Abilities main effect for the two older groups was also significant ($\alpha=.003$) as was the Physical Abilities main effect for the comparison of the Learning Partners and the Non Learning Partners ($\alpha=.049$).

The effect sizes of these significant results ranged from 0.136 for the interaction of the Physical Appearance scale by the Learning Partners versus the Non Learning Partners, and 0.278 and 0.253 for the main effects of Total Non Academic and the Total Self in the within-subjects effects. These are quite low. In the Multivariate Tests (MANOVA), the Physical Appearance by OLPs vs ONLPs interaction has an effect size of 0.486. The simple effect, Physical Ability, has an effect size of 0.680. Both of these are respectable moderate effect sizes.

Outliers removed

I redid the analyses once I removed the outliers (which I identified using the box plots). The results of the new analyses did not differ much from the analyses which included the outlier scores.

When inspected plot by plot, many box plots showed improvement with the removal of the outliers, especially with the box plots for the YLPs. However, the overall patterns did not change very much at all. There was still a marked difference with the YLPs, particularly at the second administration, with generally much smaller ranges than for other groups or for other times. The other difficulty, of course, was that these now slightly improved box plots were often for only three participants.

The Shapiro-Wilks test showed that there were more scales affected by non normality than before, and the statistic was not even calculated in 21 out of the 78 group by scale calculations when comparing the two younger groups. There were even more scales with skewness and kurtosis difficulties after the outliers were removed.

Levene's test of homogeneity again was significant for only one scale. When looking at just the variance differences, once the outliers were removed the variances were even greater than before. There would have been too large a variance difference in one of the comparisons of the LPs and the NLPs, seven of the comparisons of the OLPs and the ONLPs and eighteen of the comparisons of the two younger groups (one with a variance difference of 329.2 times).

Sphericity seemed to be improved as a result of the removal of the outliers: there were fewer significant Box's M and Bartlett's tests, but there were the same number of significant Mauchly's W, though not all on the same scales. However, for seven of the thirteen scores for the comparison of the younger participants, SPSS was unable to perform a Box's M.

It appears that removing the outliers did not do much to improve how well the assumptions were met. Little improved as far as results went as well. With the outliers gone, the results were virtually identical to those with all the scores intact.

Nonparametric Tests

Because of the difficulties I had with meeting the assumptions for the parametric repeated measures, I ran four nonparametric tests on the data in addition to the repeated measures: the Friedman test, Kendall's W, the Wilcoxon matched-pairs signed-ranks test and the Sign test. The Friedman test ranks the scores of the groups being compared and then tests to see if they are from the same population (SPSS, 1997). It is related to a one-way repeated measures (Howell, 1997). Kendall's W measures rater agreement. Each participant is considered a rater and a sum of ranks is calculated for each variable. The test statistic can range from 0 (no agreement) to 1 (complete agreement) (SPSS, 1997). The Wilcoxon matched-pairs signed-ranks test checks whether two related variables distributions of difference scores is symmetric around zero (Howell, 1997). The test gives varying weights to varying differences and the test statistic is based on the absolute values of the rank differences. The Sign test also checks for same distributions. In this

case, the test checks to see whether there are a significantly different number of positive or negative differences. This is the most assumption free test, but it is also the one with the least power (Howell, 1997).

The Friedman and the Kendall tests gave exactly the same statistics. For these two tests, the Physical Abilities scale was significant for the LPs and non significant for the NLPs. The Physical Appearance, Peer Relations, and Total Nonacademic scales were significant for the NLPs and not for the LPs. The General-self and the Total Nonacademic scales were significant for the ONLPs but not for the LPs. There were no other significant differences.

The Wilcox matched-pairs signed-ranks test showed the LPs had a significant difference between administrations one and two of the Physical Abilities scale, the NLPs had significant differences between administrations one and three, and two and three of the Physical Appearance scale and between administrations one and three, and two and three of the Total Nonacademic scale. The ONLPs showed significant differences between administrations one and two of the Physical Abilities scale, between all three pairs of administrations of the Physical Appearance scale, and between administrations one and three, and two and three of the Peer Relations scale, administrations one and three of the General-self scale and administrations one and three of the Total Nonacademic scale. The YLPs showed a significant difference between administrations one and two of the Physical Abilities scale.

The Sign test showed similar results. The LPs again had a significant difference between administrations one and two of the Physical Abilities scale; the NLPs had significant differences between administrations two and three of the Physical Appearance scale, administrations one and three of the Peer Relations scale and between administrations one and three of the Total Nonacademic scale. Other differences that were significant in the Wilcoxon Signed Ranks Tests were approaching significance on the Sign test. With the older students, the results were the same except that the Sign test did not show a significant difference between administrations one and two of the Physical Appearance scale. The difference between administrations one and two of the Physical Abilities scale for the YLPs was only approaching significance.

Thus it appears that for the most part, the nonparametric tests support the differences found in the parametric tests. The nonparametric tests, however, found more scales with significant differences (or simple effects), but this is to be expected as the power of the tests is lower than that of the parametric tests (May, Masson & Hunter, 1990).

Interaction Plots Analysis

I created interaction plots for each of the scales for each of the three comparisons. These were intended as part of the post-hoc procedures to follow the repeated measures analysis. I thought it would be useful to look at the patterns that the different groups took for all of the subscales, not just the one significant interaction.

An interesting thing occurred when I was studying these interaction plots of the estimated marginal means for my data. I was surprised to find an apparently large interaction in both the Mathematic and in the Math Academic scales when comparing the LPs and the NLPs, and when comparing the YLPs and the YNLPs, and an apparently smaller interaction in the Mathematics scale for the OLPs (for an example, see Figure 11). I

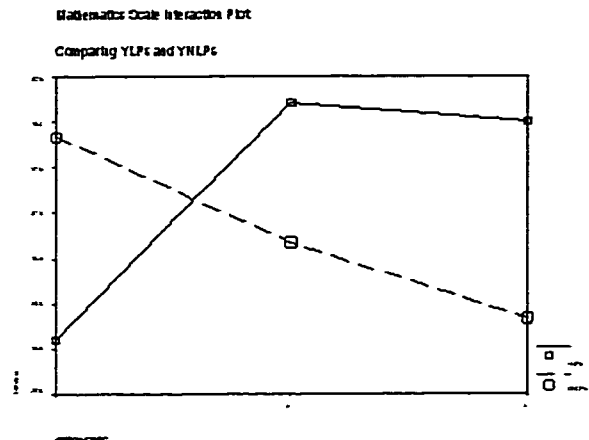


Figure 11 — Mathematic scale interaction plot for YLPs and YNLPs — unmodified.

was very excited but at the same time quite confused: these certainly did not look like the results that the numerical analysis gave. Then I looked closer and I understood. The differences between all three groups and all three administrations for the two scales were so slight that there really was no significant interaction. Once I converted all the

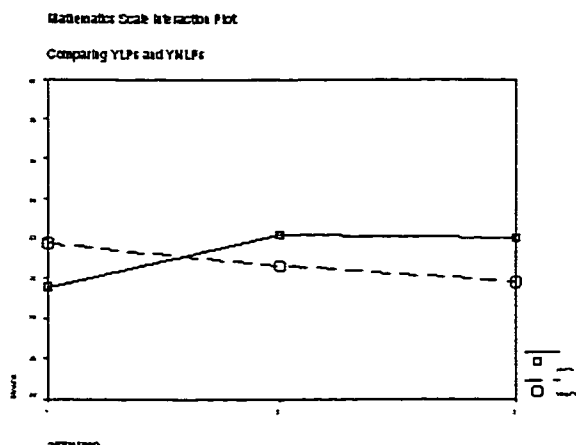


Figure 12— Mathematics scale interaction plot for YLPs and YNLPs — modified to a common scale.

interaction plots to a common scale, the interaction almost disappeared (for an example, see Figure 12). Even though these differences were very small, they did show trend, a continuing pattern of interaction for the mathematics related self-concepts.

There were other interactions, but many of

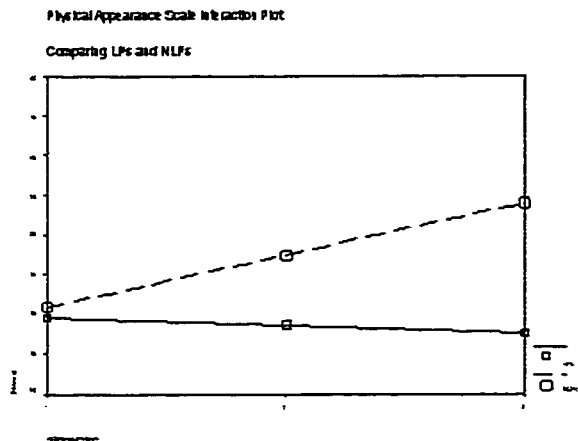


Figure 13 — Physical Appearance scale interaction plot for LPs and NLPs — unmodified.

these either were close to parallel or did not seem to follow a trend over all three administrations. For example, the scores of one group went up at administration two and down at administration three, while the scores of the other group went down at administration two and up at administration three. A notable exception to this were the plots for Physical

Appearance and Total Nonacademic for the comparison of the older groups of the LPs and the NLPs (for an example see Figure 13).

These interactions remained quite obvious even when the scales were converted to a common scale (for an example see Figure 14). It certainly makes sense that the Physical Appearance scale showed a stable interaction as

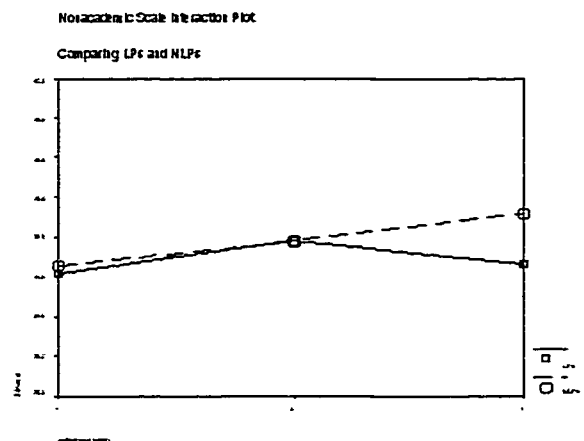


Figure 14 — Physical Appearance scale interaction plot for LPs and NLPs — modified to a common scale.

this was the only scale to show a significant statistical interaction for the two older groups in the multivariate tests and for the LPs and the NLPs in the within-subjects test of the repeated measures.

It is interesting that, at least to some extent, the patterns that emerged from the non

significant interaction plots served to answer the research questions posed in chapter 2 and were in line with the statement: "This brings into question measuring differences on a theoretically stable general self-concept or even an academic self-concept, when tutoring occurs in only one subject area. By measuring the self-concept related to the subject area being taught, it is possible that a more accurate portrayal of the effects of the CALP program on self-concept can be determined" (p. 22). The two quantitatively driven research questions also anticipated such a trend: **"did students in the CALP program have greater gains in subject specific self-concept than children in the control groups? (Q_1).** Further, **would gains in general self-concept be less differentiated between the learning partners and the control groups than the subject specific self-concepts? (Q_2)"** (p. 23). Obviously, because of the difficulties encountered with the assumption and, more importantly, because of the lack of any significance whatsoever of these interactions with that math related score, nothing can be said about them other than that these patterns may be interesting and may show an emerging trend.

To counteract the assumptions not being met, I created a series of interaction plots using the medians of the groups rather than the estimated marginal means (as in the plots discussed above). When comparing the LPs and the NLPs, interactions occurred with the several scales. The patterns were similar to those for the interaction plots for the means. Physical Appearance showed a marked interaction for the two older groups. The Mathematics scale seemed to show a relatively strong interaction for the LPs versus the

NLPs, and virtually no interaction for either the older or the younger group comparisons. While the Math Academic scales showed virtually no interaction for the median scores of the LPs and the NLPs, nor the older group comparison, there seems to be a definite trend for the younger groups comparison. Other scales showed definite trends towards interactions. In the LPs versus NLPs comparison, the Physical Appearance and Total Nonacademic scales showed such trends. When comparing the two older groups these trends emerged in the Physical Appearance, Parent Relations and Total Nonacademic, and for the two younger groups, the Parent Relations and Total Nonacademic scales.

Summary of Quantitative Results

It seems that the results are at best inconclusive, and more conservatively, entirely nonexistent. There were no significant results related to mathematics self-concept. I am unable to explain the only significant results, mainly with the Physical Appearance scale. The other minor significant results all seem to be from the same Nonacademic domain and go directly counter to the research questions which suggested that mathematics related self-concept would be more affected by this program than non mathematics related self-concepts. The evidence presented clearly shows that this is not so.

The only glimmer of evidence that any increase, however minor and insignificant, occurred in the mathematics related self-concepts of the Learning Partners (older, younger or as a whole) as opposed to the Non Learning Partners was in the interaction plots. Again, I emphasize that this evidence must be taken with great caution as it is not

significant. However, the trends did seem to emerge, on a consistent basis in these interaction plots and this may indicate that, if the limitations of this study were eradicated, perhaps these trends would have emerged as significant in the data. However, it is also very possible that this would not be the case.

Examining the Instrument

This section evaluates the SDQ-I and its accompanying manual. There were a number of discrepancies and difficulties that I found when I examined these materials. Many of these difficulties came to light only after the questionnaire had been administered and scored.

The Manual

When I began looking at self-concept, I came across the theoretical foundations underpinning the Self-Description Questionnaire. Shavelson, Hubner and Stanton (1976) identified seven critical features of self-concept: it is organized, multifaceted, hierarchical, stable, developmental, evaluative and differentiable. The multifaceted or multidimensional quality of the model's definition of self-concept was the critical issue. This model and the SDQ-I were tested and modified in a dynamic, reciprocal fashion. Research impacted both and the resulting modifications of one lead to modifications of the other. This interplay of theory and practical application made the SDQ-I a very attractive choice as an instrument to assess multidimensional self-concept and self-esteem.

The instrument was originally designed for children in grades 4 to 6 and the validation,

reliability data and norms provided in the manual are for children in those grades. Further research has shown that it is suitable for children as young as those in kindergarten, with proper modifications (Marsh, Croaven & Debus, 1991). While most sample populations have come from Australia, there have been studies conducted in England (Marsh & Smith, 1987; Smith and Marsh, 1985) and Canada (Byrne & Shavelson, 1986; Marsh, Byrne & Shavelson, 1988). While it may be dangerous to use any psychological measures indiscriminately between countries, different cultures and nationalities may have varying definitions of, structures of, attitudes about, and manifestations of self-concept. Some groups may value humility, thereby suppressing or concealing positive self-concept whereas others may value pride and self-promotion, thereby encouraging and expressing positive self-concept. Clearly the same instrument cannot be used in the same way with these different groups. While it may seem obvious which cultures and nationalities differ in which ways, in reality, we can not really know without research. Even within a smaller geographical area there may be vast differences.

Therefore, it is essential to know that an instrument has validity in the country or area in which it is being used. Even when the country of origin is an English speaking, commonwealth country like our own, it is important to make sure that the instrument can be used here with confidence. While the SDQ originates in Australia and the vast majority of validation and norming has been done there, there is evidence that the instrument is valid for use in Canada (Byrne & Shavelson, 1986; Marsh, Byrne & Shavelson, 1988).

The instrument has been validated by many studies; unfortunately, the overwhelming majority of these has involved the author and originator of the instrument, Herbert W. Marsh. The construct validity of the instrument is implied by the relationship of responses to various other predictors including demographics (sex, age, socioeconomic status) academic achievement and teacher ratings of achievement, inferred self-concept (by teachers and peers), self-attributions, other self-concept measures and programs to enhance self-concept (evaluated through experimental design).

There are some difficulties with this instrument and its accompanying manual. Sometimes instructions in the manual and in the scoring booklet are not entirely clear, or even contradictory in places.

For example, the manual example shows certain control scores being rounded off to the nearest whole, but there is nothing in the written instructions regarding rounding. Instructions say to enter the mean responses for missing items, if there are fewer than four missing items in a sub scale (of 8 items). They only caution to interpret with caution if four or more items are missing. I would submit that caution be used in interpreting any instance where more than one or two items are missed. I would exercise caution at a point much lower than a 50% omission rate. I also have great difficulty with substituting the mean score for up to three missed items. I would be more inclined to agree with an interpolation method of filling in missed items, especially for more than one item. If a person scored three 1s and two 2s on the five item she answered of, for example, the

Physical Abilities scale and missed items 8, 32 and 48, (the three items with the lowest means), the scale subscore would be 18.37 (with an item mean of 2.29625). But if that person were to continue answering with her typical responses for this scale (with a mean of 1.4), the subscale score would be 11.2. Using the item means instead of interpolation, would cause an increase of over 60% in the score. There is no reason to believe that this person would answer according to the mean for the three missed items, when she was not answering anywhere near the mean for the 5 items she did complete. If all 8 of her answers had been consistent with the mean her scale score would be 32.65. Clearly, no matter what method you use to calculate the missing items, this person's score would not approximate the mean.

Control score 5, which measures the positivity bias, is inconsistent with the remainder of the control scores. Scores on this scale are considered questionable if the score falls either in the top 5% or in the bottom 5% of scores of the normative sample. Therefore, for this score, 10% of the scores of the normative sample were considered questionable. In all the other control scales, there is only a one ended threshold, and only scores above or below the 5% of the normative sample are considered to suggest inappropriate answering. While this difference may be legitimate, there is no mention nor explanation of why 10% of the scores on this scale are considered inappropriate and only 5% are considered such elsewhere.

There were several instances throughout the manual where the author was quite vague

or confusing. For example, “most SDQ-I research has been based on factor scores, although some studies have used raw scale scores” (p. 24). It is possible that Marsh differentiates between these two terms somewhere in the manual, but I was unable to find it, even after going through the manual several times. The manual and the scoring sheets refer to raw scores. These are scale totals for the first-order scales and scale totals divided by the number of scales that accumulate into the higher-order scales. All these scales have possible scores of 8 to 40. I imagine that the author was referring to “factor loadings” discussed further in the book, but never defined as factor scores. There were numerous such small instances where it was not entirely clear what the author meant.

I also question some of the author’s assumptions. For example, Marsh says that in the research children have been assured anonymity and thus have little reason to respond other than honestly. However, I am not certain that at this age children would fully understand the concept of anonymity, and even if they do fully understand it, there is always the possibility that students (or anyone) might answer in a way that makes them look good.

My greatest single difficulty with the SDQ-I may be due to an oversight, but it is certainly a very great oversight and one that was made more than once. In chapter 4 of the manual, Marsh shows that through factor analysis, the best description of the hierarchical structure of the scales of the SDQ-I-I was found to include three second-order factors, Total Nonacademic, Reading Academic and Mathematics Academic. He goes into quite a bit of detail explaining how this is the best model for the self-concept. Elsewhere in the

manual, particularly chapter 3 where he addresses the scales of the SDQ-I, and in the scoring sheets, Marsh fails to incorporate this “best” model and uses a clearly inferior (according to chapter 4) model with only two second-order scales. In fact, this model was not even second best, but third of four possible models, the fourth of which did not include any mid-order scales, only first-order and a general score.

As mentioned above, I chose to work with the scales of both these models. Had there been significant results, the emphasis of the analysis would have been on the best fit model described in chapter 4. I felt uncomfortable working with the clearly inferior model, but I also felt uncomfortable with throwing it out completely as the scoring sheets did only take this inferior model into account. I felt secure in using the Reading Academic and Math Academic scores as they were simply recombinations of the first-order scales, just as the Total Academic scale (of the inferior model).

In general, most of the difficulties with the manual were fairly easily overcome or did not really have an impact on my understanding. They do show however, that even within the manual there are inconsistencies, some of them quite major (as in which model is used for the scoring and analysis of the instrument). Such a pattern of inconsistencies and ambiguities may impact the validity of the instrument as a whole. This may particularly be a concern as the overwhelming majority of the work on the instrument has been conducted by or with the author.

Questionnaire Responses

During the administration of the questionnaire to the younger students, I noticed that some of the students did not seem to be answering the items with regard to the content of the item. It was quite easy for me to spot these patterns because I was administering the instrument to the students one at a time using an interview method. When I noticed that the student did not seem to be responding to actual content, I noted the items in question. At the end of the administration, I went back and read these items again, noting the new response. For analysis, I only used the original response since there was no theoretical or empirical basis for my decision to go back to items in this fashion.

This is exemplified by a statement in my field notes, *Milicent did her questionnaire this afternoon. I found that she may have had difficulty with realizing what her answers are. She sometimes would say **not true**, I would then ask her very not true or mostly not true. She would then say mostly **true**. I would repeat her last answer to make sure that was what she meant. Eyeballing her answers she seemed to move from a diverse answering pattern to answers that all fell either in true or in the false (very true and very false). This transformation seemed to occur gradually from the beginning to the end. I also believe that there is quite a bit of variability between similar comments.* Such patterns were particularly apparent during the second administration, which is consistent with the results of the quantitative analysis of the scores.

Some of these children had quite substantial discrepancies between the first and

second response to an item within a few minutes. This leads me to ask why such discrepancies occurred and whether similar response discrepancies would occur for some of the older students if they were asked the same questions more than once, perhaps most especially those near the end of the questionnaire.

While Marsh, Craven and Debus (1991) found that for very young children the beginning items on the questionnaire posed some difficulty (what the authors call the practice effect), I found the opposite, that the items at the end *seemed* to give some students difficulty with responding with relevance to the item content.

The Questionnaire and the Scales

I also have some questions about the composition and placement of items of some of the scales. There seem to be patterns where I feel there should not be patterns and inconsistencies where I feel there should be more consistency.

While the items for seven of the eight first-order scales seem to be distributed throughout the questionnaire, the items for the General-self scale definitely cluster around the end of the instrument. Five of the items are in the last 9 items of the questionnaire. There seem to be patterns in the way that the items are laid out in the questionnaire. For example, in five instances items from the Physical Appearance, General-school and Physical Ability are presented in that order consecutively. In four other cases, two of those three are presented in order and consecutively. Only one question from the General-school scale is presented out of this pattern. This of course could conceivably cause

answer bias.

The items for the scales were not created equally. For the three academic factors, the items for each scale have parallel wording. There was no such parallel wording for the other scales. To be fair, it would have been difficult to have parallel wording for these other scales, but there did not even seem to be an attempt to have the same types of items. For the three academic scales, half the items referred to perceived fact or ability and half refer specifically to affect. The other scales referred almost exclusively to perceived fact as opposed to affect. It seems that the scales are not comparable, though the author does not claim that they are. It also seems that the scales could have been made more comparable rather easily.

The Physical Abilities scale refers to rather specific sports, games or athletics. Perhaps such specificity does not tap into enough of a range of physical abilities to warrant the use of that name for the scale. Similarly, the Peer Relations refers to a very limited scope of how popular the child is, but does not tap into quality of relationships nor into how the child feels about them.

Thus, I have some concerns regarding the scales. Not enough information is given in the manual regarding the location of items and the makeup of the scales to abate these concerns. I do not know what impact, if any, these concerns have on the validity of this instrument with my sample, or any other, or on the results, or rather lack of results of this study.

I am not trying to imply that the lack of results in this study is a consequence of any concerns regarding the instrument, only that results may have been affected by the composition of the questionnaire and its scales. Whether this would have resulted in more significant results or in even less significant results is impossible to tell.

Reliability Analysis

The assumptions related to repeated measures were often not met, in some cases the departure seemed quite dramatic, as in the case of the YLPs at administration two. Therefore, I decided that more investigation of the instrument was necessary. As stated earlier, in chapter 2, I had some difficulties with the instrument even during the administration phases and I questioned the validity of parts of it and of the instrument as a whole. Some of these questions were about the instrument in general, but some were about my sample's responses to the instrument in particular.

To begin to investigate these questions, I conducted a reliability analysis on each item for each of the thirteen scales to help gauge how well the items in each scales related to the overall scale and to determine the internal consistency for each scale for the participants in my study. Using this reliability analysis I discovered some interesting problems with some items, scales and groups. I determined which items were least related to each scale, thereby lowering the scale alpha. Then I redid my analysis (both the reliability analysis and the repeated measures analysis) after removing the six most offensive items. I found that while removing these items did cause the promised increases in scale alpha for those cases

where the items were causing the most difficulty, in almost as many cases, the scale alpha actually went down as a result of removing these items. In the reanalysis, 16 scales of the 39 scales had an improved alpha and 17 had an alpha that decreased. Overall, for the 39 scales, the net increase in alpha totalled .2452, or an average increase of only .0063. This very small increase did not seem to justify changing the original questionnaire.

The reanalysis of the repeated measures caused no overall useful trends, some of the statistics that assess the violation of assumptions were improved, others were worsened and still others were unchanged. Overall, the changes in the scales seemed to have a negative impact on the repeated measures analysis.

The reliability analysis gave some valuable information in addition to the reanalysis. The Cronbach's alphas for all the participants for each of the scales ranged from .5303 to .9283. The reliability tended to increase over the three administrations, in some cases by as much as .2664 for the Physical Abilities scale. The Cronbach's alphas for the second and third order scales tended to be higher than for the first-order scales, with one higher-order scale having a Cronbach's alpha of .7215 while all the rest ranged from .8182 to .9283, with six of these 15 scales higher than .9000. The first-order scales had Cronbach's alphas of .5303 to .9199. Two of these scales had Cronbach's alphas in the .5000 to .5999 range, one in the .6000 to .6999 range, seven in the .7000 to .7999 range, eleven in the .8000 to .8999 range and the remaining three scales in the .9000 to .9999 range. From this analysis it seemed that for this whole group of participants, the scale Cronbach's

alphas were quite respectable and the scales were reasonably internally reliable.

The picture changed quite drastically when I looked at the reliability analysis for the individual groups. While there are too few participants in each group to make this analysis really meaningful, it was an interesting exercise that points to a rather severe problem. Some of the Cronbach's alphas for this analysis were extremely low, as low as -.24416. Not surprisingly, the greatest internal unreliability stems from the two younger groups. Of the seventeen Cronbach's alphas below .4000, sixteen were for the younger groups, 11 of which were for the YLPs.

In comparing the LPs and the NLPs the Cronbach's alphas for the scales were quite similar. In many cases, even comparing the older and the younger groups yielded similar Cronbach's alphas.

Upon closer examination, Cronbach's alpha revealed some interesting observations. The variance of the scale reliability scores (Cronbach's alpha) for the younger students was almost six times as great as the variance of Cronbach's alphas for the older students. For the younger students as a whole, the Cronbach's alphas had a range of .8600 (from .0744 to .9344), while for the older students, the range was much smaller, only .3455 (from .6106 to .9561). The means were .8522 for the older children and .6829 for the younger.

Further, the younger students seemed to answer in a particularly unreliable fashion (as compared to the older students) during the second administration (see Figure 15). While the range was much greater for administration one (.8265 as compared to .5296 for

administration two and .5592 for administration three), this is mainly due to an outlier

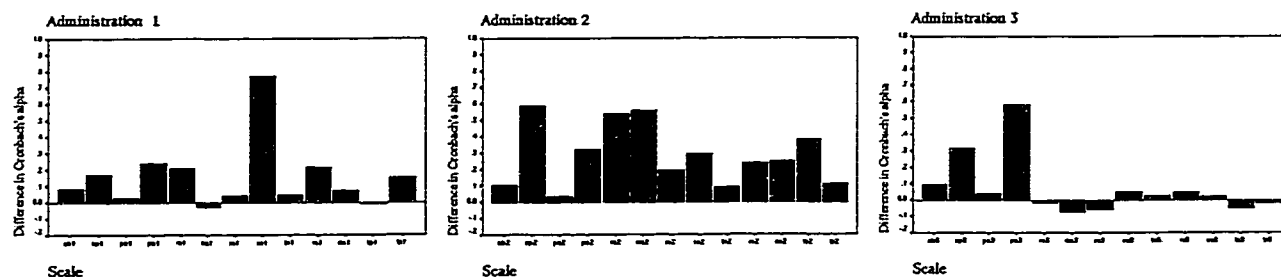


Figure 15 — Cronbach's alpha score differences between older and younger participants for each scale of administrations one, two, and three.

during administration 1. The first General-self scale (se1) had a Cronbach's alpha of only .0744 for the younger. For the first administration, the 25th percentile for the younger students was at .5854, while for the second administration it was at .3493, and for the third administration it was .7688. The differences between the older and the younger students' Cronbach's alpha for each of the scales during administration two was much higher than during administrations one or three. These differences are visually represented in Figure 16. The positive bars indicate that the older students had the higher Cronbach's alpha, while the negative bars indicate the younger students had the higher Cronbach's alpha.

There also seems to be a trend toward improved Cronbach's alphas over time. For all participants combined the mean Cronbach's alpha for administration one was .7941, for administration two it was .8073 and for administration three it was .8657. There were some segments of all the participants that did dip during administration two, most notably the

younger participants and particularly the YLPs. Also the ONLPs had a higher Cronbach's

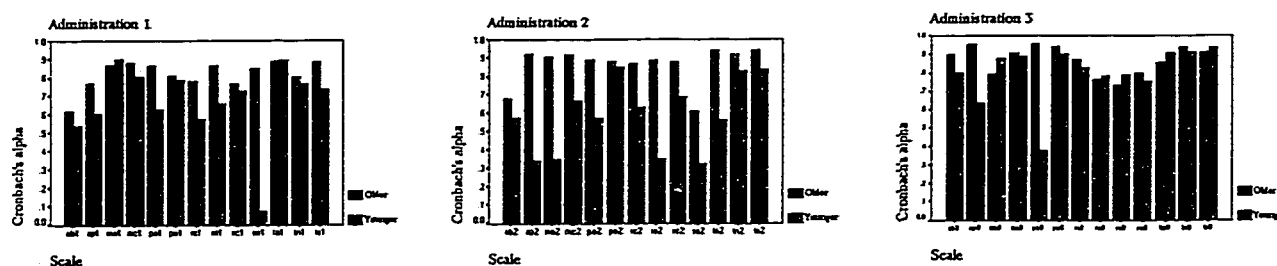


Figure 16 — Comparing scale Cronbach's alpha for older and younger participants for each scale of administrations one, two and three.

alpha in administration two than in administration three. The general trend however, seemed to increase over time.

I have no insights as to why this should be so. While I would expect the younger students to have more problematic distributions in general, I am not sure why only one of the groups, the participants, had such a difficulty at the second administration.

Chapter 6 — Qualitative Findings

In this chapter, I present and discuss the qualitative findings based on document analysis, observation and interviews. The qualitative aspects of this study were intended to help understand, expand or refute the quantitative findings. The hope was that the blend of qualitative and quantitative findings would possibly help resolve some of the discrepancies found in the literature regarding the effect of cross age learning on self-concept. Unfortunately, this did not occur. The quantitative findings were very inconclusive and thus the qualitative findings did not really help to understand, expand or refute them. Rather, the qualitative findings showed evidence of positive trends and interesting thinking patterns in the participants.

Even so, the qualitative findings were not as rich as they might have been. There were many factors which contributed to the shortcomings in terms of gathering and analysing the qualitative portions of my data. These include the complexity of the concept of self-concept, an insufficient theoretical sensitivity, an inappropriate use of observation, and an inadequate interplay of collection and analysis while in the field, “what you can’t find in your data becomes one of the limitations of your study. That is, either you didn’t collect enough information, go to the right places, persons, and so forth” (Strauss and Corbin, 1990, p.112).

The complexity of the internal process that is self-concept makes it a difficult construct to infer and to explore. Because it is a subtle, internal process, indirect approaches are

probably not appropriate ways of exploring self-concept. Even directly questioning people, especially young people, about their self-concept may leave out much in terms of understanding and even accessing a person's true self-concept. No matter what processes are used to explore self-concept, it must be inferred, either by the participant, by those close to the participant or by a compilation of information by an outsider. Because of this need for inference, data can never approach the reality of self-concept, even a person's own reality of his or her self-concept, because of the gap between inference and true self-concept.

Perhaps the greatest difficulty in this aspect of my research was a basic mistaken sense of comfort with my subject area or what Glaser (1978) and Strauss and Corbin (1990) refer to as theoretical sensitivity. While I was quite versed in the literature and meanings behind cross age learning and the nature of self-concept, I was not well enough versed in the behavioural indices of self-concept; I did not have a sufficiently thorough understanding of it and as a result, I was unable to identify some of the subtle verbal and nonverbal expressions related to self-concept that may have been expressed by the participants in the study. I did not realize, at the time, just how inappropriate observation was for understanding the self-concept, especially in such a restricted setting.

Because of my lack of extensive experience in the field I also had a certain professional theoretical insensitivity to what would happen in the field and how. I thought I had a fairly clear picture of this from reading and hearing about other people's experiences, but I was

not quite prepared for what actual fieldwork was like.

Emotionally, I experienced a sense of let-down once I was in the field. Though the participants, students, staff and environment were wonderful, after an unexpectedly difficult struggle to get to the point of entering the field and not having expected such a struggle, I had used much of my energy, and felt quite drained once there. Not only during, but even for a very long time after my field experience, I found it impossible to become engaged in the process. I felt my reserves dwindling and was unable to maintain any momentum.

This almost inevitably became part of a vicious circle; the more detached I felt, the less my data showed anything of value and the less my data showed of value the more discouraged I became and the more detached I became. In short, I failed at gleaning the theoretical sensitivity available to me through the analytic process itself. I did not sufficiently benefit from the “increasing sensitivity to concepts, their meanings, and relationships [which is] why it is so important to interweave data selection with analysis” (Strauss and Corbin, 1990, p.43). I did not take full advantage of this interplay between data collection and analysis, thereby limiting the benefits I could have had of allowing myself to better focus on relevant data in subsequent field sessions and the further verification of the emerging hypotheses that developed over the course of the study.

I also suspect that perhaps observation was not an appropriate technique to use in this case because of the limited and fairly contrived situation. Self-concept is an internal

process that may not be readily available for observation by others in a fairly short-term situation, and especially not in such a limited context as this was. Since I was looking at cross-age learning, the setting was natural. However, it was still limited and fairly contrived in comparison to other, regular settings in the participants' lives. It is entirely possible, even probable, that it may not have been an appropriate setting in which to examine self-concept. Perhaps the kinds of interactions and the roles that the students were given minimized the expression of even those self-concept related behaviours that can be apparent to observers.

These limitations, combined with the limitations imposed by the setting, resulted in limited qualitative data directly pertinent to the self-concept. The mini case studies that I developed from the transcripts and the field notes yielded only modest evidence regarding self-concept. There was also evidence of descriptions of the participants, and how they reacted or changed with respect to the program. There were descriptions of how well or how poorly some participants did, how enthusiastic or apathetic they were, or even of how much they thought they did or did not learn from the program. Of the information that came out about self-concept or self-esteem, very little came from the observation or from the document analysis. The most useful qualitative source of data regarding self-concept came in the form of the interviews, particularly the more formal ones.

The following sections describe and discuss the qualitative data collected. First, I look at the document analysis. Then, I move on to observation and interviews. The interviews

took three forms: collaboration groups, where data collection was a secondary goal to helping OLPs with issues pertaining to the learning sessions; informal interviews, asking participants and teachers about the program in an unplanned and informal manner; and four fairly formal interviews with an OLP, a YLP, and each of their teachers. These elements combined to give an overall picture, a brief description of the participants and a more detailed look at the two individually interviewed participants.

Document Analysis

Most comments on the learning logs (see Appendix C) were generic and only somewhat helpful in examining self-concept. Both the older and the younger learning partners had some difficulty completing the learning log at the end of each day's work. This log included basic information, the date, the names of the partners and the activities planned and done. There were also two four-point rating scales which showed happy and sad faces. Each child (older and younger) rated the learning session from two happy faces to one happy face to one sad face to two sad faces. Then they wrote why they gave the rating that they did. At the beginning of the program, I explained that two happy faces meant that the learning session went really well, one happy face meant that it went okay, one sad face meant that it was not very good, and two sad faces meant that the learning session went very badly. Almost all ratings were either the one or two happy faces, with a few single sad faces. There were no ratings of two sad faces. Overall, it seems that the partners felt that learning sessions were a positive experience.

The older learning partners commented that they liked it when the younger partners listened, understood, tried hard and worked diligently. Sometimes their comments implied pride, “[Sally] worked really well and is done her hundreds chart in one day⁷” (Kelly) and “Cassie learned all her 10s and now she knows them frontwards and backwards” (Hailey). Both the older and the younger partners also liked it when they accomplished something “We got things done” (Sandra).

Other times the older partners seemed to just want to get on with it, they said it went really well because “it was fast” (Kelly). They did not like it when they just supervised (and did not get fully involved), and they did not like when their younger partners were uncooperative, apathetic, not paying attention, not listening or working, or were fooling around.

In general, and not surprisingly, the older learning partners gave a rating of two happy faces when the younger learning partners put in an effort and succeeded. They gave lower ratings when one of these two factors was missing, that is, when the younger learning partner either did not try or did not succeed.

Usual comments from (or about) the younger students included that they had fun, they learned, and that it was easy. The YLPs liked it when a learning session was fun or “cool,” and when their older partners were helping them. Five of the six younger learning partners commented that they liked it when they learned, one commented that he did not

⁷Text in italics is excerpted from field notes and transcripts. Quotation marks indicate a direct quote; no quotation marks indicate comments made by me in my field notes.

like having to work and to learn. One younger partner commented that "*work is fun*" (Dan). Near the end of the program, another of the younger partners commented that he did not like that "*it was the end*" (Seth).

Outside elements also played a part in whether they had a good learning session or not. Very often, Dan commented that he was tired, and on those days he circled only one happy face or even a one sad face once. All other days he circled two happy faces. Also, many times the younger partners' comments and answers to "I really liked" and "I really didn't like" were that they did not know. One participant, Seth, was particularly confused. This example may be somewhat extreme, but it is not atypical:

Seth said that he liked today because he learned a lot. When I asked him what he learned he said "Sounding out." "Sounding out what?" "The word." "What word." "The pattern." Then I asked him again what he learned and he looked at Steven for the answer. He could not come up with a further answer. He hummed and hahhed, looked at Steven, looked around the room. Finally he said, "I learned a lot because I am good at math." I still asked him what he had learned. Finally he came out with that he didn't know what he learned. With further questioning he said that he liked it because he had fun (March 27).

While most learning logs were as brief as possible, three older partners, all girls, wrote quite a bit, at least in the first few sessions. This quickly dwindled for two of them, but

Hailey kept on with fairly comprehensive logs (two to four full lines) throughout while most older and younger partners wrote less than one line each. For example, on February 23, Kelly wrote that the session went well (one happy face) because *“she [Sally] was doing very good in the adding sheet. She needed the figure chart but when she got it she did the sheet a lot faster. Next time I think I will ask her if she could try to go without the figure sheet.”* Later her responses were much shorter, usually of the sort *“Sally worked well”* and *“It was fun.”* Hailey, on the other hand, started off with statements like *“Cassie had a little trouble with adding and counting. She would leave out a number.”* (February 23) and while some of her comments did become shorter over time, they were still longer than any other partner’s: *“Cassie is really trying to tell the difference between odd and even numbers. She has made big progress”* (March 23) and *“Cassie finished hundreds chart hunt and learned about columns”* (June 8). Compare this to typical comments from the other older partners: *“He (Dan) worked very good!!”* (Aaron, February 23); *“I like helping”* (Steven, March 27) and *“He (Timothy) did really, really, really, really, really good”* (Judy, April 20).

The document analysis of the learning logs allowed some insight into what it was that the learning partners valued and gained from the learning sessions. It also showed some of the behaviours that were manifested throughout the program. In general, it seems, that the older participants gained the most when partners were cooperative and successful, while the younger learning partners were less cohesive or at least less articulate.

Observation and Interviews

I combined the field notes on the observations with the transcripts from meetings and interviews to create mini case studies for each of the LPs. First I will begin the discussion with general comments about the OLPs and the YLPs and then I will discuss individual students who were not individually interviewed. The next section will deal with participants who were individually interviewed.

The Older Learning Partners and the Younger Learning Partners

The primary difficulty that most of the older partners had with the CALP was when the younger partners were uncooperative. All the OLPs made remarks at one point or another about not liking when their partners did not pay attention: *"I think they should pay attention because I don't like it when Sally doesn't pay attention"* (Kelly, March 30); *"I want her to like sit up and pay attention, cause she'll just sit there . . ."* (Kelly, May 1); and, *Sally seems uncooperative . . . she is being excessively slow today. I am bored, Kelly looks bored and Sally looks bored* (My notes, May 29). I asked Aaron: *"[There is] nothing you don't like about [the program]?"* *"Just when the kid fools around"* (March 30); and Judy: *"What didn't you like Judy?"* *"Okay, because he's usually always daydreaming. Like looking around and . . ."* (March 30). When I asked the group what made them feel the worst about math and about the program, they replied: *"When Seth doesn't pay attention and, like she said, I don't understand how to do it"* (Steven, June 8). *"You know what makes me feel bad about teaching? . . . Um, like when Timothy is like [looks around very*

deliberately].” I prompted for clarification, “*When he is not paying attention?*” Judy replied, “*Yah. Like looking all around the place*” (June 8). When I asked what they would change about their partners if they could, they gave responses in keeping with their complaints: “*She wouldn't . . . If I was teaching her one thing, she wouldn't start talking about something else*” (Hailey, June 8), “*Timothy would pay attention*” (Judy, June 8); “*I don't want her to just sit there*” (Kelly, June 8); “*Seth wouldn't just fool around and be so hyper going, ‘Bloooooop.’*” (Steven, June 8); “*Well, Molly would tell me if she didn't get it. Like, cause, like I'll ask her . . . And she goes, like, like I'll say ‘Did you understand?’ and she won't say a thing and then I'll do it over with her and she'll, ‘Ooh, now I get it!’ Because she hasn't told me in the first place.*” (Sandra) (June 8), “*It makes me feel frustrated and depressed because we spend all that time trying to learn something and then she's not doing too well on it*” (Hailey, May 1).

On the other hand, the thing they liked most seemed to be when their partner paid attention. When I asked Steven what part of the program made him feel best about himself, he replied: “*When actually Seth really paid attention*” (June 8) and Kelly said, “*I like it because it went fast and she paid attention the whole time*” (June 8). These comments echo those of the document analysis.

Another commonality among the Older Learning Partners was that they generally felt that while their younger partners learned from the experience, they themselves did not and were not supposed to — that it was their job strictly to teach and not to learn from the

experience. For example, Sandra remarked, *"Well, I felt, like, um, there's nothing really I learnt about, I know my pluses, I know my I times, know my take aways, I know my odd and evens, I know everything that they did."* Then I asked, *"So that was the worst part. That you didn't learn anything?"* and she responded with, *"Yah,"* and Judy agreed, *"Yah. Exactly"* (June 8). This was a particularly noteworthy statement by Sandra as I noted that Sandra is fairly low academically, *even her very basic addition is very poor* (March 20); *that is, Molly is relatively high and Sandra is quite low. Put those two factors together and they are not too terribly distant academically* (March 27). I also noted that Sandra was not always as aware of things as her partner, *Molly seems to be more aware of what is going on than Sandra is sometimes. Though I believe that Sandra is helping Molly, I am sure that if Molly were not as vigilant as she is, both would miss out on quite a bit in the learning sessions* (April 20) and *Sandra was to continue doing the Rules Rule activity but she forgot how to do it. Molly remembered* (March 20). Aaron also thought that he had not learned and was not supposed to learn, *Aaron kept protesting that he is not the one that is here to learn. He wanted to know why I was asking him questions, implying that I should only be asking Dan questions. He obviously sees his role as that of teacher and not learner in this activity* (April 6). *They [all the OLPs] seem to think that they are giving their learning partners a lot but do not see themselves as gaining much* (April 6). All but one of the Younger Learning Partners shared this attitude: *most of them [YLPs] thought that they themselves were all learning. I think that only Dan thought that the OLP's were learning.*

(April 6).

At the end of the program, Sandra did admit to having been a learner in the course of the CALP. While she did not express this thought before, during her interview she said that she thought that at least part of her role had been learner: *"A little bit of both [teacher and learner]. . . . Because I am teaching Molly something and I might be learning something and don't even know I am learning. Like I know I am learning, but I don't realize that I'm learning it"* (June 12). She claimed that she did learn quite a bit from the program, *"I knew I was going to learn something and Molly was going to learn something. . . . Um, things that my teacher would, like my teacher does stuff like . . . in class sometimes. And you know it, then you don't know it. Like you think you know it and then you come here and you realize that you didn't know it."* I probed, *"So it was sort of like a review?"* and she responded with, *"Yah"* (June 8). While she was obviously uncertain about this, she thought that she might have learned something about math. In response to the question, *What sorts of things do you think you were learning?*, Sandra said, *"Math I guess. I don't know"* (June 12). The explanation that followed did not seem to clarify things any: *"Well, I didn't know my evens and odds. But then, when we were finished, um, well we, ah, were just learning that like when I was . . . When I came down here when we were learning even and odd, we were doing it in my class too. And I um, did. I was lucky because I didn't really know before and if we did that like, even and odd, I wouldn't really know. And before we did even and odd, so I knew what to . . . what it meant"* (June 12).

Sandra also mentioned other benefits. These included getting out of class as well as a greater sensitivity to her teachers when students are not listening and she also felt more comfortable raising her hand to ask questions.

All the OLPs felt that their YLPs were learning. When I asked if they thought that their partners were learning, Kelly mumbled an *“Umhum,”* Steven said *“Yes, very good”* (March 30). Their explanations of how they knew that their partners were learning were not always clear, and I am not convinced that the OLPs really had distinct reasons for their stated beliefs: *“Um, because he is like trying a bit harder and he sort of knows them [the math items] a bit better”* (Judy) and Sandra tried to explain with, *“Well, Molly, now she knows, like when I ask her. Like, she thinks they're hard, like when you say, ‘six take away six,’ she'll start to use those little counter things. And I showed her, like, ‘if you have six take away six how many are there?’ And then she said, ‘one,’ and I said, ‘Well, can you try that again.’ And then she used the counters and then she got it. And then I asked her, um, ‘Seven take away seven’ and she does know it now”* (March 30) and Aaron felt that his partner was faster on the hundreds chart activity.

The OLPs also felt that their younger partners were more confident, *“I think they're more confident in theirselves . . . They know that they can do something easier”* (March 30). Steven and Kelly also thought that their learning partners were more confident. Aaron and Judy did not and Hailey was absent when we discussed this topic. The Older Learning Partners also felt that they were more confident as a result of the program: *“Well, now, I*

was like Kelly. Like when Kelly wouldn't put up her hand, I wouldn't either. Because I would think the class would hear me. Or something like that. But now I can ask" (Sandra, March 30).

The teacher of three of the grade 5 students commented about all three children in her class who were participating, *"Yes, I think it's helping their perceptions of themselves. I think they feel a little special"* (March 23). At the end of the program when commenting about one of these students she felt that the program had probably helped her, *"She is more at ease with herself"* (June 8), though she could not be certain that the change was a result of the program.

Most of the YLPs felt that the program helped them to learn: *"I like it. It's . . . Well, it helps me learn. . . . Because my partner Sandra is well . . . Because Sandra helps me with math"* (Molly, April 6); *"[The program is] good. Because it helps you learn"* (Dan, April 6); *"I really like working with Hailey because we're doing lots of math together and I like doing that. And I learn lots of things from math and it learns how to count for people who don't know how to count"* (Cassie, April 6); and *"Um, I really like being here because I learn lots more math games and um math things"* (Cassie, June 1). Even teachers could see the improvement: *"I got the sense that she did feel better and I could see that she did learn because she came to the math class with more knowledge"* (Molly's teacher, June 12).

All the younger learning partners said that they liked math better than before they started working with their partners (April 6). They had fun and they liked working with

their older learning partners.

The younger learning partners' teachers saw some changes in what they called confidence: "*The confidence is coming. He is really keen to get going. He seems to be a little more confident and 'with-it'*" (Seth's teacher, March 23), "*I think Molly's confidence may have increased a little bit*" (Molly's teacher, March 23), and "*I guess self-confidence wise, I know he enjoys going*" (Joe's teacher, March 23).

It seems that most of the younger learning partners agreed that they had become better people as a result of CALP. They seemed to think that being smarter made them "better," but they were generally unable to articulate why. Usually, the closest that they came to answering the question was to say, "Because I pass more school subjects" (April 6). It would seem that the YLPs equate being a better person with getting better grades. Perhaps they had an equivalence of "being smart" or learning more and being a better person so ingrained into them that they never questioned it. If this is so — how could children who do not feel smart or who feel they are not learning enough feel that they are "good" or worthy? How could such a child who has this equivalence ingrained ever have a healthy self-concept? What kind of impact must this have on the self-concept?

Do we adults continue to make these connections? I think we do and furthermore, we also equate "better" with appearance, status, wealth and other indicators that do not, on their own, make a complete person. If a person equates any single aspect with "better" than anyone with less of this single aspect seems to have their self-concept in jeopardy.

Individual Students

Each of the Older Learning Partners and the Younger Learning Partners is addressed individually in this section. Two of the students, those who participated in the individual interviews will be discussed in the following section.

Steven

I was concerned about Steven from the start, but he turned out to be more capable than I had anticipated. Steven was a somewhat confused and forgetful boy (March 16) and, as such, seemed to be the brunt of a lot of derogatory comments and unfriendly attention. This was obvious to me in meetings, *I noticed particularly acutely the dislike and disrespect that they all had toward Steven* (February 9), but has been apparent at the school for at least several years. One of Steven's former teachers commented that though Steven is very helpful and friendly, he has generally been disliked by the other children. Aaron said that he and Steven were best friends and he usually treated Steven well, *Aaron told me that Steven is his best friend, that he really likes Steven, and that they spend a lot of time together outside of school. I told him that I was still pleased at how nicely he treated Steven* (February 23). But even Aaron publically made fun of Steven on occasion, *Aaron turned his eyes up and began laughing and when I asked him if he had ever forgotten something that he had meant to say, he said he could not remember. The rest of the group also was laughing. Another time Steven said something mildly inappropriate within the training session and Aaron began laughing* (February 6). This unconcealed dislike by most students

must have had an effect on Steven.

It is almost as if Steven deliberately wanted to embarrass himself. He asked a question. When Sandra answered correctly, he said, "*Good*," but then turned to me and asked "*Is that right?*" (February 9). The question had been simple enough, but he was not sure of the answer. He also volunteered to write his answer on the piece of chart paper, but his writing was terrible and he almost pointed it out to the class (February 9). He often made inappropriate comments, gestures and noises. These were usually not rude, only inappropriate and often, apparently annoying. When he did this, it often solicited the laughter of those around him. Steven generally laughed right along with them: *Once again I noticed how Steven, even when obviously insulted by the others, especially Ryan, reacted by laughing with the group, often more uproariously than the others* (February 9).

I suspect that this is because Steven felt that he had no other choice. Perhaps he reasoned that if he laughed with the others, he could pretend or even convince himself (or even the others) that they were not laughing at him so much as with him. Steven seemed to be an alienated little boy, who sought to counteract his alienation through humour, or at least his understanding of humour.

I was originally worried about Steven's grasp of things and his effectiveness as a tutor: I was concerned that he might have difficulty communicating to his learning partner. He showed signs of not understanding things and just *having quite a bit of difficulty* (February 16) during the practice session within the first Collaboration Group. Once

Steven met with his learning partner, however, he did quite well. Steven used a lot of encouragement and praise and he was patient with Seth, *he is responding so well, I am very encouraged, mind you, they have only been working together for about 10 minutes* (February 16); *he complimented and encouraged Seth regularly. He was quite patient and, thought not too creative with his helping skills he did help Seth to find the answers to his addition facts* (February 20); and *both Steven (and Aaron when they are playing paper ball) are very supportive* (March 6). He usually guided Seth to solutions, *Steven was helping Seth without giving him the answers, and thus far they were all correct* (February 23), and *Steven is doing a good job of leading Seth to find the answers without counting from the beginning* (April 10), though, occasionally, he tended to do the work for Seth, *Steven tends to do the work rather than guiding Seth to work when they are playing the game* (April 10). Often Steven's explanations were convoluted and confusing: *Steven tends to show Seth how to do things in a way that is confusing* (March 6) and

Steven keeps expanding the work. I tried to slow him down, because when expanding the work he is making it more complicated and not necessarily correctly. Steven forgets the order of the steps, but is getting it across, sort of. Steven was trying to explain something to Seth about odd and even, but I did not understand what he was trying to say. He was doing the number three, Steven had two of the counters and was explaining: pretend that this is an ice cream and that this is a person. The person likes ice cream. . . No, pretend

that this is a dog and that this is a . . . , at that point, I stopped Steven because this was not making any sense to me, and I suspect that it was confusing Seth, though it may not have been (April 3).

His explanations even confused me. This occurred even when he was explaining non-academic things *Steven also told me a story about how and why he was tired, though it did not make complete sense (March 13).* He also tended to explain the work even when it was not appropriate and he did not tend to expand effectively. Generally, he made things even more complicated rather than clarifying or usefully expanding the material. Usually, Seth had not even mastered the initial concept before Steven began the extension activity or concept. I suspect that Steven's confused explanations reflect his own confusion regarding the concepts.

Steven also had great difficulty with his reading and this sometimes hampered his effectiveness as he often could not read the planning sheets, *Steven was unable to read the notes on his planning sheet. Aaron helped him to read the notes (March 13).* Though he was sometimes able to figure it out on his own, sometimes he got help from other older learning partners or from me.

Sometimes Steven had trouble with problem solving, even when it was for his own gain. On one occasion, he told me that he could not come because they were having a party in his class. I asked him if he could come up with alternatives so that he could attend the party and attend to his responsibility. He could not come up with any alternatives, even

after several prompts from me.

Steven's sense of role and responsibility was sometimes clouded, though this is probably completely expected for his age. He felt that it was unfair to miss a party in his class when it conflicted with his learning session but was willing to come to the session when I confronted and then discussed it with him. Other times, Steven would seem to forget that his role was that of tutor and would take on the role of playmate: *when he was playing the game, he was also encouraging, but he may have forgotten that his role was that of tutor not of playmate. He took to the game and got quite absorbed in it* (February 20). On the other hand, he could go to the other extreme and take on the role of parent or teacher: *"Seth, math is not a joke. You have to know math to get a job. You can get a really good job if you know math. You know hockey players. They have to do good in math, you have to listen." . . . Seth said Miss J helps him, Steven pointed out that Miss J would not always be around. He also talked about how soccer players (what Seth wants to be) need two jobs, "They have to be really smart because they have to have two jobs"* (April 20).

Kelly

Kelly was not so much a concern from the start, and she was quite capable in the learning sessions, but her disinterest was a concern throughout the program. Though it is difficult to say for certain, it seems that Kelly was a fairly uninvolved girl who was quite uncertain of herself. There was much evidence of her noninvolvement throughout the field notes and in transcripts, but how much of that was due to Kelly's attitude and how much

was due to other factors was uncertain. Kelly's diffidence and indifference may have been a reaction to her learning partner, Sally, who very often expressed dissatisfaction, boredom and uncooperativeness. Another factor that may have influenced Kelly could have been the class that she was part of. Because she was in the same class as Ryan, I wonder if his attitude or even the attitude of their teacher may have had an influence on Kelly. When I asked Kelly about what she thought of Ryan's attitude she implied that Ryan's attitude was not affecting her own, *"It's his own comment. . . . I think that, urn, if he wants to quit or something . . . It doesn't have anything to do with any of us. . . . So it's like his own comment and like, the French teacher, he like didn't want to come down, and like he just grabbed his book and like [she bangs]"* (March 30).

My field notes are full of references to Kelly's disaffection: *Kelly is just sitting beside her watching Sally use the counters and write the answers. She says nothing to encourage or to discourage Sally, in fact, she says nothing at all. After I whispered to Kelly to encourage Sally, she said "You're doing very good," but then did not say anything else for the longest time. Very occasionally Kelly would say "Good" in a quiet voice but never participated in Sally's learning process. She just watched her, and Sally worked on her own* (March 9); *Kelly was, as ever, merely being quite quiet and watching much more than participating* (March 13); and *Kelly, however, just sat there* (April 20).

Even at the very beginning, before the learning sessions began, Kelly was cautious. During the last training session, I noted this in the field notes: *"Kelly, you don't look very*

excited. Aren't you looking forward to meeting your learning partner and working with your learning partner?" Kelly replied, *"I want to see who it is"* (February 13). She was willing to try when I specifically asked her to be more involved with Sally: *Kelly is still fairly disinterested, but, when I specifically encourage her, she does try* (March 20); and, *Kelly is looking at them too or she stares out, not looking at anything. She does seem to be more attentive to Sally* (April 24).

Near the end of the program, Kelly seemed to become more active and involved: *Kelly is taking more of a lead* (May 15); and, *. . . is a difficult game for both learning partners to get. Kelly asked if she and Sally could play it next time and I agreed* (May 18). When paired with Jeanne and Joe she seemed to blossom (though this happened only once and thus may have been coincidental rather than the sign of real change).

Kelly was also very forgetful when it came to attending the sessions. She regularly forgot to come down: *I went up to get Kelly from the computer room. She rarely remembers to come down by herself* (April 6); *Kelly did not come down, so I went to get her. Then we went to get Sally. . . . She still does not have the routine down of coming down here at the right time and getting to work* (April 20); *I had to go up and get Kelly to go and get her learning partner* (May 15); and *I had to go and get Kelly and Jeanne, again* (June 8).

At the beginning, I had some concerns regarding Kelly's ability to deal with learning sessions with Sally:

Kelly tries to help but does not have a lot of resources with which to do it. It

seems that Sally is really high needs when it comes to math. Kelly will need a lot of help in getting through to Sally and helping her (March 9); and I am not sure that Kelly gets as much out of it as some of the other OLP's do (April 6).

However, she was able to work with her quite well, and was able to bring Sally out of her bad moods:

She went to get Sally in Miss J's class, but Sally had not done enough of her work, so that Kelly had to wait for Sally. SJ came out and said that it was amazing how Sally had suddenly become motivated. All period she had done no writing and when Kelly came to get her she all of a sudden she began to work furiously to get done so that she could go with Kelly (April 6); and When she and Kelly began playing the card game she quickly perked up (May 18).

I told Kelly that I thought she was doing a good job, that I was proud of her. I empathized with her about how difficult Sally could be at times, but that she, Kelly, was handling it very well (May 1).

Kelly did not feel that she was learning anything, but felt that Sally, her younger learning partner was. This seems to be a pattern common to both the older learning partners and the younger learning partners. They tended to believe that the younger learning partners had learned, but that the older learning partners had not, *"I know all the*

stuff that . . . we're trying to teach her" (May 1).

At the end of the program, I asked whether she would do this again if she were given a chance, Kelly said that she would participate again. She also said that she enjoyed participating and working with her partner (June 8). This was particularly interesting because Kelly was quite forthright in expressing her various dissatisfactions with the program throughout that particular collaboration meeting. She said that she did not want to go to this "class" any more. She offered several reasons throughout the conversation, but eventually she got down to the two main reasons: that Sally often just sits there and will not participate or pay attention, and the collaboration meetings. Sally's uncooperativeness made her feel frustrated (May 1). Not surprisingly, Kelly was one of the three OLPs that said they would have liked to change partners.

Hailey

One Older Learning Partner, Hailey, was particularly exceptional. Not only was she the OLP that wrote the most comprehensive learning logs, she also demonstrated a remarkable talent in tutoring right from the beginning. She was creative, gentle, caring, encouraging, and yet demanding.

Hailey was often able to spot Cassie's difficulties or frustrations and intercept them. On March 9, I noted that, *I sometimes sit here thinking, 'Cassie do this or that' and often Hailey asks that question at that time.* Hailey was also quite sophisticated in her interpretation of some of Cassie's actions. For example, after remarking that Cassie always

put the same thing on her learning log, Hailey reflected, “*Maybe she really liked it or maybe she had absolutely no idea of what we were doing*” (June 8). Though she was able to see that she had an impact on Cassie (she referred to Cassie learning on at least two occasions in that collaboration session), she saw a possible underlying meaning to Cassie’s behaviour — that is, that Cassie had no idea of what was going on.

Hailey demonstrated remarkable talent with tutoring right from the beginning. In the training and collaboration sessions, she contributed some good ideas on how to show a learning partner that you care, “*When they make mistakes, instead of making a big deal out of it, just tell them it’s okay*” (February 9). In response to my asking how they can make the work the best that it can be, Hailey suggested that the older partner check over the work when the younger partner is done. Her suggestions in these sessions were not significantly better or more numerous than her colleagues. Where Hailey truly shone, was when she worked with other people, especially her partner. However, she was very helpful to other students as well. For example, she helped Molly when I had Sandra and Hailey and their partners work together on the Great Math Race.

Right from the start, in the Getting to Know You Session, Hailey was gentle and caring with her partner. She seemed to have a natural talent for teaching, both in her ideas for passing on knowledge and in her manner. She often came up with creative ways to present information and to make it clearer to her learning partner. For example in the second learning session, Hailey explained to Cassie, her partner, that, “*adding is really just*

counting” (February 23) and she used “think aloud” techniques (April 20). She tried to get Cassie to think through the questions and answer them without using counters. Hailey lead Cassie to correct answers, but did not give them to her. She did this through hints, demonstration, encouragement and solid questioning techniques.

Even though in April, Hailey seemed to be regularly quite tired, she demonstrated good teaching qualities throughout the program, even until the very last session. She was encouraging: acknowledging correct responses and helping Cassie to better understand questions when she gave an incorrect response. Even when Cassie was wrong, Hailey was almost always helpful and encouraging.

Not only did Hailey’s teaching impress me, but she seemed to spark new understanding in Cassie, who apparently thoroughly enjoyed working with Hailey, as I noted on March 9 and when Cassie herself discussed in a collaboration group on April 6, *“I really like working with Hailey because we’re doing lots of math together and I like doing that. And I learn lots of things from math and its learns how to count for people who don’t know how to count.”*

Unfortunately, one or both partners were absent for almost the entire month of May and the occasional time other than that. When a session had to be missed, Hailey seemed disappointed. Hailey also had pulled a tendon just before the program began and was in a cast and on crutches for a significant portion of the program, but this did not seem to dull her enthusiasm and dedication.

Though she was always understanding in the learning sessions she, like the other older learning partners, did complain about some of Cassie's behaviours in the collaboration sessions: *"They're always guessing. They never actually know the answer. Like, today, when we were working with Cassie, and I asked her if it worked she just guessed at it . . . It makes me feel frustrated and depressed because we spend all that time trying to learn something and then she's not doing too well on it . . . We have to spend all that time again to try to learn something"* (May 1), and *"I was thinking about my learning log today and I just realized that Cassie filled out the exact same thing every time. Nothing she didn't like. 'I like math. We play this and it was great'"* (June 8).

Again, consistent to the other learning partners, whether or not she felt that she got through to Cassie seemed to be a determinant of how Hailey felt about the program. When I asked what part of the program made the older partners feel the best about themselves, Hailey replied, *"When I actually taught Cassie something"* (June 8). And when I asked what made her feel the worst about the program, she replied *"When I'm not quite sure how to explain something to Cassie"* (June 8).

Aaron

Aaron is a boy who has had chronic difficulties with tardiness. This created a problem for the program because not only were the training sessions scheduled for first thing in the morning (8:20 am), but, in keeping with the schedule that his teacher gave me for participants from her class, Aaron's learning sessions were scheduled for either 8:20 or

9:00 am. On the days that Aaron was late, he usually did not come to school until the morning recess at 10:15 am. This was definitely a concern for me and began to affect the program right from Day 1 of the training when he came in late. I met with him after recess and briefly went over what he had missed. He was late again on Day 3 of the training and completely missed the session. That was the same day that Judy was introduced into the program so I took both of them in the afternoon and gave them a training session. This gave them some opportunity to interact with others during the training. Aaron then missed the last training session because of illness. I allowed him to go on with the Getting to Know You Session and then caught him up on the training. I was very concerned about his absences and shared this concern with him, dealing mainly with his lateness (February 16). I explained to him how important it was for him to be on time and that when he was late it was not fair to Dan (his learning partner). I also let him know that if this tardiness continued, he would no longer be part of the program. Aaron indicated that he understood, and apparently he did because he was not late again until March 27 but even then he came at 8:40 not the usual 10:00 and was not late again for the program after that.

Aaron was often tired, sick or injured. He regularly told me that he had been up late, usually helping some farm animal to give birth. During the program he sprained his wrist (he fell off a barn) and he quite badly burned his thumb while burning brush around his house. He seems to have a lot of responsibilities at home and these may be keeping him

tired and tardy.

In the training sessions he did attend (one group and one with Judy), Aaron was able to contribute some valuable input to the discussion. He made suggestions and expounded on others' suggestions. When Judy suggested that they could help their learning partner by just giving the answer, Aaron suggested that would not be useful "*unless you explain it out while you're giving them the answer*" (February 9).

As far as being an Older Learning Partner, Aaron seemed to do quite well. He supported and encouraged Dan regularly: "*Come on Dan. Come on Dan. Come on Dan!*" (March 13). He even demonstrated this in the training. When Judy got one answer, he clapped and cheered, "*Woo!*" He also gave appropriate direction and hints to Dan, he was helpful and challenging. When Dan responded to a question by saying "*[number] take away [another number] would be about . . .*," Aaron responded with, "*Don't say 'about.'* *Don't guess, you have to know*" (March 9). He also seemed rather insightful. During the Getting to Know You Session, he asked Dan what he most wanted to learn about math.

Regularly in my field notes there are references to Aaron's helpfulness and not only with his partner. For example, "*Steven was unable to read the notes on his planning sheet. Aaron helped him read the notes*" (March 13), "*Aaron wanted to help Steven when Steven could not read Ladybugs and Leaves. But I asked him to continue working with Dan*" (March 20), and, "*Aaron barely has to help him, though he is being attentive and watching what Dan is doing*" (March 9).

Obviously, his helpfulness and encouragement had some effect on Dan. When I asked Dan if he liked math more or less than before he began working with Aaron, Dan replied, *"Oh, I like it more."* I probed a bit with, *"You like it more. Okay, but a whole lot more or just a little more?"* And Dan responded, *"A whole lot more."* (April 6).

Another interesting point is that Aaron was the main one to laugh at Steven in the second day of training. He laughed and rolled his eyes at Steven and initiated the derogatory laughter. Yet, on another day, when I complimented him on the way he was treating Steven, he said that Steven is his best friend, that he really likes Steven, and that they spend a lot of time together outside of school. I told him that I was still pleased at how nicely he treated Steven. It did surprise me however that he had been the one to initiate ridicule during the training session.

It was especially surprising because he laughed when Steven forgot what he was going to say. Yet several times throughout the audio taped sessions, Aaron would put his hand up and then say that he had forgotten what he was going to say, *"I had something, but I don't know what it was"* (February 9).

Aaron, as with most of the participants, saw the older learning partner's role as exclusively that of teacher, *not* that of learner. In fact, at one point he protested that he was not the one there to learn (April 6). This is particularly interesting as Dan, Aaron's partner was the only one who thought that the older learning partners may be learning something from working with the younger learning partners.

It was obvious that Aaron was comfortable in the setting; he was often lounging or leaning or was in some other very comfortable position. Sometimes, too comfortable. I often had to ask him to sit more properly. This may have been related to his persistent tiredness. However, even when he was very relaxed and apparently not too attentive, Aaron remained on task, noticing and helping Dan with any errors he made: *"Aaron is sitting and just recording, not saying much at all, just waiting for the next fact. He is sharp however. When Dan was doing the 10's, he accidentally said 3 and 7 equals 9. Aaron picked up on that right away, even when I did not. He makes sure that Dan has the correct number of counters out, but otherwise sits huddled over the piece of paper that he is writing on"* (April 27).

Aaron was somewhat indefinite in his feelings about the program. He said that he thought the program was fun, that it helped the other kids. He thought that it was good that they were able to get out of class. He felt that he had not learned anything but that he felt more confident and liked math better (March 30). However, later, during that same collaboration group, Aaron indicated that he did not feel more confident. Another day (May 1), Aaron was not so sure of any of the positive aspects of the program. *"... it's like messing up with my spelling. . . . And we have like tons of homework. . . . It's boring. . . . And Dan always fools around. . . . Um, I miss my sports."* Yet that same day, he said, *"We're lucky"* to be part of the program. Another day, he put his hand halfway up when I asked *"Put you hand up if you enjoyed [the program] more than you didn't enjoy it"*

(June 8).

Judy

Judy became an older learning partner after the training sessions had already begun because another student's mother asked that he start in the second session. I gave Judy a special two-part training session, first on her own and then with Aaron (who had come in late and missed the training session that day). It was convenient that both of them had missed this session because it gave them an opportunity to work with role-playing and at least limited social interaction, both of which were such integral components of the training sessions.

As a learning partner, Judy was consistently enthusiastic and encouraging, at times too much so. On February 20, for example, she asked me four times whether she was teaching that day and whether it was her turn yet. Her extreme enthusiasm did concern me at times, *I think she will do well as long as she is not too helpful and encouraging, that is, as long as her caring comes across to Timothy as genuine and as long as she does not end up doing his work for him* (February 20). Later, I noted that *she refuses to not encourage [Timothy]* (March 13). Sometimes Judy would take encouraging and praise a bit far by telling Timothy that he was close, when in reality, he was not (February 20, March 13). This pattern seemed to fade as time went on and Judy became more stringent in her expectations of Timothy.

Sometimes, when Timothy was having a particularly difficult time of it, Judy would

become frustrated (April 3) or bored (April 24) and showed this to me. However, she was quite careful not to let Timothy see her frustration: *It was extremely frustrating. Judy was very frustrated, she was rolling her eyes and making faces that I could see but that Timothy could not see* (April 3).

Timothy was able to get away with less as time went on. Though she had always guided Timothy rather than give him the answers, at the beginning Judy did take what Timothy said at face value. Later, she frequently checked that Timothy was not just guessing or giving an answer that he did not understand. In fact at times she got on a role of questioning every answer that Timothy gave, "*How did you get that?*" or "*Are you telling me or asking me?*" (March 20).

Judy did not limit her willingness to help and her encouragement to just her partner. She always volunteered to get other students, to help me, and to help other learning partners. When I asked Timothy to go and get Steven, and Timothy was reluctant, Judy went without hesitation (April 6). When one day Steven asked no one in particular if he could borrow an eraser, Judy immediately got up, and without a word, gave Steven hers (February 23).

Repeatedly in my notes I used terms such as keen, eager and thoughtful to describe Judy. I also used organized and patient. Judy not only praised Timothy's successes, but also his unsuccessful efforts. In fact, the only time any of the older learning partners used an award certificate was one day when Timothy was having a particularly difficult day, but

kept struggling with the work. Despite his effort, Timothy's lack of progress was frustrating, even to me. After all that, Judy decided that he should get an award because he worked so hard. When she got back from walking Timothy back to his class I spoke to Judy. After explaining that I was not telling her that she did not do the right thing and that I just wanted to know why she did what she did, I asked Judy why she had given Timothy the award. She replied, "*Because he worked really hard today.*" So I questioned her further: "*Did you not think that today was a particularly hard day?*" Her response was a very inspiring, "*I just thought that he was trying better*" (April 3). It was very nice to see someone acknowledging effort and Judy did this regularly.

It was interesting to see Judy involved in anything that allowed her to be competitive or to show off. Though she was as keen and enthusiastic as ever, she nonetheless did seem "*eager to go first, to win, to be competitive*" (May 18) even when I incorporated as little competition into an activity as possible. This particular observation referred to when they were doing the Great Math Race in which each partner was competing against his or her own previous score. Judy, however, was competing against Timothy to some extent. Another indicative incident was when I told her about the program and that she would be teaching a younger child; she thought that others in her class may get jealous because of this. She mentioned this on the first day and then again in May when she said that, "*More kids hate me . . . because everyone thinks I'm dumb, because I'm teaching . . . I think because they're jealous because I get to teach and they don't*" (May 1).

Her goal is to teach as an adult and when she discovered that she would be teaching a child in grade one, she replied, "*Boy! I didn't plan on teaching till I was grown up*" (February 9). She felt that participating in the program helped her toward her goal because it helped to teach her how to "*handle the children*" (March 30).

When asked about the effect that the program had on Timothy, Judy was convinced that it had helped him: "*He is like trying a bit harder and he sort of knows them a bit better*" (March 30), and, "*He's even better. . . . He's better at math*" (May 1).

Judy's teacher commented that it was good for Judy. She said, "*I think it's been great for Judy. Yah, I really do*" (March 23).

Ryan

Ryan was the boy that left the program just before the second administration of the SDQ-I on March 30. Ryan seemed to initiate conflict right from the start of the program. Though I got verbal permission for him, and based on that permission, I administered the SDQ-I to him. Ryan was very negligent about bringing the signed form in. He was excluded from the first training session because he had not brought in his form. In retrospect, I wonder if this obvious and intentional exclusion of Ryan from this session was an impetus for the viscous cycle of indecision that eventually led to Ryan quitting the program.

Ryan made it clear, as early as that first training session, that he was not certain that he wanted to be part of the program. It is possible that this ambivalence was rooted in an

uncertainty that he could succeed in the program (based on his limited previous experience) and consequently, he set himself up for failure which he then blamed on boredom. In a way he reminds me of a student I worked with in the pilot project who was convinced no one liked him and set out to prove it by antagonizing those around him until he provoked some action which proved to him that he was right, that he was unlovable and unworthy.

The other possibility is that Ryan was genuinely in conflict about whether or not to participate. I feel very strongly, however, that deeper, underlying issues were the root of his early ambivalence and later outright rejection of the program.

I believe that his indecision was based on something deeper because certain indicators show how enthusiastic he could be about the program. For example, on the first day of training, I spoke to Ryan during the recess period and asked him why he was unsure of whether he wanted to be part of the program. Ryan said that he was not a good teacher. He had been taught by his mother how to teach his brother, but it did not work because of Ryan's impatience. Ryan said that he would try to teach his brother for five minutes and his brother would not be getting it and Ryan would lose his temper. After explaining to him that I would teach him to teach, that not even the best teachers are just born, but rather they have to learn about becoming teachers and have to be trained, that both I and his teacher thought he would make a good teacher and that I would support him, he said he would get the letter back to me by tomorrow. I suggested that he get the letter signed

as soon as he got home and that he put it into his school bag right away. I asked how he planned to remember to give it to his teacher the next day. He said he would remind himself, eventually he said he would write it on his hand.

After addressing his fears of failure, Ryan demonstrated an eagerness to remember to have his form signed. In fact, he not only showed an eagerness to remember, but even followed up on it by returning the signed form the very next day. Up to this point he had taken home three different forms and I had reminded him, his teacher reminded him and his mother had even signed a form that never got to the school.

His enthusiasm shone through a tough and apathetic exterior when I required a decision on Ryan's participation in the program in late February. During my previous day at Hillside School there was a typical example of Ryan's vacillation in his attitude toward the program. Ryan came to tell me that he did not want to be part of the program any longer. I encouraged him to continue, but stressed that it was his decision. I asked why he did not want to participate and he said that he just did not want to. He also said that he might want to participate sometimes. I replied that was not an option. He had the choice of coming every time or not at all. He said he would think about it, to which I replied that he had until Thursday to decide. I also told him that I really wanted him in the program, that I thought he would do a good job. But, I reminded him that the decision to participate was his. On Thursday, February 23, the deadline I had given him, I had a chat with Ryan. I explained that I needed his decision. He had to decide for himself whether or not he

wanted to participate and that if he decided that he did want to participate that he needed to be enthusiastic. He told me that he still had not decided whether he wanted to be part of the program. I reminded him that I had given him until that day to decide and that I needed his decision. I suggested to him, that if he had not made a decision up to that point, perhaps that in itself was a decision not to participate. I also told him that I would prefer for him not to participate if he was not certain that he wanted to be part of it or if he would not be enthusiastic. Then I left him alone and told him that if he had not joined me upstairs at his partner's classroom by the time Kelly was ready, I would know that he had decided to not participate. After making it clear to Ryan that the decision was his and that I was in no way pressuring him, I went upstairs. Ryan beat me in getting upstairs. In order to do that, he must have run down the hall and up the other set of stairs. This indicated to me a hidden enthusiasm for the program.

Ryan often demonstrated creativity and even what I termed "zeal" during one of my observations. He came up with good, real-life applications of mathematics concepts (for example in training session 3 he turned the question of adding 25 and 25 to adding two quarters (25¢)) and he tried working in inventive ways with Joe, taking things on himself, *"We're going to try an experiment"* (March 20).

This seemed to be working at least to some extent, according to the observations of Joe's teacher, who commented that when Ryan came in he always looked very mature, very in control. It looked to her as if the program was helping both Ryan and Joe. She

commented that Joe “*enjoys going. [Joe] enjoys when Ryan comes and gets him.*” KP also felt that Ryan was helping Joe with respect to math as well, “*turning right now in math, whether it's due to this or something else, [Joe]'s certainly catching the concept of place value and regrouping at this point. So, I think the combination of what's happening there and what's happening in class is helping him. It just reinforces the concept. I can definitely say that. . . . he's never said, "I don't want to go." He's always ready to go [snaps fingers] when he's asked*” (March 23).

Throughout his participation in the program, Ryan seemed to be always vying for some sense of control. I perceived this struggle beginning on the third day (his second) of the training, and I suspect that his persistent threats of leaving were also, at least in part, a control mechanism.

I believe that he understood both that I really wanted him to stay in the program and that he had the ultimate power to decide whether or not to stay. I think he also understood that I was willing to bend a lot with him because I felt that he, of all people, could benefit from the program. Therein lay his power. He repeatedly expressed his uncertainty about participating: February 2, 20, 23, March 9, 20 and 27 and of course 30 (the day he quit). He did not participate on February 2, 20, March 9 and 16. On February 9 and 13 he seemed to be in a power struggle over the audio taping. And on March 6, I observed that Ryan *was not willing to try too hard*. Of the 14 days that Ryan met with either the older learning partners or his learning partner, there were only four where Ryan was present and

did not demonstrate negativity.

In group situations, he understood that he had veto power over the tape recording and, while I believe that he had no difficulty with being audio taped, he saw this as another opportunity to exert control. Even if it was only to be the one to delay and the one who necessitated explanation. *Ryan seems to be very contrary. Whenever I say or suggest something, he challenges it. Even when I am giving the students a liberty they usually do not have, as in this case. It was not okay with Ryan to talk out. Later, he protests taping of the session (February 9).* After demonstrating that he had the control to stop the audio taping, he posed no further objections and he seemed to have no problems with being taped.

I suspect that Ryan was trying to give the impression that he was doing me a favour by being part of the program. That he was trying to rack up points. I suspect he wanted a way out, so that if he decided not to do this or if he did not succeed, he could say that he never wanted to do it but was only doing it to please me.

Ryan's personal life during the time of the program was in somewhat of an upheaval over which he had no control. School officials understandably believed that this was influencing Ryan's behaviour. Ryan may have been trying to control those aspects of his life that he believed he could. Thus, I suspect that Ryan took advantage of situations in which he had veto power, both with me individually and with the group.

Ryan did not usually express his negative feelings in front of his Younger Learning Partner, but he apparently did express them in his class and in the learning sessions. I was

concerned with the effect that had on the other Older Learning Partners, particularly Kelly who was not only in his class, but also shared the same learning session time.

Ryan expressed a reluctance to be seen as a learning partner. He did not want his classmates knowing what he was doing and seemed somewhat embarrassed. When I gave him a chance to start anew with the older control group, Ryan declined.

It is also possible that Ryan's teacher's attitude apparently somewhat negative attitude towards CALP might have had something to do both with Ryan's attitude as well as with Kelly's oftentimes less than enthusiastic attitude in the learning sessions. While their teacher did not come out and say anything to me, there were indications that she may not have been as enthusiastic as the rest of the staff seemed to be. She made comments at the beginning, regarding her concern over the wording of the letter; she was less than enthusiastic in her response; she was rather rigid in time slots that her students were available for me; and she took a long time in giving me names of participants.

Seth

Seth was the younger partner of Steven. I was a little concerned about Steven and Seth working together when I saw how they both had great difficulty with academics. However, they did seem to work well together. Seth did very well with his addition; he was able to get 51 facts done in one session, with all of them being correct and only a handful of the numbers written backwards (February 23). Seth did seem to have difficulty writing numbers over 10 (February 23). Though Steven commented that "*Seth really knows his*

pluses.” While Seth was quite stronger in mathematics than Timothy, he was still relatively weak. He was unable to even remember that $10+2$ was the same as $2+10$ when the concept was repeated to him several times.

Seth said that it was “*more fun*” doing math (March 23) and that he learned more, “*I learn more here*” (April 6), in the learning sessions even though he did not like doing math. He said that he was a better person because he learned more as a result of his participation in the program but could not articulate why (April 6). He often commented on the learning log that he liked the session because he had learned a lot. When pressed about what he had learned on one occasion, he was unable to articulate what he had learned and eventually changed his reason to why he liked the session by saying that it had been fun. His never wavering response on the learning log frustrated Steven (May 1).

Even though Seth had a caring and helpful learning partner, I am sure that he got confused by Steven at times. Steven’s explanations could tend to be somewhat convoluted and even occasionally wrong.

Seth’s teacher commented that Seth was, “*more with it*” because of the program and that even though other kids are having a hard time with money, Seth was getting it. His confidence seemed to have improved and he was really keen about the program (March 23).

Sally

Sally is a very determined and organized little girl, "*[Kelly] just watched her and Sally worked on her own, never seeking help, never looking for encouragement*" (March 9), and, "*She works diligently and without complaint. She is organized in that she put a large dot beside all the addition questions that Kelly and I pointed out as needing correction and then she proceeded with correcting them*" (March 9). She was diligent when she was on a roll, *As soon as she finished her addition correction, she immediately began working on the subtraction again* (April 6).

She was able to get 70 addition questions done in one session, with only three errors. This is very confusing as she usually had great difficulty with even basic addition and subtraction. Her greatest difficulty seemed to arise when she had to do things in her head. If she was allowed to use manipulatives, she was quite efficient. Compare statements like, *She was not able to add $3 + 7$ in her head. She is completely lost when she has to add. . . . It took several interventions for Sally to be able to figure out that there were six paper balls left in the bucket when there were three outside the bucket* (March 9), and, *Sally is really poor in adding. She basically has to count the pips on both the cards to get the answer* (April 6), with statements like, *She realized after a while that to use the counters she just had to count out the top addend once and then just add the number to the bottom. Once she had the constant addend for the row counted out, she never really seemed to refer to it at all. She just counted up the counters after the constant addend* (March 9), and, *Sally is doing*

very well in solving the spaces. I have had to ask her not to count but to try to figure out the spaces based on patterns (April 20).

Sally often looked miserable and was uncooperative in the learning sessions (March 9, April 3, May 15, May 29). She would often refuse to respond to me or to Kelly: *When she came in Sally looked miserable. She would not answer my questions and only nodded when I asked if she had not wanted to leave her teacher's class. When I asked her if she would answer my question of what was wrong, she shook her head after I asked her to just answer that question (whether or not she was willing to answer a question). Then I asked her if she would be able to work with Kelly today and she nodded. I decided to start them off by playing the paper-ball game. She perked up a little bit when she and Kelly began to play the game (March 9). After getting her involved with what she was doing, she usually was able to get into her work and able to contribute to the session: Sally was in one of her moods when she came in. When I asked how she was she said, "Not fine." She was uncommunicative after that. I asked her what was wrong and she would not answer. I noticed that she had some stamps on her hands. I asked her where she got them. After a couple of attempts at getting her to answer, she finally replied with her teacher's name. I observed that her day cannot be all bad if she got two stamps from her teacher. Then we started on her Ladybugs and Leaves. She started off slowly but, when she got started, with some help, she was off and running. She was pleasant and had apparently forgotten whatever had brought her down in the first place (April 3). In fact, she often became very*

pleasant and seemed to completely forget whatever it was that had her disgruntled in the first place (April 3, April 24). At one point, Sally's resource teacher commented on how she thought that it was amazing how Sally had suddenly become motivated to do her work when Kelly came to pick her up and she was told that she could not go with Kelly until her work was finished. She seemed very enthusiastic about going with Kelly (April 6).

Rarely, her enthusiasm would diminish from a relatively high point to a low one. On May 15 she enthusiastically asked, "*Can we do it again?*" when they finished The Great Math Race, but soon after starting changed her tune, "*Do we have to do it again?*" At the end of that session Sally kept repeating that she thought the session had been boring and she became quite sullen.

Sally's case was made even more unique by the fact that her partner was Kelly who was less involved than many of the other older learning partners. Kelly often just sat and watched Sally, but, without intervention from me, she did not participate in Sally's learning process (March 9). In some cases, I suspect that Sally's apparent boredom was the result of Kelly's aloofness in addition to Sally's apparent boredom affecting Kelly.

Sally's regular classroom teacher felt that participation in the program had done nothing for Sally (March 23). About two thirds of the way through the program, Kelly said that she had detected no change in Sally. Neither she nor the teacher commented on Sally's progress at the end of the program.

Sally seemed a little confused about whether or not she thought she was learning as a

result of being in the program. About two thirds of the way through the sessions, Sally was looking back over the work she had done with Kelly and when she came upon a multiplication sheet that she had done, announced that she did not know how to do multiplication and that she never did it. During one collaboration session, she first said that she was not learning anything, and then later said that she thought that she learned more in the learning sessions than in her classroom. She also claimed that she had not only learned, but that she was better at math and liked math a lot more now than at the beginning of the program and that she was a better person as a result of having participated in the project (April 6).

Cassie

Cassie and her partner, Hailey, apparently knew each other quite well and were friends outside of school. Both were pleased when they learned that they would be working together.

Cassie seemed to really enjoy working with Hailey (March 9, April 6) and was keen to participate (March 23). She expressed that she felt she was learning as a result of working with Hailey: *I really like working with Hailey because we're doing lots of math together and I like doing that. And I learn lots of things from math and its learns how to count for people who don't know how to count* (April 6), and, *Um, I really like being here because I learn lots more math games and, um, math things* (June 1).

She claimed to learn more during the sessions than in the classroom “because it's

harder by yourself than with a partner" (April 6). She said that she felt she was a better person because of her participation, but like most of the younger learning partners, she could not articulate what that meant.

Sometimes, either she was confused or she had a very unusual way of learning math along with a keen understanding of her math learning strategies. I asked the YLPs whether there were other things that they were doing at home or in the classroom to make them feel better about themselves or about math. Cassie replied with "*Writing.*" To which I responded by asking, "*Writing. That's going to help you feel better about math? How?*" Cassie explained, "*Um, about the words and I think about the math that way*" (June 1).

Cassie had quite a bit of difficulty with her math (February 23, March 16, April 10, 24, 27), but Hailey seemed to be reaching her with unique and creative explanations of the math concepts. With Hailey's help, Cassie seemed to gain a "new understanding" (February 23). Whether her understanding was long term, however, was questionable at times (April 22). At other times, she really seemed to have the concept, even over a few days (April 24). When she was working with Hailey (and I understand even at other times), Cassie was a diligent and steady worker (April 27).

Unfortunately, because of both girls' illnesses and because of Hailey's softball tournaments, the pair missed many sessions. In May, they missed a whole three and a half weeks in one block.

Dan

In the learning sessions, Dan seemed to be a fairly bright, competent boy. He was able to do the simple math relatively quickly and independently, *Dan did all his addition facts quickly and quite accurately. He got only one wrong and he quickly and without help was able to correct that one question out of 18* (February 20). However, he said that he was not smart (March 16).

While he had some difficulty in math, he was not terribly behind. Though Dan seemed somewhat more “on the ball” (more so than most others), at times it was quite obvious that he did have difficulty with math *I asked Dan several times what $7 + 8$ and $8 + 7$ were. He had a hard time remembering, but he was able to do it* (April 6), and, *With some fairly heavy leading, Dan finally figured out the pattern. . . . Dan is still having some difficulty following the pattern* (April 24). Even when Dan became flustered, he tended to recover well. He seemed to be an efficient worker, *I said, “Dan, let’s see how quickly you can do this?” They complained a bit but then Dan started working furiously. . . . Dan gets a little flustered when he is under a time pressure. But eventually he just begins to work quickly* (April 27), and he seemed able to self-correct as well, *he was able to correct once Aaron and I showed him his error* (April 24).

He was cooperative, helpful, diligent and productive. His classroom teacher explained that he was getting more work done, was more “keen” and was keeping up better in class, *“he is certainly getting more of the work done than he would have”* (March 23). His teacher

also commented that Dan looked forward to the learning sessions and always came back from them in apparent good spirits (March 23).

Dan was the only YLP to feel that the older learning partners were learning (April 6). Aaron, his partner did not feel that he (Aaron) had learned anything. This perspective of Dan's shows a level beyond most others in the group. He was able to see past the obvious, the YLPs learning, and to see the possibilities of the program.

He said he felt that as a result of being part of the program, he was a better person because he was smarter and had learned more. Dan was the only YLP to be able to even begin to articulate why he thought that being smarter or having learned more made him a better person — because he would get better grades and *"Because I pass more school subjects"* (April 6). He felt proud when he was doing well (June 1). He could not explain beyond that. He said that he felt good about the program because it helped him learn, *"I like to come here because, um, you learn a lot of stuff"* (April 6). Dan also said that he liked math a whole lot more than he did before he started working with Aaron (April 6).

Like the other YLPs and, in fact, even the OLPs, Dan seemed to have difficulty expressing why he chose the happy face rating he did for the daily session. As with the topic about being a better person, Dan was probably the most expressive of the YLPs on this matter, but he still was vague: *Dan again said that he liked it because he learned.*

When I asked him what he had learned he said that he learned to count and stuff (April 27).

Timothy

Timothy was very weak in mathematics, *Timothy is very weak in even simple addition* (March 13), *Timothy is having an extremely difficult time with his adding* (March 27), and, *I was shocked to discover that Timothy does not know his numbers. He can barely recognize the single digits, has an extremely difficult time with ten and twelve and some difficulty with 11 and 13-19* (April 24). Even counting and very simple addition were a challenge for him. One day, after he had been away from school for two weeks, it took him five minutes and several attempts to answer the question $6+2$. He had great difficulty with recognizing and even greater difficulty with writing numbers, especially those over 10. It not only took much effort on his part, but much coaxing on the part of his older learning partner, Judy, and myself, for him to complete any amount of work. Often Timothy's learning sessions were difficult and a real struggle for Timothy, Judy and even me (March 13, April 3). Timothy was usually unable to explain to Judy how he came up with the answers that he did (March 20).

His learning partner sometimes had difficulty wording questions to help him arrive at appropriate answers and sometimes tried to explain but was not always clear (March 20). Sometimes, Judy herself was unsure of the correct answers for questions (March 27).

Despite his great difficulty and frustration, Timothy did feel that he learned more in the learning sessions than in his classroom, but he could not articulate why he thought that (April 6). He said that he also liked math more as result of his participation and that he

felt that he was a better person because he was smarter — echoing the response of other younger learning partners (April 6).

Joe

Joe was the younger learning partner of Ryan. He seemed to have some difficulty at times, including during the questionnaire administration, *Joe had some difficulty, but I think he understood and I went back over some questions when he lost his place* (January 30), but seemed to be able to regroup quite quickly. However, *Joe had some difficulty understanding the concepts at first, but then he seemed to catch on more* (March 13). Often Joe worked at a quick, steady pace with a good degree of accuracy (February 23 and May 29).

Joe was always ready and willing to go to learning sessions and, according to his teacher, his regrouping (the concept that he was working on in the sessions) had improved in class. Though he seemed to and even said he enjoyed being a learning partner (March 23), after Ryan quit and Joe was given the opportunity of continuing the program right away with a new partner, he declined. He did, however, come back to the program with the Non Learning Partners.

After less than two months of participating in learning sessions, Joe's partner, Ryan, left the program after much grumbling and vacillation. When I told Joe of Ryan's decision, I was very careful to be clear that Ryan's decision was in no way a reflection on Joe. Joe said that he understood and it seemed that he did.

Individual Interviews

Two of the learning partners, one older and one younger, and their teachers were interviewed at the end of the program. These interviews yielded more pertinent information than the document analysis, observation and formal and informal group interviews. While most of the information in this section did come from the interviews, there are also references to what these two girls said and what I observed throughout the program.

Sandra

Sandra seemed to have a fairly good sense of her abilities as an older learning partner. She seemed quite realistic and her self-concept seemed to be grounded on reality. When I asked her how well she thought she had done, she acknowledged that she had done *"Not so well. . . . Cause, I'm not used to doing, like teaching people stuff and I don't really have patience"* (June 12). She saw the improvement that I had seen in her, but attributed it to a better knowledge of Molly: *"Better. Because I knew Molly a little bit better . . . Well I knew what she knew and what she, like, needed"* (June 12). She attributed her change to interest and the possibility of a learning experience for both her and her partner: *"I did change. Because I was more interested in the program. Because I knew I was going to learn something and Molly was going to learn something"* (June 12).

Sandra's attitude changed during the course of the program. Near the beginning, she was uninvolved and concerned about keeping up in her class work. Then, after I discussed

the situation with her teacher, her mother and her, she made a real turn-around: *I'm not sure exactly what had caused a considerable change of attitude in Sandra, but it definitely seems to be there. I asked Sandra to come back after dropping Molly off and I told her how much I appreciate her change of attitude and how I notice that they both learned more today and seemed to have more fun. She was pleased with my comments, I think. She had a big grin on her face . . . Sandra seems to be in better spirits and is much more cooperative. She even seems enthusiastic (March 27), Sandra is improving vastly . . . She is much more involved with her partner than she was before. She is able to relate to her and to keep both their attention (April 6), and, Sandra is quite good at helping Molly to break some of the items down into steps. I asked her to do so specifically for square 10, and she did a great job (May 25).*

While Sandra had commented that the work she was doing was easy, she sometimes had trouble with it after she had made the comment. When I asked her about that at the interview, she could not remember ever saying it, and, in fact said, *"I never remember saying, 'this is so easy.' . . . Because I know if I say that I know, what happens if I get it wrong. . . . Because I want the person to know so they can explain it to me"* (June 12).

She claimed that she hated school, *"I hate school"* (April 24), because of the work, because, *"school is too long and it gets too boring"* (June 12), and because, *"I don't know. I'm a kid and kids don't like school"* (June 12). However, she did like coming to the CALP program, *"I like coming here"* (April 24). Later she explained that it was because she got

to miss work in the classroom and because *“like if you helped Molly she would, like you know you're helping somebody”* (June 12).

She was quite certain that the program did help her enjoy math more: *“Oh yah, except, well I like math more. . . . Since I came here [to the program] I guess. . . . Because, I am more interested in it. . . . Yah, I know, I know [inaudible]”* (June 12). On the tape, I had repeated that inaudible point, *“You know your math better”*. When I asked her if she thought that the program helped her, she said that she thought it did, *“helped me to understand better, if I didn't remember. Like, I learnt it when I was younger and I don't remember. Because still now we have tests and they have pluses on the them. And when I was little I didn't know what plus meant”* (June 12).

Sandra seemed to have some difficulty with expressing how things made her feel. The usually verbose Sandra said little about how she felt in reaction to a comment from her younger learning partner, Molly. Molly had said, *“You know, I like doing math. And Sandra helps me with my math,”* (March 23). I asked Sandra how that made her feel and why. She responded with *“Good . . . because you know you're like They said you are. . . . Helping them”* (June 12). When I asked how some positive comments that I had made throughout the program made her feel, she again responded with a brief, *“Good,”* and did not elaborate when I probed for more.

I saw certain behaviours apparently being exhibited by Sandra in the learning sessions, her teacher commented that some of these behaviours seemed atypical. Her teacher did

not find Sandra to be bored, bossy or competitive, though I did find her thus on some occasions. Her teacher did however agree that Sandra was quite forgetful, silly and even scatterbrained from time to time.

Sandra's teacher agreed that there had been some changes in Sandra, but that it was impossible to tell what had caused the changes. Her teacher felt that Sandra "*seems a little more accepting of other students. . . . She is more at ease with herself*" (June 12). However, at the same time that the CALP was starting, a special support network, "*an intensive home-teacher-Sandra intervention,*" for Sandra kicked in. Her teacher could not identify the source of Sandra's changes, "*It could be that with the tutoring she takes her school work more seriously now, but again, I can't narrow it down to just the tutoring*" (June 12).

Molly

Molly's math skills were relatively high, compared to the rest of the younger learning partners. She also tended to work quickly and accurately, *Molly got 63 facts done plus three of the four 0 facts. She had them all correct* (February 23). In fact, at times it seemed that Molly was functioning at a higher level than Sandra was. Often Molly demonstrated a better concentration, memory and understanding than her older learning partner, *Sandra was to continue doing the Rules Rule activity but she forgot how to do it. Molly remembered* (March 20), and, *Sandra called out "finished" when they were done the first page. Molly showed her that there was a back to the page* (April 20). Some of Sandra's difficulties seemed to dissipate somewhat after her change in attitude. When

pressed, Molly did acknowledge that even though she saw herself as a learner when working with her learning partner, she sometimes saw herself as a teacher when she gave Sandra answers to questions that Sandra should have known, *"Only sometimes because sometimes Sandra doesn't know the question but I know the question. Sort of"* She went on to say that sometimes she did see herself as a teacher and explained that, *"Like sometimes she has a question she's supposed to answer because she forgets sometimes, but I tell her the answer"* (June 12).

Regardless of the similarity of ability, even before the change in attitude, Molly still felt that Sandra was helping her with her math, *"I like doing math and Sandra helps me with my math"* (March 23), and on April 6, *"I like it. It's . . . Well, it helps me learn. . . . Because my partner Sandra is well . . . Because Sandra helps me with math."* Half-way through the program, she felt that she learned more in the learning sessions than in her class, that she was better at math and that she was a better person as a result because she was smarter (April 6) and at the end of the program, she said she felt better at math because of Sandra's help. Molly seemed somewhat confused about whether or not she felt any differently *about* math at the end of the program as opposed to the beginning of the program. At one point during the individual interview she said she did not feel any differently about math and at another point during the same conversation she said that she did feel differently, she felt smarter in math, better, because she learned more and she was more confident about her abilities (June 12).

At the end of the program, Molly said that she had not noticed any change at all in Sandra throughout the program. I suspect that if she had noticed the change I had noticed in March, she might have forgotten about it by June.

Molly seemed somewhat confused about what she had learned from the program, at first she said her reading was better, and then added her *"take-aways. . . . My adding, pluses"* (June 12). Later, she was a bit more articulate, *"Well, she was like helping me with my pluses. Trying to really help me to memorize them. Odds and evens, lots of math stuff, take aways, pluses and stuff like that"* (June 12). And she did say that she felt smarter, *"I felt smarter. . . . I felt better. . . . Because I learned more. Because I felt better because I was a bit smarter because Sandra helped me with my adding and take away. That's all"* (June 12), and better about math as a result of Sandra's help, *"Well, because she like really, really helps me in and . . ."* (June 12).

She said that she enjoyed the program and working with Sandra: *"I thought it was fun and I liked when Sandra was helping me because it made me feel smarter. [Smarter] in math."* Molly would not have changed partners (or anything else in the program) if she had the opportunity (April 6). She also commented that she felt that she could do better at math now. This sentiment was echoed by her teacher.

After just over a month into the program, Molly's teacher had not noticed any dramatic changes, but was only able to speculate that, "Molly's confidence may have increased a bit" (March 23). After the program was over, however, her teacher felt that Molly's self-

esteem, her ability to risk, to volunteer information and her competence in math had increased, *“before, she would struggle more with the concepts. She was always the one that stayed behind and I had to give extra work to and um, after, after the cross age learning, she seemed to be able to pick it up more, pick it up more quickly and be able to go to her follow up activity more quickly. She seemed to always have a block, always seemed to be the one to ‘I don’t get it, I don’t understand it.’ And after she was not always there with the kids who didn’t understand it. She would often pick it up and go off . . . she came to math class with more knowledge”* (June 12). She felt that Molly was also more willing to teach other kids. Her teacher attributed the changes to the Cross Age Learning Program, *“I didn’t see that behaviour before she went to Cross Age Learning”* (June 12). She said that she definitely saw a change in Molly after about six to eight weeks of the program: *“I think her self-esteem increased, I think her ability to risk in math class increased and her competence in math increased. . . . [She was] more willing to volunteer information. More willing to say, ‘I know how to do that’ and go about and do it”* (June 12).

She thought that Molly’s self-concept and confidence also improved: *“I got the sense that she did feel better and I could see that she did learn because she came to the math class with more knowledge. . . . It seemed. I mean this was not, you know dramatic. It was somewhat subtle, but I could see it. . . . She always had a fairly positive attitude, except that she often would be the one that said, ‘I don’t get it.’ And afterwards she still had a positive attitude, she never expressed a feeling that she hated math. So, I would say that it changed*

only slightly. *She was just um, she felt that she could do more.*" (June 12).

With the limited set of data that indicated any positive changes, I was a bit sceptical of this teacher's comments about how she thought that Molly's self-concept improved.

However, the teacher went on to say that she had seen no evidence of change in her attitude or self-concept with regards to others or to herself in general: *"I didn't really see that. . . . I didn't see, . . . I don't know if I can really say that I saw a change"* (June 12).

These more negative comments encouraged me and gave more credibility to the earlier comments about some changes in regards to self-concept in math. Had all the comments been positive I would have suspected that the overall comments were made to please me, to help the study or to for some other positivity biased reason.

Sandra also agreed and thought that Molly felt more confident and that she was learning. When I asked if anybody felt that their partner was more confident as a result of the program, Sandra replied that she thought that Molly did (March 30).

Again, as with Sandra, I saw things with Molly that rather surprised her teacher. I asked if Molly was usually *"particularly with it or bright with instructions and with understanding things more than other students might in the classroom"*? Her teacher seemed surprised by this information, and when I elaborated, her teacher replied, *"Wow!"* (June 12).

This teacher seemed impressed with the program: *"It seems effective in terms of, you know, I am really impressed with it and it seems effective in terms of adding to the children's*

knowledge and understanding of math concepts just in the math area and it seemed to affect Molly's self-esteem about her performance in math and her ability to take on new challenges, new concepts and new learning" (June 12). She responded with an enthusiastic "Absolutely" when I asked if she would do it again.

Summary of Qualitative Results

It seems that the Older Learning Partners felt best about themselves when they believed that their learning partners had learned something and were cooperative. These positive feelings seemed quite short lived in that the OLPs did not feel good if the learning partner was not doing well. It is also unclear whether these positive feelings were a part of or a precursor of an improved genuine self-concept. Perhaps, with enough positive elements like these, accumulated over time, there could be long lasting effects on the self-concept. Such a self-concept would not seem to match any of the self-concept factors addressed in the SDQ-I. It would have encompassed more than one of the first-order scales on the SDQ-I. It would probably include elements of General-school, General-self, Peer Relations, and possibly Mathematics Self-concept. But even elements of these scales would not be sufficient in describing such an aspect of self-concept.

The older students all felt that in the tutor-teacher role they were not supposed to receive, but only give; they were not supposed to learn, only teach. This attitude may have had a definite influence on their self-concepts, both in how they talk about their self-concepts and in their results on the SDQ-I. One OLP did, however, experience a belief

change over the course of the program: Sandra came to believe that she had been a learner as well as a teacher.

The OLPs also felt that their younger learning partners had learned and had increased in their self-confidence.

Two of the Older Learning Partners, Sandra and Kelly, definitely seemed to improve as time went on. They became better at their tutoring role. On the other hand, one of the OLPs, Ryan, quit the program at the end of March. He had vacillated between apathy, dissension and enthusiasm from the beginning of the program. Clearly, the change in him was for the worse regarding CALP.

The younger students generally enjoyed the program and felt that they had learned from it. They said that they liked math more, had fun and enjoyed working with their older learning partners. Saying that they were better people because they were smarter was a common statement by the younger learning partners. I wonder if this was not the reaction they thought that I or their learning partners wanted to hear rather than what they really felt. They often could not articulate what it was that they had learned beyond saying that they had learned.

Both the older and especially the younger learning partners had some difficulty in articulating exactly why they felt as they did about the program. Sometimes this seemed to indicate that the answers to why a learning session went as it did were either given without a lot of thought, the easy answer, or given to make someone feel better, either the

learning partner or perhaps me.

These conclusions echo those of previous research. These include such elements as decreased truancy and tardiness, (Lazerson, Foster, Brown & Hummel, 1988), an increased self-confidence (Martino, 1994), and some participants were able to better identify with their teachers (Paolitto, 1976).

The teachers whose students were involved in the program were generally cautiously optimistic about the program. They generally saw improvements that were subtle or modest and they were unable to directly link the improvements to CALP. Nonetheless, they did make positive comments regarding the program: *"I think it's worthwhile even if they didn't progress academically. It has to do something for self-esteem. Sometimes self-esteem is a hell of a lot more important than academic performance. The time element. I'm not sure whether the time element was sufficient for the tutor or tutee to progress in anything"* (Sandra's teacher, June 12), and, Judy's teacher commented that it was good for Judy. She said, *"I think it's been great for Judy. Yah, I really do"* (March 23).

Chapter 7 — Concluding Comments

The results of this study were mixed. Quantitative results showed no real results, however, the qualitative evidence was much more valuable. There were no meaningful, significant quantitative results, only trends in the interaction plots to show that there may be some patterns toward increased mathematics self-concept for the students who participated in the program. In the qualitative findings, there was evidence that the participants learned in and enjoyed the program and that those involved in the CALP had more positive feelings about themselves as a result of their participation; the students had increased confidence, self-perceptions and perhaps even self-concepts.

There were essentially three important qualitative findings. There seemed to be a reciprocal relationship between how well one partner did in a session, particularly the YLP, and how the other partner felt about the session, the program and about himself or herself. If the younger partners did well, then the older partners felt good. Many of the participants said that they felt they were better people because they were smarter as a result of the program. Both the older and the younger learning partners saw very delineated roles for the two groups: the OLPs were there to teach and the YLPs were there to learn.

The OLPs liked when the YLPs were cooperative and successful and when they both (the OLP and the YLP) accomplished something. The OLPs did not like those sessions where these elements were missing. The YLPs also liked it when they accomplished something during a learning sessions. These sentiments were expressed in the document

analysis, in the observation as well as in the interviews.

The YLPs made a connection that they felt that they were better people because they were smarter as a result of the program. Obviously, they equated doing well in school with being better people. The children felt better about themselves and this is what self-concept is all about.

The OLPs felt that their role was only to teach — they were not supposed to learn and they did not learn. In fact, at times the OLPs almost resented when it was obvious that I was trying to get them to learn something, the YLPs shared these sentiments. Only one of the OLPs admitted that she might have learned something, and only after she was specifically prompted.

For the vast majority of learning sessions, the partners felt that the sessions were a positive experience. Many participants and their teachers felt that participants had more confidence and that they did better in their class work.

In addition to these exciting findings, there were also many difficulties encountered in this study. My own limitations played a role in these difficulties but there were other aspects that also played significant roles. These included the complexity of the construct of self-concept, the limitations of the instrument designed explore the self-concept, as well as the limitations of the study imposed by practicality, by the environment and by the convenience of the participants and their school and home situations.

The phenomena which the study explored, those connected with self-concept, are very

complex, as are many phenomena within the social sciences. Self-concept is an internal, subtle construct and process and is probably not easily accessible through simple questionnaires, observation, document analysis or even through non-intensive interviewing.

Certainly, more in depth exploration of the participants, their teachers, their parents and perhaps even their peers would be required to shed more light on the construct. To get a fuller understanding of self-concept, questions would have to delve more deeply into the self-concept and interviews would have had to be more extensive — it would have been beneficial to spend more time with the interviews and to interview each interviewee more than once.

There were also difficulties with the quantitative instrument and its scoring. There were numerous inconsistencies within the instrument and its manual. While it was possible to overcome most of these inconsistencies, the amount of incongruity raises some questions regarding the overall soundness of the instrument. In addition, the control scores for the instrument showed that, for this group of participants, there were many inappropriate control scores. There were so many, in fact, that I question the validity of using this instrument with this group.

Logistics of the field site made it impossible to have the relatively equally matched groups that I had worked to achieve at the beginning of the program. Other distractions and interruptions of the regular school calendar and school day and parents requests and requirements also interfered with the program and with the adherence to strict

experimental requirements for the quantitative portions of this program. These interruptions, however, were probably very indicative of the conditions under which a program such as this one would operate in the real world. Thus they gave the more naturalistic components of the study more validity and more depth.

The time frame of the study was also limited as was the time spent at the school and other elements, such as the organization and placement of participants. The major issue here was the time spent in the school and time spent in learning sessions. Even though I was at the school from January till June, a total of five and a half months, there were only 23 learning sessions in all. Each of the participants missed a few of these sessions, and some missed several sessions. At only a half hour each day, that means that there were fewer than eleven hours spent in cross age learning by any one partner. These time restrictions made significant changes unlikely.

It is possible that the trends that were in evidence in the interaction plots could have been indications of what might be significant differences in the long term. This is particularly possibility for the OLPs and the for LPs in general, as these two groups showed increasing scores on the Mathematics Scale over time. That is, the mean Mathematics Scale score were greater at administration 2 than at administration 1, and greater at administration 3 than at administration 2. The YLPs, however, had a greater mean score at administration 2 than at administration 1, but then it levelled off and actually even went down by a fraction of a point at administration 3. The ONLPs, the YNLPs and

the NLPs in general showed quite steady or even slight downward trends on the Mathematics Scale. Perhaps the scores for the OLPs and the LPs would have continued their upward trend given more time, and this may have been enough to show a statistically significant result. Responses from participants, both in the individual interviews as well as in the collaboration groups, would presumably also reflect any longer term effects if partners had met for more than 11 hours of learning session time.

Self-concept is a fairly stable construct which does become more susceptible to change as you move down the hierarchy proposed by Shavelson, Hubner and Stanton (1976) and revised by Marsh and Shavelson (1985). Global self-concept is the most stable; the second order factors, Reading Academic, Mathematics Academic and Nonacademic self-concepts are less stable; and the first order factors, including Mathematics, are the least stable — and thus the most susceptible to change. We do not know how sensitive to change any of these levels are. Self-concept may be relatively more changeable at the first and second order factors of the self-concept, but these levels may still be very resistant to change.

We want to build genuine self-concepts and self-esteem, not inflated ones.

Accordingly, we need to work toward strong, realistic and stable levels of self-concept.

Asking a program, where children spent less than eleven hours in direct learning session interaction, to make a change toward this stable self-concept is certainly asking a lot.

It is very understandable why there have been such inconclusive results regarding cross age and peer teaching and self-concept in the literature. At the beginning of this study, I

posited that this inconclusiveness had to do with measuring general or global as opposed to subject specific self-concepts. In this study at least, inconclusive quantitative results were found — even when a more specific self-concept was targeted. However, there are indications that there may be trends toward an increased math self-concept and that over time these trends would continue, at least for the older students.

Though there were many difficulties with this study, I would not trade the experience that it has been. I would, however, do things differently if I had the study to redo. I also have uncovered many questions that I feel are worth further examination.

If I were to redo this study I would make different choices on a variety of aspects of the research process. These alternative decisions would begin with changes to the data gathering procedures. Those changes would hopefully translate to more meaningful data analysis and this would be a foundation for more valuable discussion based on relevant information and understandings.

The training of learning partners would have been more clear and more directed at issues that arose from the program. These included how to deal with the YLPs when they were distracted, disruptive or just not paying attention and how to ensure that the YLPs were understanding what was they were doing.

I would have made the intervention longer with much more contact time between partners. This is especially important when considering that self-concept is a fairly stable construct. However, even a more intensive interaction would be quite minor compared

with the countless variables, experiences and life environment that shapes human beings. Clearly, as this study such demonstrates, there are no “quick fixes” — not even with a relatively well-defined and narrow area of expected influence, in this case, mathematics self-concept.

The data showed that choice of YLP was important to the OLPs. While in the situation of this program giving them a choice would have been very difficult, I believe that it would be beneficial to both the YLPs and the OLPs to give them some say, or at least some preferences, in who their learning partners would be.

I believe we must find ways of integrating children’s view and of having them accept OLPs as learners and YLPs as teachers as well — not just seeing OLPs as strictly teachers and YLPs as strictly learners. Learning partners need to recognize themselves as co-learners. I would clarify and emphasize this role to both the older and the younger participants.

I would have used a different instrument to examine self-concept, one which did not have such glaring inconsistencies and one with which I did not have as many concerns. So many questionable control scores made me doubt the appropriateness of the SDQ-I with this group of students. I am not sure whether the large number of questionable control scores was due to the instrument or to the group of participants. There also seemed to be many inconsistencies both within the structure of the instrument (wording, ordering, and item content) and within the manual that accompanied it (the model, scoring information,

assumptions).

It might have been beneficial to use two instruments which explored self-concept. While this may have lead to fatigue on the part of the participants, if at least one of the instruments was shorter and the administration of both was carefully planned, the comparison of the two instruments could have been useful. Another possibility would have been to use a self-assessment instrument and an inferred self-concept instrument completed by parents or teachers. The comparison of the two types of instruments could have shed light on the interpretation of what happened with the participants and with the investigation the validity of the instruments with this group.

I would have made much more extensive use of interviews with much less emphasis, if any, on observation and on document analysis. I would still include the collaboration meetings, but focus groups would be added. The purpose of these focus group interviews would be strictly data gathering, rather than compromising both data and the benefits of the collaboration groups by trying to combine them. I would conduct individual interviews at various points throughout the study. There would be interviews coinciding with the three administrations of instrument as well as less formal, but directed interviews corresponding to critical incidents (for example, when Ryan left the program). The formal interviews would not last any longer than in the present study, but I would be prepared to extend a single interview over more than one session. This could combat fatigue with the interviewee as well as allow the opportunity for follow-up or more in-depth coverage of

topics when necessary.

As the observation and document analysis were not very useful in this study, I would drop or reduce the use of both in the analysis — that is, include only those incidents which were particularly pertinent to self-concept. If observation were to be included, I would include observation in the classroom setting starting from before the CALP began. This would show the child within a less constricted setting with more interaction with others in a less proscribed role. It would also give an indication of how the child behaved and interacted both before the CALP program had any effect and once it was underway. I would follow up seemingly useful observations with an informal interview with the child that was observed to try to understand the child's perspective and to gain more in-depth understanding of what transpired.

Some of these elements discussed were present in earlier versions of the proposal for this study, but were deleted due to time and practicality concerns. Reflecting back, I would have made different decisions about which elements to retain and which to discard when revising my proposal.

I found it fascinating to explore the world of controversy surrounding self-esteem and self-concept. My curiosity is piqued by the lack of a unified, accepted and universally used definition, by the differences in views of the importance of enhancing self-concept and by the whole conundrum of the causes, effects and covariates of self-concept. I believe it would be quite intriguing and beneficial to the field to delve into these issues further.

Elements of the SDQ-I have been studied to a large degree, but there are several aspects that would apparently benefit from further investigation. These include the composition of the scale items, the control scores and the accuracy and precision of the manual accompanying the instrument.

I feel that it is also essential to develop instruments that measure self-concept from a behavioural perspective. That is, instruments that can be completed by others about a person's self-concept. There are a few such instruments (Purkey, 1973). I think that it is important to explore the inferred self-concepts of people as well as each person's own assessment of his or her self-concept.

While this study certainly did not shed any great light on the question of self-concept and cross-age teaching and learning, it may have expanded the possibilities that can be further explored in the future. Thus there is room for optimism and for continuing the exploration begun in the sixties and still continuing today, the work of those such as DeRita & Weaver (1991), Gautry (1990), Gorrell & Keel (1986), Harris (1971), Hedin (1987), Heward, Heron & Cooke (1982), Labbo & Teale (1990), Lane, Pollack & Sher (1972), Maher (1986), Roswal, Mims, Evans, Smith, Young, Burch, Croce, Horvat & Block (1995), Thelen (1967), and Winson (1994).

From a personal growth and learning perspective, I found this study to be a great success. The whole process of taking the project from the seed of an idea, to the germination, and then the growth, harvesting and eating of the final produce was

rewarding and edifying. I learned a great deal about who I was, what I believed and what I could do. I also learned a great deal about the whole spectrum of elements inherent in this research project. I had to be tenacious. I had to learn how little I really knew and how much I still had to learn. I had to be humble. I had to grapple with failed data and the ensuing feelings of failure. I had to confront and transcend those feelings. I had to wrestle with statistics I did not understand and explanations and resources I understood even less. I had the opportunity to work with the supportive and truly remarkable staff, parents and students at Hillside School. I learned to deal with and persevere the necessary bureaucracies of human subjects research. I had the privilege of learning from, receiving assistance from and gathering support from a truly remarkable group of talented and inspiring people — I learned that I certainly could do none of this alone, without the help and support of so many.

So, even though the results of this study were less than spectacular and I made many mistakes along the way, I feel enriched by the whole experience.

References

- Adler, P. A., & Adler, P. (1994). Observational techniques. In N. K. Denzin, & Y. S. Lincoln (Eds.), Handbook of Qualitative Research (pp. 377-392). Thousand Oaks, CA: Sage Publications.
- Afifi, A. A., & Clark, V. (1996). Computer-aided multivariate analysis. London: Chapman & Hall.
- Allen, V. L., & Feldman, R. S. (1974). Tutor attribution and attitude as a function of tutee performance. Journal of Applied Social Psychology, 4, 311-320.
- Bar, E., & Raviv, A. (1982). Underachievers as tutors. Journal of Educational Research, 75(3), 139-143.
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The Dark side of self esteem. Psychological Review, 103(1), 5-33.
- Beane, J. A. (1991). Sorting out the self-esteem controversy. Educational Leadership, (September), 25-30.
- Bednar, R. L., Wells, M. G., & Peterson, S. R. (1989). Self-esteem: Paradoxes and innovations in clinical theory and practice. Washington, DC: American Psychological Association.
- Bell, A. (1817). An experiment in education at the male asylum of Madras: Suggesting a system by which a school or family may teach itself under the superintendence of the master or parent. London.
- Bloom, S. (1976). Peer and cross-age tutoring in the schools: An individualized supplement to group instruction. US Department of Health, Education and Welfare.
- Bogdan, R. C., & Biklen, S. K. (1992). Qualitative research for education: An introduction to theory and methods. Needham Heights, MA: Allyn and Bacon.
- Bogdan, R. C., & Biklen, S. K. (1998). Qualitative research for education: An introduction to theory and methods. Needham Heights, MA: Allyn and Bacon.

- Bradley, J., & Sutton, B. (1993). Reframing the Paradigm Debate. Library Quarterly, 63 (4), 405-10.
- Brewer, J., & Hunter, A. (1989). Multimethod research: A synthesis of styles (Sage Library of Social Research No. 175). Newbury Park, CA: SAGE Publications.
- Brown, Mary Jo McGee. (1992). Issues in educational reform: How triangulating qualitative and quantitative evaluation methods can enhance understanding. ED 352 404,
- Burns, R. B. (1979). The self concept in theory, measurement, development and behaviour. New York: Longman Inc.
- Burr, W. R., & Christensen, C. (1992). Undesirable side effects of enhancing self-esteem. Family Relations, 41, 460-464.
- Byrd, D. E. (1990). Peer tutoring with the learning disabled: A critical case. Journal of Educational Research, 84(2 (November/December)), 115-118.
- Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-concept. Journal of Educational Psychology, 78, 474-481.
- Campbell, J. D., & Lavalley, L. F. (1993). Who am I? The role of self-concept confusion in understanding the behaviour of people with low self esteem. In R. F. Baumeister (Ed.), Self-esteem: The puzzle of low self-regard (pp. 3-20). New York: Plenum Press.
- Chen, R. S., & Dunlap, W. P. (1994). A Monte Carlo study on the performance of a corrected formula for the ϵ suggested by Lecoutre. Journal of Educational Studies, 19(2), 119-126.
- Cloward, R. D. (1967). Studies in tutoring. Journal of Experimental Education, 36(1 (Fall)), 14-25.
- Cohen, P. A., Kulik, J. A., & Kulik, C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 19, 237-248.
- Coopersmith, S. (1967). The antecedents of self-esteem (A Series of Books in Psychology). San Francisco: W. H. Freeman and Company.

- Cornell, J. E., Young, D. M., Seaman, S. L., & Kirk, R. E. (1992). Power comparisons of eight test for sphericity in repeated measures designs. Journal of Educational Statistics, 17(3), 233-249.
- Darlington, R. B. (1990). Regression and linear models. New York: McGraw-Hill Publishing Company.
- Delquardi, J. C., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Classwide peer tutoring. Exceptional Children, 52(6), 535-542.
- Demo, D. H. (1985). The measurement of self-esteem: refining our methods. Journal of Personality and Social Psychology, 48(6), 1490-1502.
- Denzin, N. K. (1989). Interpretive interactionism. Newbury Park, CA: Sage Publications.
- Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin, & Y. S. Lincoln (editors), Handbook of qualitative research (pp. 1-18). Thousand Oaks, CA: Sage Publications, Inc.
- DePaulo, B. M., Tang, J., Webb, W., Hoover, C., Marsh, K., & Litowitz, C. (1989). Age differences in reactions to help in a peer tutoring context. Child Development, 60(2), 423-439.
- DeRita, C., & Weaver, S. (1991). Cross-Age Literacy Program. Reading Improvement, 28 (Winter)(4), 244-248.
- Devin-Sheehan, L., Feldman, R. S., & Allen, V. L. (1976). Research on children tutoring children: A critical review. Review of Educational Research, 46(3), 355-385.
- Erikson, M. R., & Cromack, T. (1972). Evaluating a tutoring program. Journal of Experimental Education, 41, 27-31.
- Feldman, R. S., & Allen, V. L. (1979). Student success and tutor verbal and nonverbal behavior. Journal of Educational Research, 72, 142-148.
- Fine, M. (1994). Working the hyphens: Reinventing self and other in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research. Thousand Oaks, CA: Sage Publications, Inc.
- Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. Educational Researcher, 16(7), 16-21.

- Fitts, W. H. (1972). The self-concept and behaviour: overview and supplement. Nashville, TN: The Dede Wallace Centre.
- Fleming, C. J. (1969). Pupil tutors and tutees learn together. Today's Education, 58(October), 22-24.
- Forgarty, J. L., & Wang, M. C. (1982). An investigation of the cross age peer tutoring process: Some implications for instructional design and motivation. Elementary School Journal, 82, 451-469.
- Franklin, G., Griffin, R., & Perry, N. (1994-1995). Effects of cooperative tutoring on academic performance. Journal of Educational Technology Systems, 23(1), 13-25.
- Fresko, B. (1988). Reward salience, assessment of success, and critical attitudes among tutors. Journal of Educational Research, 81, 341-346.
- Fullen, M. G., & Miles, M. B. (1992). Getting reform right: What works and what doesn't. Phi Delta Kappan, 745-752.
- Gage, N. L. (1989). The paradigm wars and their aftermath: a "historical" sketch of research on teaching since 1989. Educational Researcher, 18(7), 4-10.
- Garzarelli, P., Everhart, B., & Lester, D. (1993). Self-concept and academic performance in gifted and academically weak students. Adolescence, 28(109), 235-37.
- Gautrey, F. (1990). Cross-Age Tutoring in Frankley. Reading, 24 (April)(1), 21-27.
- Gecas, V. (1982). The self-concept. Annual Review of Sociology, 8, 1-33.
- Gerber, M. M., & Kauffman, J. M. (1981). Peer Tutoring in academic settings. In P. S. Strain (Editor), The utilization of classroom peers as behavior change agents (pp. 155-187). New York, NY: Plenum Press.
- Girden, E. R. (1992). ANOVA Repeated Measures (Quantitative Applications in the Social Sciences: A Sage University Papers Series). Newbury Park, CA: Sage Publications.
- Gold, R. L. (1958). Roles in sociological field observations. Social Forces, 36, 217-223.
- Goldstein, H., English, K., Shafer, K., & Kaczmarek, L. (1997). Interaction among preschoolers with and without disabilities: effects of across-the-day peer intervention. Journal of Speech Language & Hearing Research, 40(1), 33-48.

- Good, C. V. (1973). Dictionary of education. New York: McGraw-Hill.
- Gorrell, J., & Keel, L. (1986). A field study of helping relationships in a cross-age tutoring program. Elementary School Guidance and Counselling, 20(4), 268-276.
- Green, J. C. (1994). Qualitative program evaluation: Practice and promise. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research (pp. 516-544). Thousand Oaks, CA: Sage Publications, Inc.
- Greenwood, C. R., Dinwiddle, G., Terry, B., Wade, L., Stanley, S. O., Thibadeau, S., & Delquardi, J. C. (1984). Teacher-versus-peer-mediated instruction: An ecobehavioural analysis of achievement outcomes. Journal of Applied Behavior Analysis, 17(4), 521-538.
- Guba, E. G. (1990). The paradigm dialog. Newbury Park, CA: Sage.
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research. Thousand Oaks, CA: Sage Publications, Inc.
- Haluska, R., & Gillen, D. (1995). Kids teaching kids. Pairing up with cross-grades pals. Learning, 24(3), 54-56.
- Harris, M. M. (1971). Learning by tutoring others. Today's Education, 60, 48-49.
- Hedin, D. (1987). Students as teachers: Tools for improving school climate and productivity. Social Policy, Winter, 42-47.
- Heward, W. L., Heron, T. E., & Cooke, N. L. (1982). Tutor huddle: Key element in a classwide peer tutoring system. Elementary School Journal, 83(2), 115-123.
- Holly, William J. (1987). Student self-esteem and academic success. Prepared by the ERIC Clearinghouse on Educational Management, University of Oregon, Eugene. (ERIC Document Reproduction Service No. ED 287205)
- Howell, D. C. (1985). Fundamental Statistics for the Behavioural Sciences. Boston, MA: Duxbury Press.
- Howell, D. C. (1997). Statistical methods for psychology. Belmont, CA: Wadsworth Publishing Company.

- Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measures designs have exact F-distributions. Journal of the American Statistical Association, 65, 1582-1589.
- Jacob, E. (1987). Qualitative research methods: A review. Review of Educational Research, 57, 1-50.
- Janesick, V. J. (1994). The dance of qualitative research design: Metaphor, methodolatry and meaning. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research (pp. 209-219). Thousand Oaks, CA: Sage Publications, Inc.
- Jenkins, J. R., Jewell, M., Leicester, N., Jenkins, L. M., & Troutner, N. M. (1991). Development of a School Building Model for Educating Students with Handicaps and At-Risk Students in General Education Classrooms. Journal of Learning Disabilities, 24(5), 311-320.
- Jenkins, J. R., Mayhall, W. F., Peschka, C. M., & Jenkins, L. M. (1974). Comparing small group and tutorial instruction in resource rooms. Exceptional Children, 40(4), 244-250.
- Jensen, D. R. (1987). Topics in the analysis of repeated measurements. In H. Bozdogan, & Gupta A. K. (Eds), Multivariate statistical modeling and data analysis (pp. 147-161). New York: D. Reidel.
- Johnson, A. G. (1995). The Blackwell dictionary of sociology: Auser's guide to sociological language. Cambridge, MA: Blackwell Publishers Inc.
- Johnson, M., & Bailey, J. S. (1974). Cross-age tutoring: Fifth graders as arithmetic tutors for kindergarten children. Journal of Applied Behavioural Analysis, 7(2 (Summer)), 223-232.
- Katz, Lilian G. (1993). Distinctions between self-esteem and narcissism: Implications for practice. ERIC Clearinghouse on Elementary and Early Childhood Education, Urbana, Ill. [BBB16656]. (ERIC Document Reproduction Service No. ED 363 452)
- Keppel, G. (1982). Design and analysis: A researcher's handbook. Englewood Cliffs, NJ: Prentice-Hall, Inc.

- Kirk, J., & Miller, M. L. (1986). Reliability and validity in qualitative research. Newbury Park, CA: Sage Publications.
- Kohn, A. (1994). The truth about self-esteem. Phi Delta Kappan, 76(4), 272-283.
- Kuhn, T. S. (1970). The structure of scientific revolutions (International Encyclopedia of Unified Science No. 2, Number 2). Chicago, IL: The University of Chicago Press.
- Labbo, L. D., & Teale, W. H. (1990). Cross-age reading: a strategy for helping poor readers. Reading Teacher, (February), 362-369.
- Lancaster, J. (1806). Improvements in education. London: Collins & Perkins.
- Lane, P., Pollack, C., & Sher, N. (1972). Remotivation of disruptive adolescents. Journal of Reading, 15, 351-354.
- Lazerson, D. B., Foster, H. L., Brown, S. I., & Hummel, J. W. (1988). The effectiveness of Cross-Age Tutoring with Truant, Junior High School Students with Learning Disabilities. Journal of Learning Disabilities, 21(4), 253-255.
- LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. Review of Educational Research, 52(1), 31-60.
- Leming, J. S. (1981). Curricular effectiveness in value/moral education. Journal of Moral Education, 10, 147-164.
- Lindquist, E. F. (1953). Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin.
- Lockwood, A. L. (1978). The effects of value clarification and moral development curricula on school-aged subjects: A critical review of research. Review of Educational Research, 48, 325-364.
- Maher, C. A. (1982). Behavioral effects of using conduct problem adolescents as cross-age tutors. Psychology in the Schools, 19, 360-364.
- Maher, C. A. (1986). Direct replication of a cross-age tutoring program involving handicapped adolescents and children. School Psychology Review, 15(1), 100-118.

- Maher, C. A. (1984). Handicapped adolescents as cross-age tutors: Program description and evaluation. Exceptional Children, 51(1), 56-63.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. Journal of Educational Psychology, 79, 280-295.
- Marsh, H. W. (1990). Causal ordering of academic self-concept and achievement: A multiwave, longitudinal panel analysis. Journal of Educational Psychology, 82, 646-656.
- Marsh, H. W. (1990). Self Description Questionnaire - I: SDQ-I manual. Campbelltown, Australia: Herbert W. Marsh.
- Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). The Self-Description Questionnaire (SDQ): Age effects in the structure and level of self-concept for peradolescent children. Journal of Educational Psychology, 76, 940-956.
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. Journal of Educational Psychology, 80(3), 366-380.
- Marsh, H. W., Craven, R. G., & Debus, R. L. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. Journal of Educational Psychology, 83(3), 377-392.
- Marsh, H. W., & Hogevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. Psychological Bulletin, 97, 562-582.
- Marsh, H. W., & Shavelson, R. J. (1985). Self concept: Its multifaceted, hierarchical structure. Educational Psychologist, 20, 107-123.
- Marsh, H. W., & Smith, I. D. (1987). A cross-national study of the structure and level of multidimensional self-concepts: An application of confirmatory factor analysis. Australian Journal of Psychology, 39, 61-77.
- Marston, D., Deno, S. L., Kim, D., Diment, K., & Rogers, D. (1995). Comparison of reading intervention approaches for students with mild disabilities. Exceptional Children, 62(1), 20-37.

- Martella, R. C., Marchand-Martella, N.E., Young, K. R. & MacFarlane, C. A. . (1995). Determining the collateral effects of peer tutor training on a student with severe disabilities. Behavior Modification, 19(2), 170-191.
- Martino, L. R. (1994). Peer tutoring classes for young adolescents: a cost-effective strategy. Middle School Journal, 25(4), 55-58.
- Mauchley, J. W. (1940). Significance test for sphericity of n -variate normal populations. Annals of Mathematical Science, 11, 204-209.
- May, R. B., Masson, M. E. J., & Hunter, M. A. (1990). Application of Statistics in Behavioral Research. New York: Harper & Row, Publishers, Inc.
- McBride, P. (1995). Tutor training activities to improve student self-reliance. Paper presented at the Annual International Conference of the National Institute for Staff and Organizational Development on Teaching Excellence and Conference of Administrators (17th, Austin, TX, May 21-24, 1995). (ERIC Document Reproduction Service No. ED 390 487)
- Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: guided learning by doing. Cognition & Instruction, 13(3), 315-372.
- Miles, M. B. (1990). New methods for qualitative data collection and analysis: Vignettes and pre-structured cases. Qualitative Studies in Education, 3(1), 37-51.
- Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded source book. Thousand Oaks, CA: Sage Publications, Inc.
- Morgan, R. F., & Toy, T. B. (1970). Learning by teaching: A student-to-student compensatory tutoring program in a rural school system and its relevance to the educational cooperative. Psychological Record, 20, 159-169.
- Morris, H. H. (1981). Experimental considerations in the analysis of repeated measures. in R. H. Cox, & R. C. Serfs (Editors), Analysis of repeated measures and identifying causal relationships in physical education research (pp. 25-44). Reston, Virginia: American Alliance for Health, Physical Education, Recreation and Dance.
- Neuman, S. G. (1992). The negative consequences of the self-esteem movement. The Alberta Journal of Educational Research, 38(4), 251-253.

- Niedermeyer, F. C., & Ellis, P. A. (1971). Remedial reading instruction by trained pupil tutors. Elementary School Journal, 71, 400-405.
- Norton, D. W. (1952). An empirical investigation of some effects of non-normality and heterogeneity of the F -distribution. Unpublished doctoral dissertation, State University of Iowa.
- Osguthorpe, R. T., & Scruggs, T. E. (1986). Special education students as tutors: A Review and Analysis. Remedial and Special Education, 7(4), 15-26.
- Paolitto, D. P. (1976). The effect of cross-age tutoring on adolescence: An inquiry into theoretical assumptions. Review of Educational Research, 46, 215-237.
- Patton, M. Q. (1990). Qualitative evaluation and research methods. Newbury Park, CA: Sage Publications.
- Punch, M. (1994). Politics and ethics in qualitative research. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research (pp. 83-98). Thousand Oaks, CA: Sage Publications, Inc.
- Purkey, W. W. (1994). personal communication.
- Purkey, W. W., Cage, B., & Graves, W. (1973). The Florida Key: A scale to infer learner self concept. Journal of Educational and Psychological Measurement, 33, 979-984.
- Ramsey, P. H. (1980). Exact type I error rates for robustness of Student's t test with unequal variances. Journal of Educational Statistics, (5), 337-349.
- Rosenberg, M. (1979). Conceiving the self. New York, NY: Basic Books.
- Roswal, G. M., Mims, A. A., Evans, M. D., Smith, B., Young, M., Burch, M., Croce, R., Horvat, M. A., & Block, M. (1995). Effects of collaborative peer tutoring on urban seventh graders. Journal of Educational Research, 88(5), 275-279.
- Scheff, T. J., Retzinger, S. M., & Ryan, M. T. (1989). Crime, violence, and self-esteem. In A. M. Mecca, N. J. Smelser, & J. Vasconcellos (Eds.), The social self-importance of self-esteem (pp. 165-199). Berkeley: University of California Press.
- Schrader, B., & Valus, A. (1990). Disabled learners as able teachers: A cross-age tutoring project. Academic Therapy, 25(5), 589-597.

- Schwandt, T. A. (1994). Constructivist, Interpretivist Approaches to Human Inquiry. In N. K. Denzin, & Y. S. Lincoln (Editors), Handbook of qualitative research (pp. 118-137). Thousand Oaks, CA: Sage Publications, Inc.
- Scruggs, T. E., & Osguthorpe, R. T. (1986). Tutoring interventions within special education settings: A comparison of cross-age and peer tutoring. Psychology in the Schools, 23(2), 187-193.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. Review of Educational Research, 46(3 (Summer)), 407-441.
- Shavelson, R. J., & Marsh, H. W. (1986). On the structure of self-concept. In R. Schwarzer (Ed.), Anxiety and cognitions. Hillsdale, NJ: Lawrence Erlbaum.
- Smelser, N. J. (1989). Self-esteem and Social Problems: An introduction. In A. M. Mecca, N. J. Smelser, & J. Vasconcellos (Eds.), The social importance of self-esteem (pp. 1-23). Berkley, CA: University of California Press.
- Strauss, A., & Corbin, J. (1990). Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, CA: SAGE Publications, Inc.
- Sutton, B. (1993). The Rationale for Qualitative Research: A Review of Principles and Theoretical Foundations. Library Quarterly, 63 (4), 411-30 .
- Tabachnick, B. G., & Fidell, L. S. (1983). Using multivariate statistics. New York: Harper & Row, Publishers, Inc.
- Tesch, R. (1990). Qualitative research: Analysis types and software tools. New York: Falmer.
- Thelen, H. A. (1967). The humane person defined. St. Louis, MI: Paper presented at the Secondary Education Leadership Conference.
- Thompson, D. (1996). The Oxford Dictionary of Current English. New York: Oxford University Press Inc.
- Vasconcellos, J. (1989). Preface. In A. M. Mecca, N. J. Smelser, & J. Vasconcellos (Eds.), The social self-importance of self-esteem (p. xi-xxi). Berkeley: University of California Press.
- Weber, P. (1968). Elements of statistics for market research. Worchester, England: Ebenezer Baylis and Son, Ltd.

- Weinstein, G., & Fantini, M. D. (1970). Toward humanistic education: A curriculum of affect. New York: Praeger.
- Wells, L. E., & Marwell, G. (1976). Self-esteem: Its conceptualization and measurement. Beverly Hills, CA: Sage Publications.
- Whitt, E. J. (1991). Artful Science: A Primer on Qualitative Research Methods. Journal of College Student Development, 32 (5), 406-15 .
- Winson, A. (1994). Study buddies. Project seed. Prepared at the Lyseth Elementary School, Portland, Maine. (ERIC Document Reproduction Service No. ED 376 985)
- Wolcott, H. F. (1992). Posturing in qualitative inquiry. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), The handbook of qualitative research in education (pp. 3-52). New York: Academic Press.
- Wylie, R. C. (1989). Measures of Self-concept. Lincoln, NB: University of Nebraska Press.
- Wylie, R. C. (1968). The present status of self theory. In E. F. Borgatta, & W. W. Lambert (Eds.), Handbook of personality theory and research (pp. 728-787). Chicago: Rand MacNally.
- Wylie, R. C. (1974). The self-concept: A review of methodological considerations and measuring instruments Vol. One). Lincoln, NB: University of Nebraska Press.
- Zumwalt, J. (1994). Exploring peer tutor differences and biases. Paper presented at the Annual Meeting of the Conference on College Composition and Communication (45th, Nashville, TN, March 16-19, 1994). (ERIC Document Reproduction Service No. ED 370 099)

Appendix A — Time Line for CALP Implementation

Time Period	A			B	
Phase	1	2	3	4	
Associate orientation	*				
Selection and consent	**				
Self-concept measure (SDQ-I)	*		*	*	
Observation	*****	*****	*****	*****	*****
Training of older participants (Learning Partners)		*****			
(Non Learning Partners)					
CALP sessions					
(Learning Partners)			*****	*****	*****
(Non Learning Partners)					*****
Participant discussion groups			* * *	* * *	
Post-program interviews — participants, teachers					**

A - January to beginning of Spring Break

B - End of Spring Break to beginning of June.

Appendix B — SDQ-I

SELF-DESCRIPTION QUESTIONNAIRE - I



Your Name: _____

Circle one: Boy Girl

School: _____

Grade: _____ Age: _____

Teacher: _____

Date: _____

This is a chance to look at yourself. It is not a test. There are no right answers, and everyone will have different answers. Be sure that your answers show how you feel about yourself. **PLEASE DO NOT TALK ABOUT YOUR ANSWERS WITH ANYONE ELSE.** We will keep your answers private and not show them to anyone.

When you are ready to begin, please read each sentence and choose an answer. (You may read quietly to yourself as I read aloud.) There are five possible answers for each question: "True," "False," and three answers in between. There are five boxes next to each sentence, one for each of the answers. The answers are written at the top of the boxes. Choose your answer to a sentence and make a check mark in the box under the answer you choose. **DO NOT** say your answer out loud or talk about it with anyone else.

Before you start, there are three examples below. A student, Bob, has already answered two of these sentences to show you how to do it. In the third example you must choose your own answer and put in your own check mark.

EXAMPLES

FALSE MOSTLY FALSE **SOME-
TIMES
FALSE/
SOME-
TIMES
TRUE** MOSTLY TRUE TRUE

1. I like to read comic books 1 ☐ ☐ ☐ ☐ ☒ 1

Bob checked the box under the answer "True." This means that he really likes to read comic books. If Bob did not like to read comic books very much, he would have answered "FALSE" or "MOSTLY FALSE."

2. In general, I am neat and tidy 2 ☐ ☐ ☒ ☐ ☐ 2

Bob answered "SOMETIMES FALSE, SOMETIMES TRUE," because he is not very neat, but he is not very messy either.

3. I like to watch TV. 3 ☐ ☐ ☐ ☐ ☐ 3

For this sentence you have to choose the answer that is best for you. First you must decide if the sentence is "TRUE," or "FALSE," or somewhere in between. If you really like to watch TV, a lot, you would answer "TRUE" by making a check mark in the last box. If you hate watching TV, you would answer "FALSE" by making a check mark in the first box. If your answer is somewhere in between, then you would choose one of the other three boxes.

If you want to change an answer you have marked, you should cross out the check mark and put a new check mark in another box on the same line.

For all the sentences be sure that your check mark is on the same line as the sentence you are answering. You should have one answer and only one answer for each sentence. Do not leave out any of the sentences. Once you have started, **PLEASE DO NOT TALK.** Turn over the page and begin.

		FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE	
1. I am good looking	1	☐	☐	☐	☐	☐	1
2. I'm good at all SCHOOL SUBJECTS	2	☐	☐	☐	☐	☐	2
3. I can run fast	3	☐	☐	☐	☐	☐	3
4. I get good marks in READING	4	☐	☐	☐	☐	☐	4
5. My parents understand me	5	☐	☐	☐	☐	☐	5
6. I hate MATHEMATICS	6	☐	☐	☐	☐	☐	6
7. I have lots of friends	7	☐	☐	☐	☐	☐	7
8. I like the way I look	8	☐	☐	☐	☐	☐	8
9. I enjoy doing work in all SCHOOL SUBJECTS	9	☐	☐	☐	☐	☐	9
10. I like to run and play hard	10	☐	☐	☐	☐	☐	10
11. I like READING	11	☐	☐	☐	☐	☐	11
12. My parents are usually unhappy or disappointed with what I do	12	☐	☐	☐	☐	☐	12
13. Work in mathematics is easy for me	13	☐	☐	☐	☐	☐	13

		FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE	
14. I make friends easily	14	☐	☐	☐	☐	☐	14
15. I have a pleasant looking face	15	☐	☐	☐	☐	☐	15
16. I get good marks in all SCHOOL SUBJECTS	16	☐	☐	☐	☐	☐	16
17. I hate sports and games	17	☐	☐	☐	☐	☐	17
18. I'm good at READING	18	☐	☐	☐	☐	☐	18
19. I like my parents	19	☐	☐	☐	☐	☐	19
20. I look forward to MATHEMATICS	20	☐	☐	☐	☐	☐	20
21. Most kids have more friends than I do	21	☐	☐	☐	☐	☐	21
22. I am a nice looking person	22	☐	☐	☐	☐	☐	22
23. I hate all SCHOOL SUBJECTS	23	☐	☐	☐	☐	☐	23
24. I enjoy sports and games	24	☐	☐	☐	☐	☐	24
25. I am interested in READING	25	☐	☐	☐	☐	☐	25
26. My parents like me	26	☐	☐	☐	☐	☐	26

		FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE	
27. I get good marks in MATHEMATICS	27	☐	☐	☐	☐	☐	27
28. I get along with kids easily	28	☐	☐	☐	☐	☐	28
29. I do lots of important things	29	☐	☐	☐	☐	☐	29
30. I am ugly	30	☐	☐	☐	☐	☐	30
31. I learn things quickly in all SCHOOL SUBJECTS ..	31	☐	☐	☐	☐	☐	31
32. I have good muscles	32	☐	☐	☐	☐	☐	32
33. I am dumb at reading	33	☐	☐	☐	☐	☐	33
34. If I have children of my own, I want to bring them up like my parents raised me	34	☐	☐	☐	☐	☐	34
35. I am interested in MATHEMATICS	35	☐	☐	☐	☐	☐	35
36. I am easy to like	36	☐	☐	☐	☐	☐	36
37. Overall, I am no good	37	☐	☐	☐	☐	☐	37
38. Other kids think I am good looking	38	☐	☐	☐	☐	☐	38
39. I am interested in all SCHOOL SUBJECTS	39	☐	☐	☐	☐	☐	39
				SOME- TIMES FALSE/ SOME- TIMES TRUE			
		FALSE	MOSTLY FALSE		MOSTLY TRUE	TRUE	
40. I am good at sports	40	☐	☐	☐	☐	☐	40
41. I enjoy doing work in READING	41	☐	☐	☐	☐	☐	41
42. My parents and I spend a lot of time together	42	☐	☐	☐	☐	☐	42
43. I learn things quickly in MATHEMATICS	43	☐	☐	☐	☐	☐	43
44. Other kids want me to be their friend	44	☐	☐	☐	☐	☐	44
45. In general, I like being the way I am	45	☐	☐	☐	☐	☐	45
46. I have a good looking body	46	☐	☐	☐	☐	☐	46
47. I am dumb in all SCHOOL SUBJECTS	47	☐	☐	☐	☐	☐	47
48. I can run a long way without stopping	48	☐	☐	☐	☐	☐	48
49. Work in READING is easy for me	49	☐	☐	☐	☐	☐	49
50. My parents are easy to talk to	50	☐	☐	☐	☐	☐	50
51. I like MATHEMATICS	51	☐	☐	☐	☐	☐	51
52. I have more friends than most other kids	52	☐	☐	☐	☐	☐	52

		FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE	
53. Overall I have a lot to be proud of	53	☐	☐	☐	☐	☐	53
54. I'm better looking than most of my friends	54	☐	☐	☐	☐	☐	54
55. I look forward to all SCHOOL SUBJECTS	55	☐	☐	☐	☐	☐	55
56. I am a good athlete	56	☐	☐	☐	☐	☐	56
57. I look forward to READING	57	☐	☐	☐	☐	☐	57
58. I get along well with my parents	58	☐	☐	☐	☐	☐	58
59. I'm good at MATHEMATICS	59	☐	☐	☐	☐	☐	59
60. I am popular with kids of my own age	60	☐	☐	☐	☐	☐	60
61. I can't do anything right	61	☐	☐	☐	☐	☐	61
62. I have nice features like nose, and eyes, and hair ..	62	☐	☐	☐	☐	☐	62
63. Work in all SCHOOL SUBJECTS is easy for me ..	63	☐	☐	☐	☐	☐	63
64. I'm good at throwing a ball	64	☐	☐	☐	☐	☐	54

		FALSE	MOSTLY FALSE	SOME- TIMES FALSE/ SOME- TIMES TRUE	MOSTLY TRUE	TRUE	
65. I hate READING	65	☐	☐	☐	☐	☐	65
66. My parents and I have a lot of fun together	66	☐	☐	☐	☐	☐	66
67. I can do things as well as most other people	67	☐	☐	☐	☐	☐	67
68. I enjoy doing work in MATHEMATICS	68	☐	☐	☐	☐	☐	68
69. Most other kids like me	69	☐	☐	☐	☐	☐	69
70. Other people think I am a good person	70	☐	☐	☐	☐	☐	70
71. I like all SCHOOL SUBJECTS	71	☐	☐	☐	☐	☐	71
72. A lot of things about me are good	72	☐	☐	☐	☐	☐	72
73. I learn things quickly in READING	73	☐	☐	☐	☐	☐	73
74. I'm as good as most other people	74	☐	☐	☐	☐	☐	74
75. I am dumb at MATHEMATICS	75	☐	☐	☐	☐	☐	75
76. When I do something, I do it well	76	☐	☐	☐	☐	☐	76

Appendix C — Learning Logs

Date: _____

Older Learning Partner: _____

Younger Learning Partner: _____

OLDER LEARNING PARTNER LOG

Plan for today:

Done

Today's session went



because _____

Other comments:

YOUNGER LEARNING PARTNER LOG

Today's session went



because _____

I really liked:

I really didn't like:

Appendix D — Sample Activities

Card Facts Card Game

Materials: Numbers only card pack

Steps:

- Deal half of the cards (face down) to each learning partner.
- Each player turns over his or her top card.
- The partner with the smallest card adds the two cards together and says:

$\frac{\text{Smaller card}}{\text{Smaller card}} \text{ plus } \frac{\text{Bigger card}}{\text{Bigger card}} \text{ equals } \frac{\text{Total}}{\text{Total}}$

If both numbers are the same, the younger learning partner adds them together.

Or you can subtract the bigger number from the smaller number.

- Each partner should make certain that the other partner has the right answer. One partner may make a mistake or may try to trick the other partner.
- When your partner has said the sentence correctly, put these cards in 'throw away' pile.
- Repeat steps 2, 3 and 4 until all the cards in the deck are used up.

Start		12 How many less that the total on the dice do you need to make 1?	13 +	14 If you added an odd number to the total of the 2 dice, would your answer be odd or
1 +		11 -		15 -
2 Tell if each number on the dice is odd or even. Is the total odd or even?		10 How many more than the total on the dice do you need to make 14?		16 Add the two numbers together and subtract 2.
3 -		9 +		17 +
4 Subtract the smaller number from the bigger number and add 9.		8 Is the total of the two numbers bigger than or less than 7?		18 How many more than what is on the dice do you need to make 17?
5 +	6 Which number on the dice is bigger?	7 -		Finish

The Great Math Race

In the square beside each number:

put a check mark in the first column for the game if the younger learning partner got the answer on the first try
put a check mark in the second column for the game if the older learning partner got the answer on the first try.

Write to total number of check marks for each of you for the game in the last row.

[illegible]

Appendix E — Teacher Information Package

THE EFFECTS OF CROSS AGE LEARNING ON THE SELF-CONCEPT OF PRIMARY AND JUNIOR STUDENTS

January, 1995

Magda Stryk
University of Ottawa
564-5573 (Office)
236-0926 (Home)

TABLE OF CONTENTS

INTRODUCTION	1
DESCRIPTION OF PILOT PROJECT	1
BACKGROUND INFORMATION	1
METHOD	3
Participants	3
Data Collection	3
Quantitative	3
Qualitative	3
Observation	3
Interviews	3
Document Analysis	4
Procedure	4
Phase 1 - Preparation	4
Consent and Confidentiality	4
Testing	4
Grouping	5
Phase 2 - Training	5
Phase 3 - Learning Sessions	6
Phase 4 - Control Group Treatment and Qualitative Data Collection	6
CONCLUDING REMARKS	6
REFERENCES	7

This approach is based on a concept of the school child as capable of participating in a meaningful, dignified way in his [or her] education. He [or she] is seen as capable of learning more than specific content in coursework, but rather as a learner of how to learn. . . . Teaching oneself should be enhanced by learning how to teach others.

Morgan and Toy, 1970

INTRODUCTION

Cross Age Learning Partners (CALP) is a custom designed program for children who are having difficulty, where children in junior grades tutor children in primary grades. It is based on cross age and peer tutoring, in which one student is 'expert' of the content area. *Tutoring* is where a person with a minimum of special training, helps one or more students learn a specific task, under the guidance of a teacher or supervisor.

This study will examine the effects of CALP on the self-concepts of children in primary and in junior grades who are having difficulty. These are children who are capable of learning in a regular environment but who are having difficulty with their academic work and whose academic performance is below the expected level for their grade, in this case, in mathematics. Self-concept is an individual's perceptions of self as a person. This includes abilities, appearance, performance given tasks and other phases of daily living.

DESCRIPTION OF PILOT PROJECT

In the 1992-1993 school year, CALP was initiated, planned and implemented at an elementary school in the National Capital Region. The feedback that was received was positive. Children seemed to improve both in their academic performance and in their self-concepts. Children, teachers, parents and the school administration seemed to like the program and to be pleased with the results. Because this original implementation of the CALP program was a volunteer project and not a study, the data were not systematically collected and were entirely descriptive. These initial results, however, stimulated further exploration of cross age learning in the form of the proposed study and several changes to the initial design of the program were made resulting from experience with the program.

BACKGROUND INFORMATION

Research suggests that both older and younger learning partners benefit from cross-age and peer tutoring programs. An overwhelming majority of studies have supported the concept of cross age tutoring as a worthwhile academic approach.

Studies of the effects of peer and cross age tutoring on self-concept, however, have yielded contradictory evidence; and it has been impossible to determine whether or not growth in self-concept can be anticipated from such programs. This study seeks to explore and begin to reconcile some of these contradictions.

It seems that the inconclusive results regarding growth in self-concept have not necessarily been due to an absence of effect. Instead, it is possible that one of the reasons the results have been indeterminate is due to the use of inappropriate self-concept measures. Until now, the instruments used to measure self-concept in tutoring studies have only looked at general self-concept.

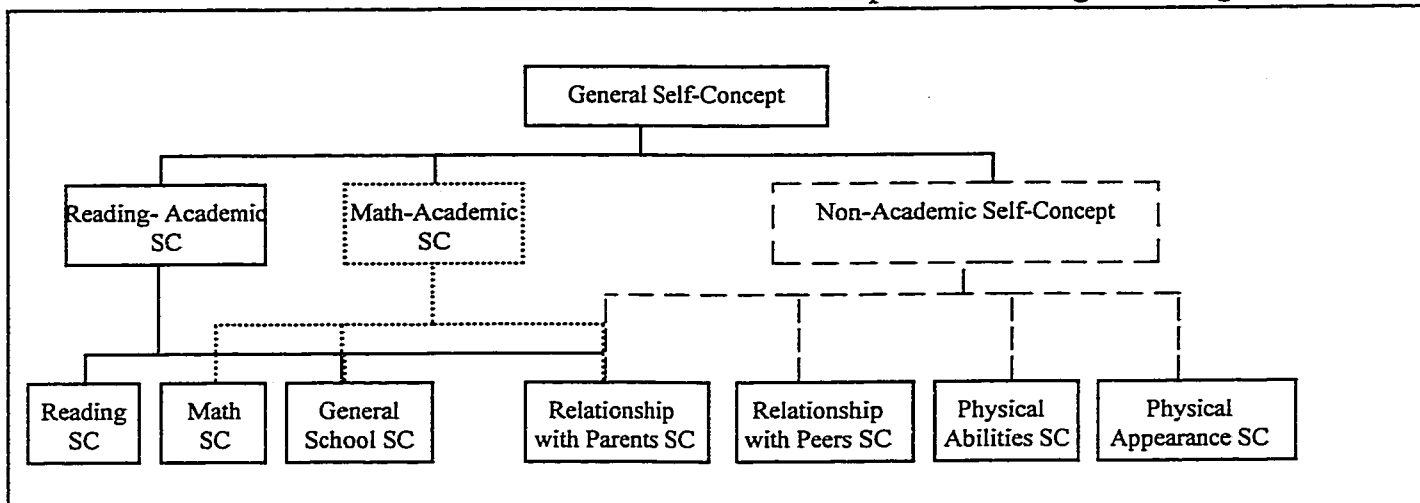


Figure 1 - Marsh and Shavelson's (1985) revised model of Self-Concept

However, according to the model of March and Shavelson (1985), the self-concept is multifaceted and hierarchical. The general self-concept is *stable* but as one moves down the hierarchy, self-concept becomes more susceptible to change because it becomes more situation specific. The general self-concept then, is only one global aspect of a self-concept hierarchy, and is divided into a reading-academic self-concept, a math-academic self-concept and a non-academic self-concept (see Figure 1). Interestingly, reading and math self-concepts (subsets of reading-academic self-concept and math-academic self-concept respectively) have been found to be uncorrelated.

This brings into question measuring differences on a theoretically stable general self-concept or even an academic self-concept, when tutoring occurs in only one subject area. By measuring the self-concept related to the subject area being taught, a more accurate portrayal of the effects of the CALP program on self-concept can be determined. Therefore, the proposed research seeks to answer the question **will students in the CALP program have greater gains in subject specific self-concept than children in the control groups?** (Q_1). Further, **will gains in general self-concept be less differentiated between the learning partners and the control groups than the subject specific self-concept** (Q_2). That is, will the students in the CALP program have greater gains in subject specific self-concept than in the theoretically more stable general self-concept? If significant, will the difference in gains in general self concept between the two groups have a relatively small effect size when compared to the difference in gains in subject specific self-concept?

The qualitative aspects of this research will examine a question of process: **In what ways do the qualitative data help to explain, expand or refute the quantitative findings?** (Q_3). This question may possibly help to resolve some of the discrepancies found in the literature regarding the effect of cross age and peer tutoring on self-concept.

METHOD

This study begins in January 1995, and learning sessions will run from February to the beginning of June, 1995. The participants, the data collection strategies, the procedure and the analysis are outlined below.

Participants

Primary (grades 1 to 3) and junior (grades 4 to 6) children who are identified as having difficulty by their teachers will be invited to participate in this study. Classroom teachers do this regularly in their classrooms, when putting children into groups, when writing report cards or when designing activities. As such, they are in a good position to make the selection decisions in this study. This teacher selection will also prevent the disappointment that could foreseeably occur if all children were tested, but only a few were invited to continue in the research.

Data CollectionQuantitative

The Self-Description Questionnaire (SDQ-I), will be used to examine the self-concept of all children participating in the study. This questionnaire is based on Marsh and Shavelson's Model (1985) and looks at the different factors within the general self-concept. It will be given to the participants 3 times during the study: at the beginning, after four weeks of learning sessions, and just before the control group children begin participating in the CALP program.

Qualitative

In addition to the collection and analysis of the SDQ-I a qualitative research approach will be used to provide the participants' perspectives of the CALP experience and to further understand the underlying reasons for the results obtained, that is, to help understand why what has happened has happened. Three qualitative techniques will be used in this study: observation, interviewing and document analysis.

Observation — I will make observations during the training sessions, the collaboration groups, the learning sessions and testing sessions.

The records of the observations will include specifics regarding who is involved, the interactions, the routines and rituals involved in the interactions and the setting, the social organization, the temporal and spatial characteristics of the interactions, and my interpretations of these.

Interviews — Children's interviews will be used to obtain more specific information on the growth of self-concept and academic achievement, to validate observations, test and classroom results and to build stronger data through a one-on-one approach. Parent and teacher interviews will be used to validate and enhance understanding of student responses, observations, test and classroom results as well as to obtain different perspectives on the effectiveness of CALP.

Interviews, meetings and some sessions will be tape recorded. All interview tapes will be transcribed. Tapes from meetings and sessions will be reviewed, and those that are judged to yield data pertinent to the study will be transcribed; others will be stored, as is, with a detailed contact summary sheet to outline what the tape contains should the information on it be deemed necessary. For tapes that are transcribed, the names of the participants will be changed during transcription and then the tapes will be destroyed.

Document Analysis — Throughout the study, learning logs will be collected for document analysis to ascertain how the participants feel that the learning sessions are unfolding. Classroom work will also be gathered to look for patterns or trends that the learning partners may be exhibiting in their class work.

In all transcriptions and reports all names of participants, teachers, parents, others, the school and the school board will be assigned a pseudonym.

Procedure

Phase 1 - Preparation

The first step in the program was the teacher orientation session during which teachers were familiarized with the program, the study, its aims and the logistics of carrying out this research.

Consent and Confidentiality — A letter of permission has been sent to the parents and guardians of the children that teachers have identified as having some difficulty in mathematics. At any point throughout this study, if a child or his or her parents want that child to withdraw, the child will be free to do so.

Pseudonyms will be given to all participants in the program. No real names will be used in any reports, except in conversation about the project with school staff or in discussion and in correspondence with parents of the participating children, in which case the names or details of other participating children will not be revealed.

Tape recordings that are fully transcribed (using pseudonyms) will be destroyed. Those recordings that are not fully transcribed will be kept private and any summary sheets will use pseudonyms when names are needed. Field, research and other notes generated or transcribed by me will have pseudonyms substituted for real names as soon as possible. Documents that are collected from and about the participants and all other materials with identifiers will be kept in a safe, locked and private location that only I have access to (my home).

Testing — Before any grouping decisions, training or learning sessions occur, all children who agree and who receive parental permission to participate in the study will be administered the Self Description Questionnaire (SDQ-I; Marsh, 1988) as a measure of self-concept. As well, samples of classroom work for these students will be collected.

Four weeks after learning sessions begin, the SDQ-I will be administered (within-program) and sample classroom work will also be collected. Within-program testing is essential because variations sometimes exist between results from testing partway through and at the end of a tutoring program. At the end of quantitative data collection (the end of phase 3), the SDQ-I will be administered and classroom work will be collected once more.

Grouping — Children will be placed into one of four groups: younger learning partners, the YLPs (5-8 children); younger non learning partners, the YNLPs (5-8 children); older learning partners, the OLPs (5-8 children); or older non learning partners, the ONLPs (5-8 children). Children at each grade level will be put into one of the two groups for that level based on matching pretest scores as close as possible. That is, if six children in one class will be participating in this study, each of these six will be matched with the child who has the most similar pretest scores. One of these children will be put into the experimental group and the other child will be put into the control group. A letter will be sent to parents informing them of which group their child belongs to.

Children placed in the experimental groups (YLPs and OLPs) will participate in the Cross Age Learning Partner Program right away. These children will participate in testing (three times), learning sessions and collaboration groups (after the OLPs are trained). Older learning partners will participate in training sessions immediately after the first SDQ-I administration.

Children in the control groups (YNLPs and ONLPs) will participate in the three testing sessions. After the last administration of the SDQ-I, they too will participate in CALP program as described below.

Phase 2 - Training

Junior children in both groups will go through a two-week tutor training program which will incorporate training techniques from successful tutoring studies and programs. Those in the experimental group will be trained immediately following the first testing session and those in the control group will be trained immediately following the last testing session. The training session will include learning to work in a flexible manner with a learning partner, learning to encourage participation from a learning partner, discussions, modelling appropriate behaviour, role playing and simulations, and learning to keep and use a planning and comment log.

Following training, OLPs and YLPs (and later, ONLPs and YNLPs) will be matched, based on grade level. After an orientation meeting for all the learning partners (and later, all members of the control groups) the cross age learning sessions will begin.

In addition to pre-service training, OLPs (and ONLPs) will receive in-service training in the form of collaboration groups. In these group discussions learning partners will discuss ideas, problems, and tutoring in general. Primary students will meet in separate collaboration groups to discuss their tutoring experiences. I will act as facilitator during these discussion.

Phase 3 - Learning Sessions

Tutoring sessions will occur two times a week for thirty minutes in an empty classroom, library or some other appropriate location. Learning partners will sit in pairs and work. I will supervise these sessions.

The cross age learning partners will leave their classrooms during class time to engage in the learning partner sessions. The children in the control groups will not leave the classroom, but will engage in normal classroom activity until after the last testing session. Learning sessions will be held during class time to ensure all children who can benefit from the program will have the opportunity to do so.

Older learning partners will move from teaching fairly structured lessons, designed by you, the classroom teacher, or by me (depending on the your preference), to lessons they design themselves, under my guidance.

Each day, after the session, all learning partners will fill out a learning log. Older partners will help their younger partners to complete their logs.

Phase 4 - Control Group Treatment and Qualitative Data Collection

After the third and final administration of the SDQ-I in April, the qualitative data analysis will continue for the experimental group.

At this time, the junior students in the control group will be trained as Older Learning Partners. Following this training, the students in the control groups will engage in the CALP program. I will supervise these students, but qualitative data analysis will focus on the students in the experimental group.

CONCLUDING REMARKS

In summary, the total requirements from the school and the participants will be approximately a half hour initial meeting with teachers, an additional half hour to one hour with each teacher to discuss student needs and scheduling, a half hour interview with two teachers at the end of May, interviews with two participating children, interviews with two parents, three half hour testing periods for the children, two half hour periods per week out of the classroom for children when they are participating in the CALP for learning sessions, space for the sessions and meetings; use of an overhead projector for some meetings and use of some classroom materials for the sessions.

A special report will be prepared for the school board, the school and any interested participants. This report will be a concise and practically oriented document. A copy of the full thesis will be also available upon request. Journal articles and presentations may also result from the research. Confidentiality will always be assured.

It is hoped that this study will lead to a mutually beneficial relationship between the participants, the school, the school board and myself. It is hoped that the children that participate in the CALP

program will experience increases in self-concept. For the time that the children in the control groups are not participating in the CALP program, they will benefit from having more individual attention due to a smaller class when their classmates are participating in CALP. After the study is concluded, parents of all the participants will be advised of their children's self-concept profiles and the changes of those profiles. They will also receive an information packet with tips for enhancing self-concept.

REFERENCES

- Marsh, H. W. (1988). Self-Description Questionnaire-I. New York, NY: Psychological Corporation.
- Marsh, H. W., & Shavelson, R. J. (1985). Self concept: Its multifaceted, hierarchical structure. Educational Psychologist, 20, 107-123.

Appendix F — Consent Letter

Dear Parents/Guardians:

Hillside School and the Hillside Area School Board⁸ have approved that a Cross Age Learning Partner, or CALP, program be done at Hillside School as a research project.

Younger children will be paired with older children in a peer tutoring format. Specifically, grade 1 to 3 students who are having difficulty in math will be paired with children in grades 4 to 6 who are having difficulty in math. We expect that the CALP will improve both older and younger students' self-concepts in the subjects being tutored.

Half the children in the study will be in CALP right away and half will be in the control group who will start CALP a few months from now. Until then, the children in the control group will do the same testing as the children in the CALP group.

All children in CALP and in the control group will be asked to complete a self-concept measure. Children will do these tests three times over the next few months: before the project begins, four weeks into the project, and several (8 to 10) weeks after the project begins. Classroom work and test marks will also be looked at.

Junior students who are in the study will receive four 30-minute periods of training over two weeks. Those in the CALP group will be trained beginning near the end of January, those in the control group will be trained near the end of April. The primary learning partner will work with the older learning partner in learning sessions that will take place twice weekly for 30 minutes each. The primary children in the CALP group will start their learning sessions in February, those in the control group will start in May. The learning sessions will be supervised by the children's teachers and by myself. When they are taking part in the learning sessions, the students will also meet in groups every two weeks to discuss cross age learning. Some children will also be asked to do informal interviews about the program and two children will be asked to do formal interviews. Some of the group meetings and the formal interviews will be audio taped. Some audiotapes will be transcribed and others will be summarized. **NO REAL STUDENT, PARENT OR TEACHER NAMES WILL EVER BE USED IN ANY REPORTS, DISCUSSIONS, OR OTHER FORMS OF COMMUNICATION.**

Some parents will also be asked to do an informal interview. This interview will ask about the cross age learning program and how you think your child's self-concept and academic achievement have been affected by the program. The interviews will be tape recorded and transcribed, and then the tape will be destroyed.

All information in this study will be kept strictly confidential and will be used for research purposes only. **NO INFORMATION FROM THE STUDY WILL GO INTO SCHOOL RECORDS**, though you may wish to share some of the information you receive about your child with his or her classroom teacher. Participation in this study will have no direct effect on your child's school marks. The names of the participants will not be used in any reports. You will be able to get copies of a report of the results of the study when the study is complete.

Participation in the study is completely voluntary, and your child can withdraw from the study at any time. Only children with written permission to participate in the study will take part.

⁸In this copy of the letter, the names of the school and school board are pseudonyms.

I expect that children in the learning partner program will grow in self-concept and in their academic achievement as a result of participating in this study.

Children in both groups should benefit by participating in CALP program. While they are not in learning sessions, children in the control groups will benefit from participation in this research by:

- 1 Receiving more individual attention in the class during the times that their classmates are in the learning sessions because there will be fewer students in the class.
- 2 You will become better informed about your child's self-concept. You may want to pass this information on to your child's classroom teacher. He or she will then be better informed and thus better prepared to work with your child.
- 3 At the end of the project you will receive an information packet about how to enhance your child's self-concept.
- 4 Your child's school has also committed to establishing some form of cross age or peer learning if results from this study are as positive as I expect them to be.

Eventually, all students may benefit from studies such as this one, if programs such as CALP prove to be beneficial and can be established in schools on a permanent basis. You and your child can participate in helping to get a better understanding of the effects of cross age learning by taking part in this study.

I am a fully qualified elementary school teacher with teaching and research experience. The study is to fulfil part of the requirements for my Master of Arts (Education) degree. If you have any questions about the study, please contact me at the numbers listed below.

Sincerely,

Magda Stryk,
Researcher,
University of Ottawa
564-5573 (University of Ottawa)
236-0926 (home)

Tom Wright,
Principal,
Wakefield School

I have read and I understand the attached letter and

☐ I give permission ☐ I do not give permission for my child, _____
to participate in the Cross Age Learning Partner Study as outlined above.

Signature of parent or guardian

I agree to participate in the study.

Signature of the student

Appendix G — Older Partner Training Handout

- Day 1 - Introduction
 - Introductions around the room
 - Explain CALP OLPs' role within it.
 - Goals of the program.
 - Overview of the program.
 - training
 - learning sessions
 - record keeping
 - collaboration sessions
 - What is an Older Learning Partner?
 - responsibilities
 - Learning Partner Hints - Overview of training
 - problems learners might have and how we can help
 - getting to know your partner
 - planning
 - record keeping
 - showing your partner that you care
 - letting your partner know how he or she is doing
 - making the work as good as possible
- Day 2 - Learning Problems
 - What are some of the problems that learners might have that would make learning math hard? - brainstorming
 - What can we do to help? - brainstorming
 - Getting to know your partner
 - interview your learning partner and fill in the questionnaire - it will help you find out about your learning partner, you should try to refer to these interests when doing activities to keep your partner's attention
 - make sure you know your partner's name, teacher's name and grade, you might also want to find out your partner's birthday and other days and things that may be important to him or her
 - Planning
 - as soon as possible after your lesson each day plan what you will do for the next lesson
 - plan a learning activity with something more fun planned in the middle of the learning session (see the learning log)
 - Record keeping

- Day 3 - Learning Sessions
 - Showing your partner you care
 - brainstorming
 - my tips
 - sit beside your partner, not across from him or her
 - be friendly
 - when you talk to your partner, look right at him or her to show that you are listening, use his or her name
 - remember that your words are only a part of what you are saying - use good body language, smiles, nods, a nice voice
 - don't interrupt your partner, give him or her enough time to finish his or her thoughts
 - give your partner enough time to think about his or her answer
 - Letting your partner know how he or she is doing
 - Role playing
 - praise good work and good effort
 - keep encouraging, even when things are not going very well that is usually when people need the most encouragement, try "Let's try again," "Maybe we should try this a different way," "Let's keep working on this. I'll help you."
 - emphasize what is right not what is wrong, if you are marking something, put a mark beside the right answers, not the wrong, and then write the total number of right answers at the top
 - repeat some of your learning partners' correct answers, not the wrong answers
 - praise the right things that your partner does
 - use awards and congratulations sheets when your partner has had a really good session or week
- Day 4 - Learning Sessions Continued
 - Making the work as good as possible
 - Role playing
 - if you and your partner are working on a long job, break the work into smaller tasks, when each task is finished, let your partner know that he or she has finished the job and praise the work or at least give an encouraging remark
 - help your partner help him or her self, help to discover how to find answers, do not just give the answers
 - ask questions that will lead your partner towards the answer
 - if your partner gives the wrong answer, ask the question again, in a different way, do not repeat the wrong answer
 - ask thinking questions, not only questions that ask for facts, try, "Why do you think this happened?" or "Was that the best decision?"
 - help your partner to check his or her work before it is handed in to his or her teacher
 - have your partner evaluate how well she or he has worked and how good his or her work

Appendix H — Getting to Know You Sheet

What is your name?	Age?
What is your teacher's name?	Grade?
When is your birthday?	
How many brother and sisters do you have? How old are they?	
Do you have any pets? What kind? What are their names?	
What is your favourite thing to do?	
What is your favourite thing to do in school?	
What is your favourite subject in school?	
How do you feel about school?	
Do you play any sports?	
What is your favourite TV show? Movie?	
How do you feel about older kids?	
What do you most want to learn about math?	
What do you most want to get out of this learning partner program?	
Is there anything else you want to tell me?	

Appendix I — Permission to Use Figures

Aug 06 08:41:00 1998

To: Magda Therrien

From: 'h.marsh@uws.edu.au' (INTERNET)

Subject: Re: Request for permission to use figures.

Attached: 1) UNIX File: ORIGINAL.HEADER - 1360 bytes

I'm happy for you to reproduce the materials listed below for purposes of your thesis. I would appreciate it if you could include a note indicating the SDQ items cannot be reproduced in any form without permission of the author (i.e., me).

Your thesis sounds interesting. Could you send me a summary. HERB

At 06:26 PM 7/30/98 EDT, you wrote:

>Dear Dr. Marsh,

>

>For the quantitative portions of my thesis, entitled 'Cross age learning in primary and junior grades and the self-concept,' I have used the SDQ-I, which I purchased from you in 1995. I would like to use figures from your SDQ-I manual and from the article that you wrote with Dr. Shavelson in 1985 in my thesis and I am requesting written permission to do so.

>

>The figures have been adapted and a copy of the adaptations is enclosed. The figures in question are as follows:

>Figure	Name	Page Number
---------	------	-------------

>

>>From the Self Description Manual I

>

>Figure 2	Physical Abilities Scale (an adaptation)	5
>Figure 3	Physical Appearance Scale (an adaptation)	5
>Figure 4	Peer Relations Scale (an adaptation)	5
>Figure 5	Parent Relations Scale (an adaptation)	6
>Figure 6	Reading Scale (an adaptation)	6
>Figure 7	Mathematics Scale (an adaptation)	6
>Figure 8	General-School Scale (an adaptation)	7
>Figure 9	General-Self Scale (an adaptation)	7

>

>And a copy of the SDQ-I for an appendix.

>

>>From Self-concept: Its multifaceted, hierarchical structure.

>

>Figure 3	Higher order structures of self-concept in Models 4-7 (an adaptation)	114
-----------	---	-----

>

>I would appreciate receiving permission to use these figures in my thesis.

>

>Thank you in advance for your time and prompt attention.

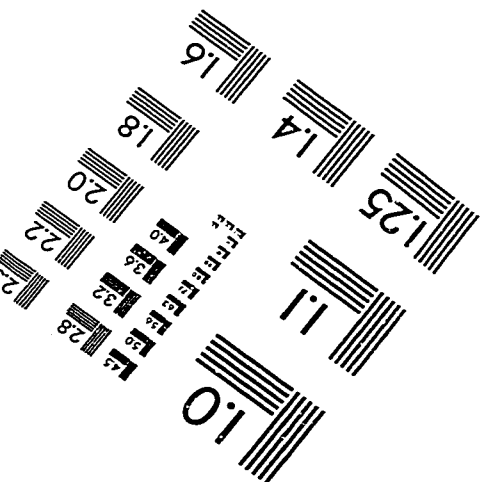
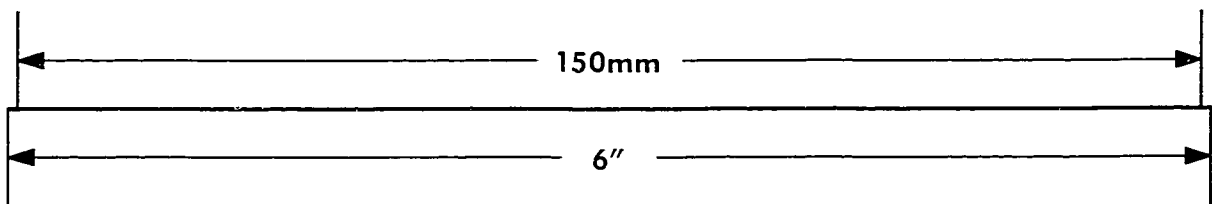
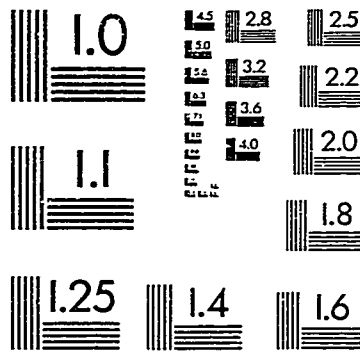
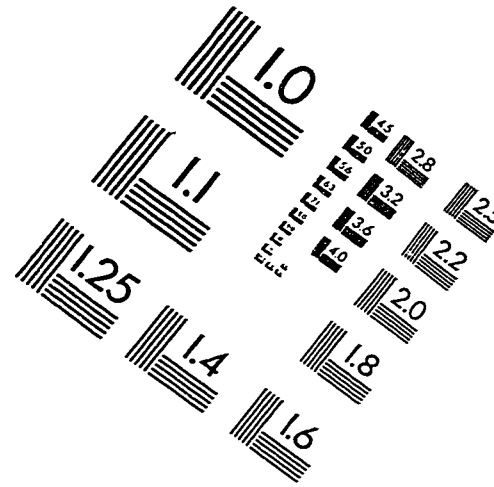
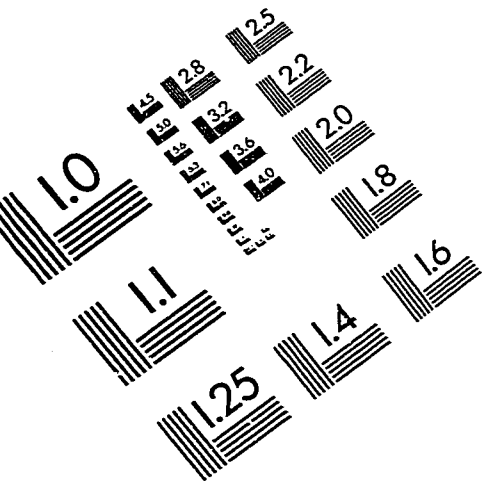
>

>Sincerely,

>

>Magda Stryk Therrien

IMAGE EVALUATION TEST TARGET (QA-3)



APPLIED IMAGE, Inc.
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved

