

**A LONGITUDINAL STUDY OF THE EFFECTS OF IN-SERVICE PROFESSIONAL
DEVELOPMENT ON PRIMARY STUDENTS' LITERACY SUCCESS**

by

Mary Ann Collins-Williams

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Human Development and Applied Psychology
Ontario Institute for Studies in Education of the
University of Toronto

©Copyright by Mary Ann Collins-Williams (2001)



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 *Voire 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.

0-612-58986-2

Canada

**A LONGITUDINAL STUDY OF THE EFFECTS OF IN-SERVICE PROFESSIONAL
DEVELOPMENT ON PRIMARY STUDENTS' LITERACY SUCCESS**

Mary Ann Collins-Williams

Doctor of Philosophy (2001)

Department of Human Development and Applied Psychology

Ontario Institute for Studies in Education of the University of Toronto

ABSTRACT

The purpose of this study was to evaluate a professional development (PD) system that combined the key components of effective literacy instruction with the key features for promoting transfer of skills to the classroom. There has been very little empirical research on the impact of teacher education – both pre-service and in-service – on teacher effectiveness. This evaluation of the impact of in-service PD on student literacy outcomes was undertaken in a four-year longitudinal study with three main parts: (1) the first involved tracking teachers to determine whether they became better teachers of beginning reading and writing after their year-long participation in the PD; (2) the second involved tracking students who were consistently in the classrooms of teachers who participated in the PD to determine whether they showed cumulative benefits in the form of better literacy outcomes after being in the classrooms of in-serviced teachers; (3) the third involved an examination of the literacy outcomes of “at risk” students to determine whether classroom teachers’ participation in PD reduced the proportion of students in their classrooms requiring special education in reading and writing. The major outcomes were as follows: First, based on the analyses that tracked in-serviced teachers, the findings indicate that PD was effective in improving the instructional skills of participating teachers, with the associated result of raising the literacy achievement levels for students in their classrooms. Second, based on the analyses that tracked children, there were cumulative effects

to the literacy skills of Grades 1 and 2 children. Children with more exposure to the instruction of in-serviced teachers were more successful in reading and writing than those with less exposure. Third, results indicated that participation of teachers in PD can contribute to a substantial reduction in the numbers of children who are "at risk" of failing age appropriate levels of achievement in reading and spelling. These results suggest that through participation in PD teachers can significantly improve their skills and this is associated with improvement in the acquisition of literacy skills of their students.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Dale Willows, for her support, encouragement, and invaluable expertise. Her unwavering commitment, thoughtfulness and confidence in me sustained me throughout the journey of this thesis.

I am also grateful to the members of my thesis committee: to Dr. Esther Geva for her guidance; to Dr. Tom Humphries for his thoughtful ideas; to Dr. Phil Nagy for his statistical expertise; and to Dr. Ruth Scott who, as the external examiner, provided helpful comments for the final thesis. I am also indebted to Alisa Kenny for her unfailing assistance with the statistical analyses. I wish to express my gratitude to the school administrators and teachers who participated in this study.

My friend Debbie was a great comfort as we shared the highs and lows associated with the challenges of writing a dissertation.

I deeply appreciate the loving support and encouragement from my family. My mother, who believed in me, as always. My sister, Diane, who celebrated my accomplishments and supported me during my tribulations. My sons, Michael and Paul, who provided me with laughter and love. My husband, Don, who encouraged me when the journey seemed too long, and believed in me when I doubted myself. Thank you all for supporting my dream.

TABLE OF CONTENTS

Title page	i
Abstract	ii
Acknowledgements	iv
Table of Contents	v
Tables	viii
Figures	xi
Appendices	xii
CHAPTER 1	
Introduction	1
CHAPTER 2	
Literature Review	5
<i>1. What Should Teachers Know About Literacy Acquisition?</i>	7
Stages	7
Reading Failure is Preventable	9
Components of Effective Literacy Instruction	11
Balanced Approach to Literacy Instruction	17
Summary of What Teachers Should Know About Literacy Education	17
<i>2. Pre-Service Teacher Education</i>	18
Description of a Pre-service Program	18
Are Teachers Properly Prepared to Provide Effective Literacy Education?	19
General Issues Regarding Reform in Pre-service Teacher Education.....	21
Pre-service Teacher Education Programs: Preparation for Literacy Instruction	26
<i>3. In-Service Professional Development</i>	39
What is Professional Development?	40
Lack of Research-based Teaching Methods in Classrooms	42
Four Dimensions in Teaching	44
Transfer of Learning	46
Problematic Conditions for Effective In-Service Professional Development	61
<i>4. Teachers' Impact on Children "At-Risk" for Literacy Failure</i>	62
Teachers' Knowledge and Classroom Practice	62
<i>5. Impact of Professional Development on Literacy Acquisition</i>	65
Description of Six In-Service PD Programs for the Improvement of Literacy Instruction...	66
Conclusions Concerning Professional Development to Promote Better Literacy	80

CHAPTER 3	
1. <i>The Current Professional Development Approach</i>	83
The Context for the Professional Development Initiative	83
Rationale for Use of <i>The Literacy Diet</i> Metaphor	83
Content of the Professional Development	86
Implementation of the Professional Development	90
2. <i>Current Research Objectives</i>	94
Research Context	94
Research Questions	96
CHAPTER 4	
Overview of Research Methodology	97
Test Instruments	98
Procedures: Testing and Scoring	100
Participant Characteristics	101
CHAPTER 5	
Tracking the Teachers: Study of Teacher Effectiveness	102
Tracking the Teachers: Part I	103
Participant Characteristics	103
Analyses	107
Results and Discussion	109
Summary of Findings Concerning Teachers Overall	121
Tracking the Teachers: Part II	124
Participant Characteristics	125
Analyses	126
Results and Discussion	128
Summary of Findings Concerning Individual Teachers	142
CHAPTER 6	
Tracking the Children: Study of Student Change	144
Tracking the Children	145
Participant Characteristics	145
Analyses	149
Results and Discussion	151
Conclusions Concerning Tracking Student Change Across Years with In-serviced Teachers ...	161

CHAPTER 7	
Children “At Risk” of Literacy Failure	163
Risk Factor: Teachers’ Knowledge and Classroom Practice	163
Participant Characteristics	164
Analyses	165
Results and Discussion	166
Summary of the Impact of Teachers’ Participation in PD on Students’ Rate of Literacy Failure ...	175
CHAPTER 8	
Summary and General Discussion	176
Overview of the Major Findings	176
Contributions of Current Research	178
Limitations and Future Directions	181
Final Comment	184
References	185
Appendix A.....	204
Appendix B.....	207

Tables

Table 1	Stages of Literacy Development	8
Table 2	Summary of Evidence on the Importance of Early Success in Reading	9
Table 3	Summary of Components of Effective Literacy Instruction	14
Table 4	Dimensions of Teaching	45
Table 5	Features in the Successful Transfer of New Skills to Classroom Practice	48
Table 6	Percentages of Children Making Expected Gains in Reading Comprehension	64
Table 7	Description of Four Blocks Literacy Model	68
Table 8	Description of the In-service Training/Consultation Model for Reading Instruction	70
Table 9	Description of an In-service for Reading Comprehension	72
Table 10	Description of a Two-Year PD to Improve Teachers' Instruction of Reading	75
Table 11	Description of the Four-Year Madeline Hunter Follow-Through Project	77
Table 12	Description of the Reading Instruction Study	79
Table 13	Key Components of A Balanced Literacy Diet	87
Table 14	Overview of Data Collection on a Time Line	97
Table 15	Literacy Measures Administered During Each of the Five Testing Periods	98
Table 16	Overview of Teachers Who Participated in the Study	104
Table 17	Characteristics of the Teachers	105
Table 18	Summary Table of Participant Characteristics	108
Table 19	Assessment of Teacher Effectiveness: Means and Standard Deviations of Scores for Grade One and Grade Two Children on Three Literacy Measures	110
Table 20	Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness: Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the Teachers' Professional Development and in Follow-up Years	112

Table 21	Summary Table of Participant Characteristics in the Within-Teachers Longitudinal Study	126
Table 22	Teacher A: Assessment of Teacher Effectiveness: Part II Means and Standard Deviations of Three Literacy Measures	128
Table 23	Teacher A. Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness: Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the In-Service PD and in the Follow-up Years	129
Table 24	Teacher E: Assessment of Teacher Effectiveness: Part II Means and Standard Deviations of the Three Literacy Measures	132
Table 25	Teacher E. Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness: Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the In-Service PD and in the Follow-up Years	134
Table 26	Teacher F: Assessment of Teacher Effectiveness: Part II Means and Standard Deviations of the Three Literacy Measures	137
Table 27	Teacher F. Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness: Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the In-Service PD and in the Follow-up Years	138
Table 28	Cohorts of Children in Grade 1 and Grade 2 Exposed to the Instruction of In-serviced Teachers for 1, 2, 3 and 4 Years	146
Table 29	Description of the Cohorts in the Children's Longitudinal Study: Children Taught by In-serviced Teachers	147
Table 30	Grade 1 and Grade 2 Teachers from 1994 to 1998	148
Table 31	Summary of Participant Characteristics: Tracking the Children	149
Table 32	The Means and Standard Deviations on the Achievement Measures of Word Recognition, Spelling and Reading Comprehension for the Grades 1 and 2 Children Across Years of Exposure to Experimental Teachers	152
Table 33	Assessment of Student Change: Orthogonal Comparisons between the Baseline and Experimental Groups of Children who had 1, 2, 3, or 4 Years of Exposure to In-serviced Teachers	154
Table 34	Percentage of Grade 1 Students' Word Recognition, Spelling, and Reading Comprehension Scores in each Quartile	168
Table 35	Percentage of Grade 2 Students' Word Recognition, Spelling, and Reading Comprehension Scores in each Quartile	170

Figures

Figure 1. Overall Framework for Evaluation of Professional Development	95
Figure 2. Assessment of Teacher Effectiveness: Mean Number of Words Recognized	114
Figure 3. Assessment of Teacher Effectiveness: Mean Number of Spelling Scores	117
Figure 4. Assessment of Teacher Effectiveness: Mean Reading Comprehension	120
Figure 5. Assessment of Teacher Effectiveness: Teacher A	131
Figure 6. Assessment of Teacher Effectiveness: Teacher E	135
Figure 7. Assessment of Teacher Effectiveness: Teacher F	139
Figure 8. Mean Number of Words Recognized	156
Figure 9. Mean Spelling Scores	159
Figure 10. Mean Reading Comprehension	160
Figure 11. Word Recognition (Burns & Roe Raw Scores) Grade One	167
Figure 12. Word Recognition (Burns & Roe Raw Scores) Grade Two	169
Figure 13. Spelling WRAT Raw Scores Grade One	171
Figure 14. Spelling WRAT Raw Scores Grade Two	172
Figure 15. Reading Comprehension (Gates-MacGinitie Raw Scores) Grade One	173
Figure 16. Reading Comprehension (Gates-MacGinitie Raw Scores) Grade Two	174

CHAPTER 1

Introduction

Learning to read represents the major route to learning in advanced societies. It is critical to a child's success in school and later in life.

Communications in the past were more verbal with the use of the telephone or person to person discussions. However, currently communications demand more reading and writing with messages sent by electronic mail, Internet, fax and printed documents. Thus, there are now widespread expectations that all children will learn to read. In many literate countries only 40 to 60 percent of the population achieve literacy (Snow, Burns & Griffin, 1998). To be employable, high school graduates need to be able to read challenging material and to solve sophisticated calculations and problems. Experts warn that the primary and secondary school systems are still preparing students for an old economy. The Collegium of Work and Learning, a Toronto-based educational think-tank, reports that, "less than half of the workforce will be in traditional full-time jobs within the next decade" (Sheppard, 1998). According to Snow et al. (1998), the demands on today's education system are far greater than a quarter-century ago.

Research is beginning to demonstrate that consistently high quality regular classroom instruction which addresses the needs of all children is essential if all children, irrespective of background, are going to learn to read and write (Allington, 1994). A large number of children who should be capable of reading and who have been exposed to instruction, are not reading. Since 1975 the number of students identified for special education has risen by more than 250% in the United States (Slavin, Madden, Karweit, Dolan, & Wasik, 1991). The number of children classified as learning disabled has doubled in the last ten years (Gaffney, 1998).

Why the huge rise in the number of children requiring special services, often in the area of reading? Some researchers state that many children are experiencing reading difficulties due to poor instruction and a poor classroom curriculum. These students are called “curriculum casualties” (Snow et al., 1998, p. 25). With our current higher expectations that every child should learn to read and write, there is a movement to focus more resources on the prevention of reading and writing problems within our schools. Thus, with a focus on prevention, there would be fewer students requiring remediation for literacy problems. The challenge facing our society is to ensure that our teachers are adequately prepared to effectively teach literacy skills to all our children.

An underlying question of concern in this thesis is: Are teachers adequately prepared to teach reading in our classrooms? Several researchers have concluded that reading failure is preventable for all but a small percentage of children, and that prevention of reading failure depends on effective instruction (Adams & Bruck, 1995; Chard, 1999; North, 1999; Pikulski, 1994; Snow et al., 1998). However, according to a position paper of the Orton Dyslexia Society (1997), “reading failure is epidemic” (p. 20). National studies in the United States indicate that one in six elementary students encounters reading problems (Gaffney, 1998).

Why are students encountering difficulties learning to read? There is a growing body of literature arguing for improved pre-service education for our novice teachers (Booth, 1992; Lyon, Vaassen & Toomey, 1989; Manzo, 1991; Moats, 1995; Nolen, McCutchen & Berninger, 1990; Richardson, 1996; Sindelar, 1995; Soltis, 1990). It is suggested that there are too many primary teachers entering the field without a good education on how to teach or assess basic reading skills (Adams & Bruck, 1995).

Furthermore, due to the increase in knowledge emanating from a burgeoning literature of reading research, there is now more to teach prospective teachers as well as experienced teachers on how reading and writing skills are learned and how these skills should be taught. Thus, in-service professional education would appear to be as important as pre-service education for improving the instruction of literacy skills. As a result, another growing body of research examines the need for in-service teacher development and the features of effective in-service teacher programs (Darling-Hammond, 1996; Gersten, Morvant & Brengelman, 1995; Guskey, 1986; Joyce & Showers, 1988; Little, 1993; Richardson, 1998; Smylie, 1989).

This thesis reviews the research literature examining the preparation of teachers, both pre-service and in-service, for effective instruction of literacy skills to primary students. In reviewing this literature it became clear that there is a dearth of controlled experimental studies in both pre-service and in-service teacher education. Many papers report opinions, and but there are few research-based evaluations of teacher education programs.

The purpose of the present longitudinal study was to evaluate the long-term impact on the reading and spelling performance of students in the primary grades of a particular in-service professional development system called *The Balanced and Flexible Literacy Diet: Putting Theory into Practice* (Willows, 1994). The overall goal of the research was to determine whether teachers internalized, applied and maintained the knowledge acquired through professional development (PD). An additional goal was to determine whether the proportion of students considered to be “at risk” of literacy failure is reduced when their classroom teachers participate in PD on literacy education.

In order to achieve these research goals, both teacher change – the impact of the in-service education on the instructional effectiveness of participating teachers – and student change – the impact of length of exposure to the instruction of in-serviced teachers on literacy learning outcomes – were examined. Two types of longitudinal analyses were undertaken on the data set. One type tracked teachers who had participated in the PD by comparing scores of the children in their classrooms across four time periods (pretreatment year, treatment year, follow-up year 1, follow-up year 2). The other type of longitudinal analysis tracked cohorts of children across 3 years (from Junior Kindergarten through Grade 1) and across 4 years (Junior Kindergarten through Grade 2) in the classrooms of teachers who had participated in the PD. In addition, the study explored the impact of the PD on teachers' success in raising the literacy achievement of the students considered most "at risk" of literacy failure – those in the lowest performance ranges on literacy measures.

CHAPTER 2

Literature Review

The purpose of this review of the literature on teacher education was to determine the amount and type of evidence available to guide the development and implementation of programs for pre-service and in-service professional development in literacy education. In this search of the literature it was found that the vast majority of the published articles concerning teacher education are descriptive, theoretical, and/or opinion papers and very little of the literature is based on either quantitative or qualitative research evidence. There are, however, some very insightful analyses of the challenges and impediments in the development of effective professional educators.

The literature is divided between two major aspects of professional teacher development. One aspect examines the instruction of “student teachers” before they have had significant classroom experience, while they are training to become teachers. These individuals are essentially full-time students. This stage is known as “pre-service teacher education.” The other aspect examines the instruction of practicing teachers who have had significant classroom experience and who are professional teachers. These individuals are essentially full-time teachers, and part-time students who require maintenance and expansion of their instructional skills. This stage is known as “in-service teacher education.”

Given these different contexts and different types of learners in these two contexts of professional development, pre-service and in-service teacher education are presented as two different sections in the review of the literature. Although the present study is focused on in-service teacher education, the review of the literature begins with an examination of pre-service teacher education since pre-service experiences are the foundation upon which teachers base their practices. Some teachers may not pursue further education beyond pre-service and it is

only from the knowledge gained from their pre-service education that they learn how to teach reading. Therefore, it is essential to discuss those factors that contribute to the growth in teachers' knowledge and skills in pre-service education. Furthermore, teacher education is viewed as a continuum of lifelong learning from pre-service and continuing into in-service (Rolheiser, 1996),

Although this thesis is focused on the impact of teacher education on students' acquisition of literacy skills – teacher effectiveness and student success – much of the literature review is focused on general principles and issues of teacher development, rather than specific subject areas. These general principles would likely impact most, if not all, curriculum areas.

Therefore, a discussion of these general principles is included in the literature review as a foundation for the specific application of professional development in improving instruction in literacy.

This literature review is presented in five major sections: (1) The first provides an overview of some key issues arising from basic research on the acquisition of literacy skills and points to some important principles of effective reading instruction. (2) The second describes pre-service teacher education programs, documenting the need for reform in teacher preparation programs in order to improve classroom literacy instruction. (3) The third section examines the features of effective in-service PD. (4) The fourth section examines the impact of teachers' instruction on children who are "at risk" for literacy failure. (5) The last section reviews specific in-service teacher education programs developed for improving the literacy outcomes of students. It presents conclusions concerning professional development for promoting better literacy education and introduces the next chapter in which the rationale of the professional development approach, and research objectives of the current study are presented.

1. What Should Teachers Know About Literacy Acquisition?

In this first section, the research addresses the importance of early success in attaining reading skills, and the areas of difficulty for some students in the acquisition of reading skills. This research sets the stage for asking the questions: Is reading failure preventable? What are the principles of effective reading instruction?

Stages

Many researchers view reading as the product of a complex interaction of abilities and instruction that changes with development (Chall, Jacobs & Baldwin, 1990; Frith, 1986; Juel, 1987; Marsh, Friedman, Welch & Desberg, 1981). From this perspective, reading for the preschooler, first grader, third grader, and high school student is essentially different. As students progress through reading stages, the processes involved and the skills needed change. Thus, the factors influencing achievement also differ with reading development.

Chall (1983, 1996) proposed several stages of reading development. Table 1 presents 4 stages of literacy development – including both reading and writing – adapted from Chall’s reading stages.^{*} Similarly, Juel (1987) also discussed three stages of reading development in the early elementary grades. In her “selective cue” stage the child attends to random features either in the environment where the print occurs (e.g., the red colour of the stop sign) or to some feature in the print itself (e.g., the length of the word). Usually around first grade, the child enters the “spelling-sound” stage, attending primarily to the graphic information of the printed words – “glued to print.” Around grade 3 to grade 4 the child reduces attention to the print or becomes “unglued” as he or she becomes more automatic in the processing of letters in every word.

^{*} Note: Although reading and writing processes are obviously interconnected, a large portion of the research literature has focused on reading. Some researchers, however, (e.g., Ehri, 1986; Juel, 1988; Shanahan, 1988) have focused specifically on the interrelation among the “literacy processes” involved in reading and writing.

Table 1 Stages of Literacy Development

<p>Stage 0 PRELITERACY STAGE (6 months to 6 years)</p> <ul style="list-style-type: none"> - pretends to read familiar books - names letters of alphabet - recognizes some signs - plays with pencils and paper - prints own name - develops phonological awareness - acquires concepts of print 	<p>Stage 1 BEGINNING LITERACY (6 years to 7 years)</p> <ul style="list-style-type: none"> - learns relations between letters and sounds - learns to print letters and numbers - learns to recognize high frequency words “by sight” - learns to sound out regularly spelled words - learns to use contextual and pictorial clues - learns to write through copying and inventive spelling
<p>Stage 2 BEGINNING FLUENCY (7 to 8 years)</p> <ul style="list-style-type: none"> - reads simple, familiar stories independently - consolidates “sight” vocabulary - expands letter-sound knowledge - develops fluency (speed and accuracy) in reading - writes and spells with less effort 	<p>Stage 3 LITERACY FOR GROWTH (9 years and up)</p> <ul style="list-style-type: none"> - reads “for pleasure” - reads to gain new knowledge - expands vocabulary through reading - writes and spells more automatically - writes for communication with others - writes for personal expression of ideas

Note. From “Learning to Read and Write” (Willows, 1993, p.15). Adapted from Chall’s (1983) *Stages of Reading Development*. Reprinted with permission.

The overarching goal of reading instruction is for children to learn to read fluently and to understand the text. The stages of literacy development build on each other so the later a child learns to read and write, the more difficult it is for that child to cope with the increasing demands of literacy in the higher grades. Therefore, early success with reading appears critical. According to Kameenui (1998), if low achieving students can be brought up to grade level within the first three years of schooling, their reading performance stays at grade level henceforth but “if we fail to bring a student’s reading to grade level within the first few years, the likelihood of their ever catching up is slim, even with extra funding and special programs” (p. 321). Sources of evidence supporting this statement are summarized in Table 2.

Table 2 Summary of Evidence on the Importance of Early Success in Reading

SOURCES	CONCLUSIONS
American Federation of Teachers (1999) Adams & Bruck (1995) Chall et al. (1990) Gough & Tunmer (1986) Juel (1988) Kameenui (1998) National Reading Panel (2000) Pressley (1998) Snow et al. (1998) Stanovich (1986) Vellutino et al. (1996)	<ul style="list-style-type: none"> - need to foster early awareness of the nature of print - more phonemic awareness training should occur in preschool and kindergarten - phonemic awareness predicts reading achievement - poor decoding skills lead to little reading - automatic word recognition skills predict reading comprehension ability - when children master the early stages of reading then reading becomes more rewarding and children will read more - more reading means better reading - more reading leads to more vocabulary, new knowledge, new ideas - "Matthew Effect" which means the rich get richer and the poor get poorer - the gap in the amount of exposure to print between good and poor readers grows wider with each grade - around Grade 3 the major factor determining vocabulary growth is reading - good readers made greater gains in listening comprehension - Simple View of Reading - product of decoding and listening comprehension - poor readers tend to become poor writers - reading helps develop awareness of the components of a good story - spelling depends on word specific knowledge that is acquired through reading

Reading Failure is Preventable

Researchers are beginning to report that reading failure is preventable for all but a small percentage of children (Pikulski, 1994). The Committee on the Prevention of Reading Difficulties states that "most reading difficulties can be prevented" (Snow et al., 1998, p.13), and the prevention of reading failure depends on effective instruction (Allington &

Cunningham, 1996; Allington & Walmsley, 1995; Pikulski, 1994; Torgesen, Wagner, & Rashotte, 1997).

In support of the concept of “curriculum casualties,” Vellutino et al., (1996) suggest that the majority of children who might be diagnosed with a reading disability actually lack experience and have received poor instruction rather than possess cognitive deficits. In their study, 9% of the children participating in their program were defined with a reading disability according to exclusionary criteria usually used to identify these children (i.e., excluding the children with low general intelligence, sensory deficits, emotional disorder, motivational problems, frequent absences from school, etc.). After one semester of individual one-to-one remedial tutoring only 3% scored below the 30th percentile and 1.5% scored below the 15th percentile in reading achievement.

From the perspective of “curriculum casualties” many children may have processing weaknesses that will interact with factors in their environment, such as quality of reading instruction. However, these children may not have the severity of processing weakness that would account for their extreme difficulties in learning to read and write. Although, those children who fall behind their peers in learning to read may have some mild processing weaknesses, the evidence from the Vellutino et al. (1996) study suggests that an appropriate curriculum may compensate.

Although there are still major gaps in the knowledge of instructional methods for teaching effectively the 3% to 5% of children with the most severe reading disabilities, there is growing evidence that it is possible to produce significant improvements in the rate of reading growth in

groups of children with a broad variety of learning problems when “well known principles are consistently applied in their reading instruction” (Torgesen et al., 1997, p. 217).

Thus, there is increasing evidence that the teacher could be a key player in the reduction of reading failure in our schools.

Components of Effective Literacy Instruction

Why do some children have difficulty learning to read? According to Snow et al. (1998) there are three major areas where children may have problems in developing fluent reading skills.

The first challenge posed to a beginner reader is understanding the alphabetic principle. This is a challenge since the units represented graphically by the letters of the alphabet have no meaning and are “phonologically abstract” (Snow et al., 1998, p. 22). The child must come to understand that written spellings systematically represent spoken words. The idea that the graphic letters represent the small sound units within spoken words called phonemes is very important in the development of reading. The breaking of this spelling-to-sound code is essential if the child is to gain the reading independence that will eventually lead to reading practice, which is absolutely necessary for attaining fluency.

After 20 years of research the results consistently demonstrate that the major problem facing reading disabled children is their poor phonological awareness (Adams, 1990; Stanovich & Siegel, 1994). Phonological awareness is “the ability to recognize that spoken words consist of smaller units of sound” (Cisero & Royer, 1995, p. 275). Phonological awareness develops over a number of years, progressing from early rhyming abilities to explicit awareness of the individual phonemes in words (Moats, 1995). Phonemes are difficult for children to become aware of because when humans say words, the phonemes overlap and influence each other so

individual phonemes are difficult to isolate. The reader must be aware of the phoneme as the basic unit of the spoken word. However, phonemic awareness is not an insight, but a skill.

According to Snow et al. (1998) the second challenge for children learning to read is the transfer of the comprehension skills of spoken language to reading. Supporting this second challenge is research by Torgesen and Goldman (1977) where poor readers were less likely to use cognitive, text-level memory strategies. Furthermore, research has shown a strong relationship between reading fluency and reading comprehension (Stanovich, 1980). Poor readers do not decode the sounds of words automatically, and thus have less memory capacity to allocate for comprehension processes. Thus, there is a powerful causal connection between phonological awareness and reading for meaning (Adams, 1990; Kameenui, 1998).

The third obstacle to learning to read is the loss of motivation to read, since the acquisition of reading fluency, and thus, vocabulary development occurs through reading. Motivation is crucial for learning to read (Baker & Wigfield, 1999; Guthrie & Wigfield, 1997; Snow et al., 1998; Stanovich, 1986). If learning to read is too effortful, then children will not enjoy reading (Cunningham & Stanovich, 1993).

A child who reads well, will read more, learn more word meanings, and hence read better. The child who reads slowly, “sounding out” out each word, will not enjoy reading, will read less, and will fall farther behind the better reader, not only in reading skills but also in the attainment of other cognitive skills. There is a spiraling effect of being a poor reader, since reading affects a broad range of tasks and skills. As the child progresses in school most academic tasks depend on reading and thus, all school work could become problematic for a poor reader. Stanovich (1986) called this spiraling effect the “Matthew effect,” or in other words, the rich get richer and

the poor get poorer. A reading disability can result in feelings of failure, and helplessness that permeate children's behaviours and attitudes in school and in other parts of their lives. There is a reciprocal relationship between reading and academic and personal successes in the life of a child. Motivation to read is critical for accelerating growth in literacy.

How can these three obstacles to learning to read be surmounted? There has been an enormous amount of research addressing the question "What are the most essential components of effective literacy instruction?" There is converging agreement among researchers that there are some key components that should be included in a balanced literacy program. Chard (1999) argues that understanding the theory of reading, the developmental stages of reading, speech-sound awareness, speech to print mapping, the importance of automaticity in word and text reading, the importance of vocabulary development, syntax and academic language, comprehension strategies, and knowledge of classic, contemporary literature, with knowledge of methods for the assessment of written language progress are necessary for effective instruction in the complex demands of today's classrooms. Similarly, Pressley and Wharton-McDonald (1998) found in a survey of the teaching practices of teachers nominated as exceptional in teaching literacy, that many components of reading instruction were integrated into their classroom instruction which combined the features of whole language with explicit skills instruction. Within a literature rich environment including in-class library, displays of student work, and displays of stories and poems, the teachers reported developing many specific competencies such as concepts of print, letter recognition, the alphabetic principle, letter-sound associations, decoding strategies, punctuation, sight words, vocabulary, spelling, comprehension strategies, elements of text, critical thinking skills, and writing. These components reflect a combination of various major approaches to literacy instruction, whole language or language experience, whole word programs and phonics programs. The features of all the programs are

integrated so teachers are sensitive to the individual instructional needs of their students. These teachers reported changes in their teaching practices as their students advanced through the grades. The components included in the in-service system evaluated in this thesis are summarized in Table 3 and are consistent with the research discussed by Chard (1999) and Pressley et al. (1998). Table 3 presents a list of the main components required for effective reading and writing instruction along with some of the key sources of evidence.

Table 3 Summary of Components of Effective Literacy Instruction

SOURCES OF EVIDENCE	COMPONENTS
Adams (1990) Brady, Fowler, Stone & Winbury (1994) Cunningham (1999) Foorman, Francis, Fletcher, Schatschneider & Mehta (1998) Gough (1996) Griffith & Olson (1992) Juel (1987, 1988) Lundberg, Frost & Petersen (1988) Moats & Foorman (1997) Yopp (1992)	1. Phonemic Awareness Instruction - use a variety of strategies for manipulating the sounds in words - phonemic awareness is a developing skill - phonemic awareness skills can be taught to children - children taught phonemic awareness in kindergarten developed better reading skills in Grade 1 and Grade 2 compared to a control group
Adams (1990) Clay (1991) Cunningham & Stanovich (1993) Gaskins, Ehri, Cress, O'Hara & Donnelly (1997) Scanlon & Vellutino (1997) Stahl (1992)	2. Teach Alphabetic Understanding - systematic instruction of the alphabetic code of written English - concepts of print
Bolton (1994) Daniels (1994) Hoover & Gough (1990) Siegel & Ryan (1984) Scanlan & Vellutino (1997) Willows & Ryan (1986)	3. Language Development - if a child has problems in receptive, and/or expressive language functions he or she will almost inevitably experience difficulties in the written domain as well - the processing of written language is directly dependent on the phonology, morphology, semantics and syntax of oral language - "simple view" of reading - decoding and linguistic comprehension

Table 3 (cont'd)

SOURCES OF EVIDENCE	COMPONENTS
Armbruster & Nagy (1992) Biemiller (1999) Juel (1987) McKeown & Beck (1988) Stanovich (1986) Torgesen, Wagner, & Rashotte (1997)	4. Vocabulary Development - reading slump occurred with at-risk children in Grades 4 and 6 due to poor vocabulary development - semantic maps, semantic feature analysis - content area vocabulary - rich instruction with several activities with words, frequent use of new words, extend use of new words to environments outside of classroom
Bromley, Irwin De Vitis & Modlo (1995) Maria (1989) Neuman (1999) Pressley (1999)	5. Background Knowledge - prior knowledge activation improves reading comprehension - good comprehenders, based on their prior knowledge, anticipate the contents of the text - low SES children are prone to lack of appropriate background knowledge - need to bridge the gap between what children know and what they need to know
Armbruster, Anderson & Ostertag (1989) Fielding & Pearson (February 1994) Hansen (1981) Palincsar & Brown (1984). Pressley (1998) Stanovich (1986) Torgesen & Goldman (1977)	6. Teach Reading Comprehension Strategies - monitoring one's progress as a reader - questioning, summarizing, predicting - identification of main ideas - constructing mental images
Allington (1983) Samuels (1997) Shanahan & Samuels (2000) Stanovich (1986)	7. Promote Reading Fluency - real reading - repeated readings
Baker & Wigfield (1999) Gambrell (1996) Guthrie (1996) Guthrie & Wigfield (1997) Johns & VanLeirsburg (1994) Turner & Paris (1995)	8. Motivation for Reading and Writing - the engaged reader: motivated, knowledgeable, strategic and socially interactive - four key features are associated with motivation to read: access to books in the classroom, opportunities to self-select books, familiarity with books, social interactions with others about books -link students' intrinsic motivations to classroom activities

Table 3 (cont'd)

SOURCES OF EVIDENCE	COMPONENTS
Brown & Ellis (1994) Cunningham (1995) Ehri (1986) Fresch & Wheaton (1997)	9. Instruction in Spelling - spelling is a problem solving process - spelling growth is developmental and continues into adulthood - stages of invented spelling: precommunicative, semiphonetic, transitional, conventional - develops from letter-name stage to using more sophisticated strategies such as analogy
Bereiter & Scardamalia (1982) Bromley (1999) Freedman (1998) Hansen (1998) Hayes & Flower (1980) Scardamalia (1981)	10. The Writing Process - brainstorming, suggestive prompts, means-end planning, seeking alternatives in sentence structure, feedback loop-evaluation - classrooms which support oral language development also foster children's writing - choice in writing for a variety of purposes and authentic audiences - large blocks of time for writing - direct instruction in composing and conventions of grammar, spelling, form and handwriting - writing across the curriculum regularly
Afflerbach (1998) Allington & Cunningham (1996) Housego (1994) Fullan (2000) Winograd & Arrington (1999)	11. Assessment - ongoing, reliable and valid assessment of reading and writing skills - provides information that informs educational practice - assessment is an opportunity for teaching and learning - connection between reading assessment and reading achievement - early reading assessment serves the critical need of identifying what skills a student possesses and what needs to be developed - the varying developmental stages of readers and writers require that teacher evaluation and observation play a central role

Balanced Approach to Literacy Instruction

Evidence is increasingly pointing to the need for a “balanced” approach in the instruction of literacy (Pressley, Rankin & Yokoi, 1996; Spiegel, 1998). Effective primary instruction is multifaceted and does not support one approach or another but includes the components of effective literacy instruction discussed in the previous section (Pressley et al., 1996). A balanced approach to the teaching of literacy is not constrained or reactive to a particular philosophy or program, but requires a teacher to make thoughtful choices each day about the best way to help each child in his or her classroom to become a better reader and writer. Balanced literacy acknowledges that not every child learns in the same way, not every literacy activity requires the same strategies, not every teacher has the same talents, and not every school has the same combination of teachers and students (Spiegel, 1998). Every teacher selects the “right” task for each child, and is able to easily change the particular emphasis of instruction. Therefore, the balanced literacy approach to instruction assumes that the teacher has a lot of knowledge about child development, knowledge about the cognitive processes in reading/writing, and knowledge of what strategies work best for what child (Ferguson, 1999). “It views the teacher as an informed decision-maker who develops a flexible program and it is constructed around a comprehensive view of literacy” (Spiegel, 1998, p. 117). As Pressley et al. (1996) state, “no two balanced programs are identical” (p. 365). A balanced program is built upon research-based information and it views the teachers as an informed decision-maker who has a comprehensive view of literacy and recognizes the contributions of different approaches in teaching reading and writing.

Summary of What Teachers Should Know About Literacy Education

In sum then, there is a growing consensus – based on the research on literacy acquisition – that every elementary school teacher should have a foundation of knowledge that includes: an

understanding of the stages of literacy development, a recognition that reading/writing failure is preventable, and a comprehensive knowledge of the components necessary to provide a balanced and effective literacy program.

2. Pre-Service Teacher Education

This section will begin with a description of a typical pre-service teacher education program in Ontario. Next, the general challenges facing the faculties of education for preparing teachers with the required skills to teach all students will be explored. Following this more general discussion, issues relating directly to the preparation of student teachers for the teaching of reading and writing will be described. Finally, a pilot project aimed at improving pre-service teacher education in Canada's largest Faculty of Education will be described.

Description of a Pre-service Program

Pre-service teacher education in Ontario occurs in two settings: courses in the university and practicum teaching experiences in classrooms of schools. In Ontario, teacher candidates are assigned a minimum of 40 days of practical experience in one or two schools (Beynon, 1999).

In order to select their university courses teacher candidates must choose one of three programs, primary/junior (for teaching grades JK - 6), junior/intermediate (for teaching grades 4 - 10) and intermediate/senior (for teaching grades 7 - OAC). Courses will vary depending on the age of students to be taught since a primary/junior (P/J) teacher must be a generalist, teaching all of the curriculum, whereas a junior/intermediate (J/I) candidate must be a generalist as well as a specialist in one subject at the intermediate level, and an intermediate/senior candidate must be a specialist and teach two subjects at the senior level.

Teacher candidates are assessed on the quality of work within their university courses and the quality of their application of newly acquired teaching skills in the classrooms of their practicum settings.

The question arises: Are the principles of good reading instruction through a balanced literacy approach being taught to our teachers? If not, why?

Are Teachers Properly Prepared to Provide Effective Literacy Education?

According to MacGinitie (1991), there is a need to shift our teaching methods from what is new to what is working. Prevention of reading failure depends on effective instruction. “Effective instruction depends on a deep, thorough and flexible understanding of the knowledge and processes involved in reading and of how they vary across the development of children” (Adams & Bruck, 1995, p.18).

It is widely believed, however, that classroom teachers lack this type of understanding about literary acquisition (Adams & Bruck, 1995). Chall et al. (1990) state that few teachers have sufficient theoretical knowledge and practical skills to teach at-risk children successfully.

Furthermore, in an article titled *Cognitive Science Meets Beginning Reading* Stanovich (1991) states that very little research on reading has “filtered through to reading teachers, parents, and educational administrators...remedies for dyslexia are still more likely to emanate from cuckoo land than from research literature” (p. 79). From this perspective, the gains from three decades of research could dramatically increase the levels of reading achievement. So for many researchers and educators, “the solution rests in adding new research-based information and methods to the preparation of teachers” (Brady, 1999, p. 1).

To add new research-based information in the courses taught to our teachers will require an alliance between the educational researchers and educators of our teachers. The educators must keep abreast of the new research in the area of reading, so this knowledge can be taught to our teachers. Research knowledge needs to be taught to new as well as experienced teachers. In this way the tremendous gains in knowledge on effective literacy instruction will filter into our classrooms.

If research is demonstrating that most reading and spelling disabilities are originating from an impairment in the knowledge of the structure of language and of the alphabetic writing system, then teachers must address these issues. However, a survey by Moats (1995), involving 89 teachers with 5 to 20 years of experience, indicated that they had “an insufficient grasp of spoken and written language structure and they would be unable to teach it explicitly to either beginning readers or those with reading/spelling disabilities” (p. 45). Similarly, a survey of 440 teachers by Lyon, Vaassen and Toomey (1989) found the majority of teachers reporting that their pre-service education program did not provide effective instruction within the university or school practicum settings. Only 20% of regular teachers and 10% of special educators felt they had adequate content preparation in reading and reading instruction (Moats & Lyon, 1996). A more recent 1999 survey by the U.S. Secretary of Education supported these findings as again only 20 per cent of the respondents expressed confidence in their abilities to meet the needs of students with reading disabilities (North, 1999). The teachers felt underprepared at the end of their education and stated that their expertise developed primarily through their own teaching experiences. However, poor teaching habits may result from a lack of adequate content and pedagogical preparation. Thus, these papers support the view that pre-service teacher education is not fulfilling its role to properly prepare our teachers for effectively teaching literacy skills to all students.

The next question arises: Why is pre-service teacher education not fulfilling its role of properly preparing our teacher candidates to teach literacy?

General Issues Regarding Reform in Pre-service Teacher Education

This section is focused on issues in pre-service teacher education which are impacting the instruction of teacher candidates. This background is important because it provides the broad context into which specific issues concerning teacher preparation in literacy education are formed.

What has been occurring in the teacher preparation programs in the last 15 years? There have been numerous shifts in education reform. In the 1960's reform was focused on curriculum development, relevant curriculum for minority students and upgrading teacher education. During the 1970's there was the "back to basics" movement and in the 1980's and 1990's improvement in the instruction of teachers has become the central concern of reform (Metcalf-Turner & Fischetti, 1996).

There is large variation in the types of programs among the many Faculties of Education in the United States and Canada. These programs vary according to level of funding, types of course work and type and amount of time expected in practicum classroom experience. Some states require only one year of pre-service education, whereas others require two years (Abell, Dillon, Hopkins, McInerney & O'Brien, 1995).

With this large variation in available pre-service programs, licensure was introduced to protect students from being taught by ineffective teachers. Each state in the United States has adopted

its own standards for licensing teachers. However, the quality and diversity of teacher preparation programs is a major problem. In a survey in 1986 (Howey & Zimpher) it was found that the number of full-time faculty members in Schools and Departments of Education in the United States varied from 1.5 to 200.

There is an extensive literature on the teaching and learning processes, as previously discussed. However, there does not appear to be a body of research guiding the development of pre-service teacher preparation programs and in-service instructional programs. Darling-Hammond (1996) reports that as recently as 10 years ago there was little emphasis on teacher knowledge as critical for educational improvement. Policymakers looked for the “right set of test prescriptions, textbooks adoptions and curriculum directives to be packaged and mandated to guide practice” (p. 5). There was little reflection on how the ideas from legislation entered the classroom. Raising standards for literacy, mathematics and science will be of little use if there are not enough teachers prepared to teach these subjects well.

Changes are taking place in teacher preparation programs with performance-based approaches to licensing and accreditation of programs being considered (Chambers, Presseault & Sheehan, 1998; Del Schalock & Myton, 1988; Salinger, 1995). The new National Board for Professional Teaching Standards in the USA was created to provide assessments for certifying accomplished teachers. “The invention of 21st century schools that can educate all children well rests, first and foremost, upon the development of a highly qualified and committed work force” (Darling-Hammond, 1996, p. 5).

Several national boards in the USA are involved in developing more meaningful standards for teaching, including performance-based assessments for teacher licensing, and national standards

for teacher education, licensing, and certification. These include the National Board for Professional Teaching Standards (NBPTS), the Interstate New Teacher Assessment and Support Consortium (INTASC) and the National Council for Accreditation of Teacher Education (NCATE). To address how to improve student learning in the USA 12 leading national organizations formed the Learning First Alliance in 1997. In Canada, the *Standards of Practice for the Teaching Profession* (Ontario College of Teachers, 1999) were approved, in principle, by the Council of the Ontario College of Teachers in December 1998. One of the mandates of the College is to regulate the pre-service and in-service education of teachers based on these standards. The College has initiated an accreditation process for all the Faculties of Education in Ontario so there will be similarities in the course work and standards among these Faculties.

The question arises: Are the standards adopted by these American boards and the Ontario College of Teachers based on research and theory? There is a note of caution in the literature, stating that state/provincial program approval in teacher education, and teacher licensure may depend little on research to influence its decisions in planning (Lilly, 1992). There is conflict between the higher institutions of education, the universities, and the state departments of education. There are issues of who controls entry into the teaching profession, the universities or the state and who decides on the quality or competencies of a teacher.

Lilly (1992) and Reiman and Thies-Sprinthall (1993) articulate the need for the integration of research with policy analysis, and the need for the cooperative efforts of policy makers, school teachers, university teachers and researchers. Thus, research and the teaching of student teachers should not be fragmented in the higher institutions providing teacher education programs. Research on education should not be separated as a different department or area from the education of teachers. Teacher education programs need to be based on research. "All too often

a huge disconnect occurs between the content of teacher preparation programs in reading and what research shows about how students become skilled readers” (Reback, 1999, p. 16).

There are, however, a number of identified conditions that have been viewed as undermining teacher education: “low prestige of educational departments, state-mandated curricula and accrediting requirements, and emphasis on research over teaching” (Simpson, Whelan & Zabel, 1993, p. 11).

Many teacher preparation programs evolved from normal schools, to teacher colleges to state colleges to regional state universities (Shen, 1995). Part of this evolution in many of these teacher education institutions has been the shift in emphasis from teaching to research. In a survey of 29 colleges and universities with teacher education programs, 30% of the schools, colleges and departments of education (SCDEs) had little or no involvement in the preparation of teachers since they were focused on research (Shen, 1995). Faculty members more heavily involved in preparation of teachers (53.3% of faculty) were less likely to engage in research. Therefore, the goals of the SCDE’s are fragmented. All the SCDE faculty in this survey expect their involvement in research to be vital for obtaining tenure and thus, they predict the importance of research will be greater than the preparation of school teachers.

However, it is crucial to uphold the notion that teaching and service to schools merit reward. There is little status given to supervisory duties, and therefore, there is little reward for the university to support its student teachers (Hoover, O’Shea & Carroll, 1988). Without more equity in the reward structure, the group of faculty heavily involved in teacher preparation will be marginalized. According to a survey of 835 teacher educators at 49 teacher preparation courses “they do not believe their academic peers accept them as academic equals”

(Reynolds, 1995, p. 219). The respondents agreed that teacher education has a “second-rate status in the university” (Reynolds, 1995, p. 222). Therefore, there is a morale problem within teacher education institutions in the United States. There is a cycle of low status and low expectations. Teacher educators believe their colleagues see them as lacking in scholarly productivity and perceive education as a haven for the less academically able. The issue of scholarly rigor worries members of the faculty. The faculties of education require “balanced” teacher educators. These are teachers in the faculties who are aware and understand the recent research findings on good classroom instruction and at the same time are available to teach and nurture teacher candidates. Presently, teacher educators have an immensely important function, but they have little status or power to effect real change in their institutions (Reynolds, 1995).

Due to the lower prestige of teacher educators, many educators change their roles from instructing teacher candidates towards research or administration. However, long-term commitment of qualified personnel would enhance pre-service teacher education programs. To attract future competent teachers additional financial support for teacher education programs will be essential. However, teacher education programs are poorly funded compared to other programs at the universities (Ebmeier, Twombly & Teeter, 1991). In a survey over 10 years of six major research universities in the United States, the schools of education were the lowest funded unit in universities, with minor exceptions, (e.g., foreign languages and letters) and “have lost ground over the last 10 years both in terms of constant dollars and in relation to other disciplines” (Ebmeier et al., 1991, p. 233).

Pre-service Teacher Education Programs: Preparation for Literacy Instruction

This section presents a discussion of pre-service teacher education programs with narrower focus specifically on the preparation of student teachers for the instruction of reading and writing skills.

Pre-service teacher candidates may come into their training with a strong foundation in content areas such as Math, Science and History based on their High School and Undergraduate studies. However, rarely will they have any particular knowledge of early linguistic development and reading and writing processes other than the fact that they are themselves users of the oral and written language. In spite of the fact that the most critical area of instruction is reading, teachers receive little education in this important subject (Reback, 1999).

Role of the University

As previously described, pre-service teacher education programs consist of two areas: the development of a knowledge base reflected in course offerings at the college or university with an emphasis on teaching theory, educational research, and specific content (e.g., reading, math, science) and a skills development portion of the curriculum reflected in the practicum teaching sites, where the student teachers are expected to put into practice the knowledge they have gained in the university courses (Obiakor & Utley, 1997).

For this study, the question arises: Are the universities supplying our pre-service teachers with the theoretical knowledge of how children learn to read and write?

Nolen, McCutchen and Berninger (1990) surveyed 48 state departments of education regarding the current certification requirements for teaching reading and writing at elementary, secondary

and adult education levels. Their findings indicated that the requirements are minimal. Most states in the USA require only one or two reading courses for a teacher to have the credentials to teach in elementary school and virtually no states require reading courses as necessary for secondary teachers to become certified (Reback, 1999). Moreover, the course offerings in reading are often “limited to children’s literature and/or a single course in the methods and materials of teaching reading” (Reback, 1999, p. 16). Despite research evidence to support the importance of key literacy components such as alphabetic understanding, phonemic awareness, development of language, vocabulary instruction, development of background knowledge, instruction of reading comprehension strategies, an environment to support motivation for reading, promotion of reading fluency, real reading and writing experiences, spelling instruction and ongoing assessment, it appears that many of these components are not being taught to student teachers in the pre-service preparation programs.

As an example, research has shown the importance of ongoing, reliable assessment of reading achievement (Afflerbach, 1998; Fullan, 2000; Winograd & Arrington, 1999). The task of becoming an insightful and consistent assessor of students is a challenging task for most teachers, and yet, state certification require one or no courses on assessment (Afflerbach, 1998).

How will the results of reading research become part of everyday classroom practice? To successfully incorporate the knowledge acquired from the research in reading into teaching practice, there have to occur both bottom-up changes in the classroom and top-down changes in teacher education and teacher certification faculties (Nolen et al., 1990). A core set of research-based principles in reading courses in teacher preparation programs would promote better instructional practices in the classroom. Moats (1995) states, “it is critical that teachers be able

to interpret errors, present linguistic concepts accurately, with appropriate examples, and relate spoken and written language” (p. 76).

Why do gaps exist in teachers’ knowledge of language structure? Complete and explicit awareness of phonemes is an undeveloped metalinguistic skill for the average adult and many adults think of words in written rather than spoken form, unless taught to do otherwise (Moats & Lyon, 1996; Scarborough, Ehri, Olson & Fowler, 1998). Moats (1995) surveyed 89 teachers’ knowledge of language structure. Their knowledge of phonics was surprisingly weak, and only 27% were able to identify the component morphemes of specified words. Moats (1995) outlined the advantages of a solid language knowledge base in the teaching of reading: being able to interpret and respond to student errors, being able to choose the best examples for the teaching of decoding and spelling, being able to properly organize and sequence information for reading instruction, being able to use knowledge of morphology to explain spelling, being able to incorporate components of language arts into other areas of the curriculum. After participating in Moat’s (1995) study, 91% of the teachers reported the need for a course on the structure of language for all teachers who have the responsibility of teaching reading and writing. Current pre-service teacher education programs do not offer this kind of course. Thus, Moats (1995) and Morris (1985) argue that pre-service teacher education faculties do not adequately prepare teachers for the classroom demands of teaching reading and writing.

Afflerbach (1998), Allington and Cunningham (1996), Fullan (2000) and Winograd and Arrington (1999) suggest that good teaching cannot exist without valid assessment. Fuchs, Fuchs, Hamlett, Phillips and Bentz (1994) estimate the range of instructional levels in some general education classrooms may be as high as five grade levels. A study by Housego (1994)

of 16 graduates from the University of British Columbia's Faculty of Education Two-Year Elementary Program in their first year of teaching reported findings of the student teachers' Feelings of Preparedness to Teach Scale (PREP Scale). As a group these beginning teachers found themselves least prepared for assessment tasks. According to a survey by Rogers (1991) of education students in 33 Canadian universities, approximately 60% of prospective elementary teachers and three-quarters of future secondary school teachers did not complete an undergraduate measurement and evaluation course prior to beginning to teach. Moreover, he argued that there is a conflict between the focus of the topics in measurement courses and the need to emphasize topics and activities related to classroom assessment. "Assessments in the classroom are designed to support instruction and are feedback about the student's rate of learning" (Rogers, 1991, p. 181). Fullan (2000) states that instruction and assessment interact with each other so teachers can produce better results from their teaching. However, Rogers (1991) suggests that policy makers are not aware of the lack of assessment education in the pre-service teacher education, and stresses the importance of a uniform policy for assessment courses in all teacher education programs.

Thus, researchers are arguing for the need for a more intensive focus on educating student teachers on the theory of reading development, on the linguistic structure of language, and on assessment of reading progress. With this knowledge classroom teachers will have the ability to flexibly modify programming across the wide range of developmental stages in reading among their students and, as a result, their students will receive more effective instruction in literacy skills.

Summary of the Role of the University in Pre-Service Teacher Education

There is increasing evidence for the need to improve literacy instruction in the primary grades. It is argued that improvement in literacy instruction will occur with more compulsory research-based courses in pre-service teacher education programs on reading development, on the structure of language, on different teaching strategies for different levels of reading development, and on ongoing assessment of the acquisition of literacy skills.

Practice Teaching

The role of the university in pre-service teacher preparation has been examined. Next, the role of practice teaching in teacher education programs will be presented by surveying the research literature for the past ten years. As in other areas of teacher education, a great deal of the literature on practice teaching is based on opinions rather than research data.

In this literature teaching is viewed to be experiential, but this raises the question, "Is experience always the best way to learn?" According to Dewey, "an overemphasis on practice would lead to unquestioned acceptance of cooperating (associate) teachers' techniques and away from the development of reflective inquiry" (quote in Zaborik, 1988, p. 9). Practice teaching has been presented as the "capstone experience in teacher preparation for more than 75 years" (Veal & Rikard, 1998, p. 108). It is viewed as critical in the development of the student teachers' pedagogical skills. "Implicit in this arrangement is the belief that student teachers learn the craft of teaching through guided practice with an experienced teacher" (Rust, 1988, p. 56).

Guidance for the student teacher in the field settings is typically a shared responsibility between the associate teachers and the university supervisors. The teacher whose classroom the student teacher is placed for practice teaching is called the associate teacher. Each student teacher is

also assigned to a faculty supervisor from the university who evaluates the student teacher's instructional skills in the practice setting and who is the liaison between the university and school settings.

The associate teacher must fulfill a number of functions: provide supervision, offer opportunities for the student teacher to learn classroom instruction methods, demonstrate superior teaching skills, and exhibit good interpersonal communication skills. How are associate teachers chosen? Associate teachers are usually appointed by the superintendents for a district, based upon the recommendations of principals of schools. These associate teachers are usually selected for their excellence in teaching, years of experience in the classroom, respect and esteem of their colleagues, and whether they teach the same grade or subject as required in the student teacher's program. While these characteristics represent a basis for describing their effectiveness in the classroom, these characteristics do not assure that the associate teacher will develop a good supervisory, helping relationship with their student teachers or provide modeling for theoretically sound, current, knowledgeable teaching practices.

Much of the field experience of student teachers is supervised by persons who themselves are not well versed in a conceptual understanding of reading processes and methods. Teachers who have taught for over ten years may not be aware of the more current theories and methods for teaching literacy. Tidwell and Mitchell (1994) suggest that "it is not uncommon for teachers to maintain fundamental perspectives and understanding they obtained during pre-service education" (p. 48). Consequently, Manzo (1991) argues that student teachers may rarely observe or practice the well-researched principles of good reading instruction that have emerged in the research in the last ten years or so.

What is the role of the university supervisor? According to Zimpher, deVoss and Nott (1980) the role of the university supervisor is to define and communicate the purposes and expectations to be fulfilled by the student teacher and associate teacher, to decide on the timing of the phasing in of the student teacher's instructional activity in the classroom, and to evaluate, with constructive criticism, the instructional skills of the student teacher. These supervisors must be actively present during the time of the student teaching in order to challenge the student teacher's beliefs and practices and also to model teaching practices and ideas.

How are university supervisors chosen? According to Kagan, Freeman, Horton and Rountree (1993) the university supervisor in the USA is often a graduate student with limited teaching experience who completes evaluations of the student teacher, assigns a grade, and often requires the student teacher to demonstrate particular instructional strategies in the classroom.

Therefore, supervision is highly individualistic. Student teachers complained of the lack of theoretical knowledge in their particular subject by the university supervisor. Rust (1988) states the essentials of "good" pre-service supervision is having supervisors with "teaching experience, the ability to reflect on practice, and the ability to talk about teaching" (p. 62).

Knowledge in the student teacher's specific area of teaching would be a great asset in the supervision of teaching practice.

Thus, there is not only a great variation in the quality of supervision from the associate teacher, but also from the university. There is a need to invest in the education of university supervisors and associate teachers. Shared supervision or collaboration between the associate teacher and the university supervisor, where the supervisors met regularly to talk about their work with the student teachers, and the involvement of associate teachers in the planning of the teacher education program would provide the necessary strong link between the research, theory-based

course work of the universities and the real world of the schools (Rust, 1988). The university supervisor is the acknowledged expert on the research of effective teaching and the associate teacher is the acknowledged expert on effective practice within the context of the classroom.

For teacher education it seems essential to bring supervisors together regularly to talk about teacher development and its implications for their work with the student teachers. In these discussions, supervisors would talk about their work, and so begin to break down the barriers of “privatism that isolate so many fine teachers” (Rust, 1988, p. 63). Kagan et al., (1993) perceived these discussion groups as part of in-service education which would enrich the professional lives of the teachers, by giving them an opportunity to share their perceptions, beliefs and practices, and broaden their responsibilities.

As Nolen et al. (1990) state, the responsibility for our children’s literacy belongs not just to the classroom teacher but also to the institutions that educate and certify them, “The time has arrived for teacher educators, policy makers and school district personnel to work collaboratively to promote sound and research-based inductive programs” (Reiman & Thies-Sprinthal, 1993, p. 179). The associate teacher must be informed and engaged with the ideals of the university teacher preparation program. Associate teachers need to view themselves as teacher educators.

Summary of Practice Teaching

Reviews on supervision of practice teaching demonstrated a large variation in the competency of the associate teachers and supervisors with the resulting large variation in the practice of effective instructional techniques by student teachers. With the tremendous growth in research on the principles of good literacy instruction in the past 10 years the likelihood of a student

teacher being placed with an associate teacher who is aware of this research is small. Therefore, the need for research-based in-service PD on effective literacy instruction becomes important for new as well as for experienced teachers, especially for those teachers interested in becoming associate teachers.

There has been growing recognition in the last ten years of the need for reform in the preparation of our teachers. School-university partnerships have been central to many reform proposals such as the professional development schools. In Canada a range of alternative models for the delivery of the one-year Bachelor of Education program are emerging at the Faculty of Education at the University of Toronto (FEUT). The following section describes one of the pilot projects funded by the Ministry of Education in Ontario for improving pre-service teacher education.

Canadian Research in Pre-Service Teacher Education

At the largest teacher education faculty in Canada, OISE/University of Toronto, there was dissatisfaction with the current eight-month Bachelor of Education program for preparing our future teachers. It was felt that the program was not coherently planned for either the teacher candidates or the teacher educators. The teacher candidates were not able to see the connections between the varied courses at the university and their various field experiences in the practicum component of the program. It remained a challenge for the individual candidates to integrate the theory, skills and content from the faculty courses with the practice provided in the school field work (Rolheiser, 1999). Similarly, there was a lack of cohesion among the courses offered by the university. Any team of instructors could design their courses, so many aspects of the pre-service program could be neglected or shortchanged. Furthermore, individual faculty members could teach courses without the knowledge of the content that other instructors were

focusing on. Therefore, there could be overlap or omission of important components in the program. Finally, the researchers at OISE/UT felt there was also a lack of consistent outcomes to guide the program planning and to measure what teacher candidates should know and practice as teachers (Rolheiser, 1996).

Therefore, in 1995, the Faculty of Education in the University of Toronto began a pilot project based on a two-year model of teacher education. The pilot project involved a partnership agreement between the university-based teacher educators and school-based teacher educators in 11 professional development schools. To ensure a voice of all partners in the decision making of the program the Academy Advisory Committee (AAC) was formed. The members consisted of a program coordinator from the university, a field based coordinator from the schools, three university-based teacher educators, the principals from the 11 schools, a teacher liaison from each school and three elected teacher candidates. The goal of this committee was to ensure all partners were part of the planning and development of the pre-service program.

There are differences between the cultures of the school and university environments. In the school culture the primary client is the student, whereas in the university culture the focus is on the needs of the teacher candidates. Therefore, it was a challenge to work together so all parties could collectively change teacher education (Rolheiser, 1999).

The focus of the pilot project was collaboration in the design and delivery of the two-year pre-service program. The faculty members from the university and the teacher candidates worked in an apprenticeship relationship with school administrators, associate teachers and mentor teachers. Teacher candidates remained with a cohort of peers through the duration of the two-year program so there was a sharing of experiences and a support group. School staff also

worked together to support the teacher candidates with a larger number of staff participating in a variety of ways rather than the traditional 'classroom teacher/student teacher' relationship. This project attempted to build a connected, coherent, mutually developed program by all people involved in the education of teacher candidates.

Specifically in the area of literacy, the school districts strongly expressed the need for better prepared teachers (interview with Dr. Rolheiser, September 26, 2000). Therefore, a subcommittee of the AAC was formed to struggle with design components of the instruction of literacy in the two-year program which would better prepare teacher candidates for the challenges of teaching literacy. A group of university professors, a teacher, a principal, and a field coordinator formed this committee. From this committee textbooks were suggested for the curriculum. However, there was a range of philosophies about the teaching and learning of literacy. It was difficult to merge the spectrum of views. As a result, there was a constructivist-driven school-based literacy project used to develop teacher candidates' inquiry skills in the first four months of the pilot. This component was followed by a number of skills-oriented workshops focused on comprehension and word decoding skills. Therefore, there was an initiation of some of the components of effective reading instruction discussed earlier.

Other changes to the pre-service program included the assessment of the teacher candidates, with more collaboration between the teacher candidates, the field personnel, and the university personnel. The goal was to provide the teacher candidates with more ownership for their own development as teachers.

An evaluation of this two-year pilot project was very positive (Watson, Hart, & Jacka, 1997). The teacher candidates' self assessments from the two-year program were higher than the teacher candidates in the traditional eight month program in terms of overall preparedness to teach. The two-year teacher candidates felt the second year provided time for the consolidation and deepening of skills and knowledge. The additional field experience was valued by the two-year students.

During the pilot project the direction of pre-service education in Ontario was in a state of change. A political decision ushered in the merging of pre-service teacher education and graduate studies or professional development, so the Faculty of Education and the Ontario Institute of Education became OISE/University of Toronto. This merge caused a stage of disorientation within the two faculties, and currently further experimentation in the design of the pre-service program has plateaued (Rolheiser, 1999). Reform can be threatened by political change and inertia. However, in September 1999 OISE/UT extended the time of its pre-service programs, and it endorsed the proposal for all programs to be ten-months in length. In addition, an alternative two-year teacher certification program emerged resulting in two new Masters of Teaching degrees, one focusing on elementary teaching and one focusing on secondary school teacher preparation. There remains the challenge of redeveloping the extended 10-month program with the successful components from the two-year pilot project.

Summary of Canadian Project at OISE/UT

The research at OISE/UT acknowledged many of the problems discussed in the US literature and Canadian literature on pre-service teacher preparation. The pilot project implemented many of the solutions expressed in pre-service research. There was a more coherent, in-depth program

with strong collaboration between the personnel in the schools and the universities. It demonstrated the successful implementation of theory into practice.

There is a desire to change teacher preparation. The partnership between the school districts and the university highlighted the urgent need for better preparation of teacher candidates in the instruction of literacy. The recognition of this need started a dialogue between school and university educators on effective literacy instruction. There is a movement towards improving teacher preparation in the instruction of literacy. Since pre-service education lays the foundation of instructional skills, it is invaluable for improving literacy instruction in our schools.

Currently, teacher education is viewed as a “continuum of lifelong learning, with school districts and universities working in concert to support learning from pre-service through induction, and continuing into in-service and leadership” (Rolheiser, 1996). The merging of the Faculty of Education and OISE at the University of Toronto is a concrete example of this view. The following section focuses on the next part in the journey of a teacher’s learning cycle, in-service professional development.

3. In-Service Professional Development

The nature of in-service PD encompasses a broad range of approaches. In-service teacher education can be a one-hour workshop, a full day workshop, a couple of days of a workshop or an ongoing program throughout a year or more. Not only is there a large variation in the types of in-service PD, but also there is a large variation in the participants. In-service PD involves teachers who are learners with varied teaching experiences and thus, have existing beliefs and schemas about teaching. In pre-service teacher preparation, the teacher candidates are learners who do not have the experience of teaching and thus, have fewer beliefs and schemas about teaching. In contrast, in-serviced teachers, because of their more extensive experience, may feel threatened by ideas that conflict with their own practices. As a result they may be defensive, feeling that acceptance of the new theories and findings would be an admission of a lack of knowledge and competence. Therefore, they may not accept new information on teaching practices as readily as pre-service candidates.

In this regard, longer-term professional development provides the necessary time for experienced teachers to synthesize new knowledge and skills with their existing belief systems. Hargreaves' (1995) dimensions in teaching and Mevarech's (1995) stages of learning during professional development will be discussed in order to better understand the perceptions and reactions of experienced teachers to in-service PD.

Another area of interest in the literature of in-service PD is the problem of transferring the learning from the in-service staff development into actual classroom instruction. Most in-service PD lacks coherent presentations of knowledge and skills over an extended period of time. Research supports the need for certain key elements within a professional development program to ensure that the transfer of learning occurs between the presentation of new ideas and

skills in the in-service PD and the implementation of these ideas and skills in the classroom. These elements include: theory-demonstration-practice-feedback-coaching, linking change in teacher practice to successful student outcomes, teacher support groups, increasing teacher self efficacy, goals or scope of change, and strong school leadership (Bennett, 1987; Gersten & Brengelman, 1996; Guskey, 1986; Joyce & Showers, 1988; Little, 1990; Showers, Joyce & Bennett, 1987). Research supporting the incorporation of these elements to professional development will be reviewed.

The focus of this thesis on in-service PD arises out of the dilemma that if teachers are entering the teaching profession from pre-service education ill-equipped in the area of literacy instruction, and if teachers who are serving as their models in the field also lack an up-to-date understanding of the processes underlying literacy acquisition and the factors contributing to difficulties in learning to read and write, then it is clear that in-service PD for both new and experienced teachers is essential.

What is Professional Development?

Fullan (1995) describes professional development as “the sum total of formal and informal learning pursued and experienced by the teacher in a compelling learning environment under conditions of complexity and dynamic change” (p. 265). Professional development is continuous learning and is not an add-on or episodic event. It is “the essence of teaching and learning to teach better” (Fullan, 1995, p.257).

Education is a professional field where change is constantly occurring since “the old factory system of mass production and standard schooling systems is being replaced by flexible patterns of teaching and learning to prepare a workforce for flexible technologies in smaller units of

enterprise” (Hargreaves, 1995, p. 12). With these changes in education, it is becoming more important for teachers to remain informed of new developments in research on learning.

In the literature there are at least two different perspectives on professional development. Professional development can be viewed within a deficit or a growth model. For many reformers, pre-service teacher education is incomplete, and thus there is something lacking which needs to be corrected. Professional development or in-service teacher education fills the lack of knowledge by providing workshops or programs to teach the required knowledge. This is the deficit model. However, it is suggested in the literature that we are expecting too much from pre-service education, which is brief and has many of the previously discussed limitations. Furthermore, pre-service candidates lack teaching experience and thus, they may not fully understand all the information taught in the courses presented during their pre-service training. Therefore, this more traditional perspective of in-service teacher education is being supplanted by the growth model of professional development. Even the term, professional development is replacing the term in-service teacher education in current literature.

In the growth model, professional development is viewed as an integral part of the career life cycle of teachers. There is qualitative research on the career cycles of teachers which chronicles changes in teachers’ sense of competency, in satisfaction with their work, and in their motivation to improve their instructional goals (Huberman, 1995). These identified phases may impact on the successful outcomes of professional development. In this perspective, professional development is viewed as a self directed journey to find more meaning in one’s work. Thus, professional development is an ongoing, daily occurrence with teachers, as they inquire and reflect upon their teaching practice and assess the impact of their instruction on their students. In this model professional development can be a structured effort to keep

professionals aware of the research advances in their professional field so their teaching skills remain current. As Allington and Walmsley (1995) suggest, there is never any one “right” way to teach, just as there is no “quick fix” to teaching literacy so a teacher has to find the knowledge and skills which provide the opportunities for learning in their particular classroom for specific groups of children.

Supporting this growth model of professional development are papers describing teaching as always changing. Researchers such as Fullan (1995) state, “new knowledge, new ways of knowing and learning and global interdependencies are changing all the time in unknown ways” (p. 254). Consequently, Smylie (1995) argues that if schools are to improve, with better instruction for students, then schools are not only places for teachers to work, but also places for teachers to learn. Professional growth must be valued and expected (Joyce & Showers, 1988). It should be ongoing, and not a specific event (Guskey, 1998).

Lack of Research-based Teaching Methods in Classrooms

According to Joyce and Showers (1988) teachers should be expected to experiment with the content and process of teaching during their careers. As Monroe and Smith (1985) state, “regardless of how adequate a pre-service teacher education may appear, it cannot prepare the teacher for all the on-job changes and technological advancements which seem inevitable” (p.123). However, it is rare for research evidence to find its way into the classroom although educational researchers are constantly gaining new insight about learning processes and teaching (Gersten & Brengelman, 1996; Richardson, 1994). It is suggested that teachers must keep abreast of their expanding and emerging knowledge base in order to be prepared to refine their conceptual and instructional skills.

How do teachers learn about current research in education? Of the 14 sources of learning identified by the National Education Association, 1,789 teachers in the United States chose classroom experience, consultation with and observation of teaching colleagues and individual study as the most effective learning sources. In-service PD was rated as the least effective source of learning for these teachers (Smylie, 1989).

There are problems with the sources of learning chosen as most effective by these teachers. For instance, there is a restriction of new learning by relying on direct experience. Classrooms are complex and dynamic environments and it was found that teachers engaged in over 1,000 exchanges with students in a day (Smylie, 1989). Therefore, they cannot be aware of everything that occurs in their classroom. It is difficult for teachers to gain a perspective on their work, as there is little time for reflection on their teaching practices, and furthermore, trial and error is the predominate means of learning. Their interpretation of direct experience may just confirm ill-formed judgments. The interactions with colleagues may serve to confirm current beliefs rather than inject new knowledge to challenge and enhance present practice. The physical isolation of teachers and the lack of time limit the frequency of interactions with peers. This is compounded by the prevailing norms of autonomy and status, which may further constrict work-related requests and offers for assistance (Smylie, 1989).

Consequently, most teachers use a narrow range of practices and only expand this repertoire with the provision of well designed education (Joyce & Showers, 1988). Thus, Joyce and Showers (1988) argue that the learning benefits to students are so great with ongoing teacher development that “the failure to create a strong staff development system is a tragic dereliction” (p. 27). However, there are 17,000 school districts in the United States and few have education programs that ensure changes in the learning outcomes of students (Joyce & Showers, 1988). In

fact, school districts in the United States spend half of 1% of their resources on staff development (Darling-Hammond, 1996). Nevertheless, high quality staff development is a central component of many reform suggestions for improving education. “Ultimately, it is the teacher who must change if a reform is to take place” (Richardson, 1998, p. 314). For a teacher to change through professional development, according to Mevarech (1995) “a process of change in teachers’ mental models, beliefs and perceptions regarding children’s minds and learning” must take place (p. 152). It is suggested that this kind of change takes time and effort as old values and beliefs which were the basis of their teaching practices for many years have to be weakened if new beliefs about children’s learning are to be accepted as the basis of the new instructional practices.

Four Dimensions in Teaching

Hargreaves (1995) describes four different dimensions in teaching (summarized in Table 4). He suggests that successful professional development will assess growth in the teachers’ knowledge of subject matter, pedagogy, and subject-matter pedagogy. Then, with a richer knowledge base, teachers come to hold new beliefs and develop new ways of thinking about learning.

Knowledge and technical competence is the main focus of most professional development. According to Hargreaves (1995) this focus is not sufficient, since more in depth knowledge of subject matter and improved instructional skills is often not utilized by teachers because the professional development does not address the emotional impact of teaching, moral purposes of teaching, and the political aspects of teaching in our society. “Even when new techniques have demonstrable merit, education in them may be ineffective when it does not address the real conditions of teachers’ work, the multiple and contradictory demands to which teachers must respond, the cultures of teachers’ workplaces, and teachers’ emotional relationships to their teaching, to their children, and to change in general.” (Hargreaves, 1995, p. 26). Hargreaves

(1995) suggests that emotions are pivotal to the quality of teaching, as when burnout and cynicism occur with the attendant reduction of joy in teaching.

Table 4 Dimensions of Teaching

1. Moral Purpose in Teaching	- the importance of high quality instruction and high moral tone in schools
2. Political Awareness	- reflection about the social conditions of one's work as a teacher, and learning about the power structure for the allocation of resources in order to be empowered to assist students and parents
3. Emotions	- "Without desire teaching becomes arid and empty. It loses its meaning" (Hargreaves, 1995, p. 21) - find ways to positively cultivate teachers' engagement to their work
4. Skills and Knowledge	a) Pedagogical Knowledge - overall knowledge about teaching and learning b) Subject Matter Knowledge - relationships between knowledge about a certain subject and instructional strategies c) Pedagogical Content Knowledge - extensive repertoire of strategies and the ability to adapt these strategies to meet the learning needs of specific students d) Knowledge of Curriculum and Curricular Materials - in depth knowledge of curricular materials and how it is organized across all grades in the school system

Note. Hargreaves (1995) and Borko and Putnam (1995)

Thus, Hargreaves (1995) argues that all the dimensions of teaching must be addressed in order to attain lasting, ongoing change in teachers' beliefs and practices: skills and knowledge, moral purposes, political awareness and emotional involvement in teaching. "With the setting of school learning goals, ongoing collaboration to stimulate desire and passion, awareness of political conditions such as the imposed and escalating demands on teachers, and with increased mastery and competence of knowledge and skills, then professional development will improve the progress of students' learning" (p. 34).

Transfer of Learning

The history of staff development is characterized by a lot of criticism (Guskey, 1986).

According to a review by Wilkinson (1994), many professional development days for teachers were spent listening to inspirational lectures on improving instruction but there was little impact on their teaching practices. Research has examined the reasons for the ineffectiveness of in-service education (Darling-Hammond, 1996; Guskey, 1986; Showers, Joyce & Bennett, 1987; Showers, 1990).

One-shot workshops did not have any follow-up where there were opportunities for ongoing work on specific classroom problems, and there were no opportunities for practice and feedback of newly acquired knowledge. When only intellectual content was provided in the in-service program with no specific procedural component for actually demonstrating strategies of what to do, there was a low level of change in the classroom practices of teachers (Joyce & Showers, 1988). Therefore, Darling-Hammond (1996) argues that administrators need to examine their priorities for staff development since budgets cannot sustain both intensive, ongoing professional development and numerous one-shot activities. If staff development programs are “systematic attempts to bring about change – change in the classroom practices of teachers, change in their beliefs and attitudes, and change in the learning outcomes of students,” then funding has to support the most effective means to attain these changes with ongoing professional development (Guskey, 1986, p. 5). There is a realization of the profound differences between short-term implementation and sustained use of newly acquired knowledge in the classroom (Darling-Hammond, 1996). How is the knowledge acquired in professional development transferred to the day-to-day classroom instruction?

There is a growing recognition of problems with transfer of education. “Transfer of education is defined as the degree to which knowledge and skills acquired by education are effectively applied in the workplace of the school or classroom” (Veenman, Van Tulder & Voeten, 1994, p. 304). Some educators in professional development automatically assumed that once teachers developed the skills, the skills would automatically be transferred into classroom instruction. Some service evaluation reports were simply statements of the participants’ satisfaction with the workshop and there were few systematic studies examining the effectiveness of in-service education programs on student learning outcomes. The cost of research examining the outcome of in-service professional development was a barrier.

Research examining the transfer of training should ask the following questions: “Did participants use new skills in the classroom, did they use them appropriately, did they integrate new skills with existing repertoire, was there long-term retention of the products of training?” (Joyce & Showers, 1988, p. 70).

According to Darling-Hammond (1996), teachers engage in staff development because they want to become better teachers which means enhancing the learning outcomes of their students. However, teachers often felt a sense of frustration after attending an in-service PD course as they were unable to use the new teaching activities or new curriculum materials to improve their instruction in the classroom (Bell & Gilbert, 1994).

A synthesis of research on the transfer of new skills to classroom practice has demonstrated the importance of several key features in the professional development of teachers. These features are summarized in Table 5.

Table 5 Features in the Successful Transfer of New Skills to Classroom Practice

KEY FEATURES	RESEARCH EVIDENCE
<p>1. Theory-Demonstration-Practice-Feedback-Coaching</p> <p>Bennett, 1987 Joyce & Showers (1988) Showers, Joyce & Bennett (1987)</p>	<p>- .39 effect size with theory-demonstration-practice - 1.68 effect size with the added component of coaching - 80% of coached teachers transferred new strategies to active teaching practices</p>
<p>2. Feedback: Linking Change of Instructional Practice with Successful Student Outcomes</p> <p>Guskey (1986) Showers (1990)</p>	<p>- student promotions rose from 34% to 95% over 2 years - ITBS- student scores rose from below average to average and above by the second year of the study</p>
<p>3. Teacher Support Groups</p> <p>Fullan (2000) Gersten & Brengelman (1996) Gersten, Morvant & Brengelman (1995) Guskey (1995) Joyce & Showers (1988) Mevarech (1995) Showers, Joyce & Bennett (1987) Sparks (1988) Smylie (1995) Veenman, Van Tulder & Voeten (1994)</p>	<p>-introduction of new skills causes anxiety and discomfort among teachers so teachers need support to persist during the beginning stages of implementation of innovative skills (qualitative studies) - in a regression analysis teacher support groups was one of the significant variables ($p < .001$) predicting change in classroom practice</p>

Table 5 (cont'd)

KEY FEATURES	RESEARCH EVIDENCE
<p>4. Teacher Self-Efficacy</p> <p>Sparks (1988)</p>	<p>- a qualitative study-nonimproving teachers showed lower self efficacy ratings</p>
<p>5. Goals: Scope of Instructional Change and Its Relations to Practicality in the Classroom</p> <p>Gersten & Brengelman (1996) Grant & Mindell (1989) Guskey (1986) Hargreaves (1995) Smylie (1989) Sparks (1988) Tillema (1995)</p>	<p>- "Reality Principle"- there was more teacher change when the in-service provided practical instructional ideas for the classroom</p> <p>- significant correlation between ratings by teachers of importance of the in-service education and change in classroom practice ($r = .68$)</p> <p>-goals of professional developers and teachers should be created jointly</p>
<p>6. Strong School Leadership</p> <p>Joyce & Showers (1988) Smylie (1989) Veenman et al. (1994)</p>	<p>- function of the principal is a significant variable in predicting the impact of professional development on classroom practice ($p < .001$)</p>
<p>7. Effects of School Context</p> <p>Fullan (2000) Little (1990) Placier & Hamilton (1994)</p>	<p>- qualitative studies - culture of learning predicts more successful application of new teaching skills in the classroom</p>

Note: ITBS is the Iowa Test of Basic Skills

1. The Theory-Demonstration-Practice-Feedback-Coaching Model

There are many different ways to design a professional development program. An exploration of theory through discussions, readings and lectures is necessary to understand the rationale behind a skill or strategy. However, Joyce and Showers report (1988) that increasing amounts of time on education did not impact the transfer of new teaching skills to the classroom. The

effect size was .00 for implementation of new skills in classroom instruction when more time was spent on instruction of theory. Instruction of theory with the added features of demonstration, practice and feedback within the education site substantially increased the amount of change observed in the classroom to an effect size of .39. The four features of theory, demonstration, practice and feedback appear necessary to develop the levels of cognitive and interactive skills for a new teaching skill to become incorporated into teaching practice in the classroom (Showers, Joyce & Bennett, 1987). Teachers have complained about the overemphasis on theory in their education, but without a clear understanding or grounding in the theory of a new teaching strategy teachers would be unable to use the new skills and strategies “in anything but a superficial manner” (Joyce & Showers, 1988, p. 74).

Understanding the theory underlying the new content knowledge and instructional skills enables teachers to apply a new strategy flexibly and appropriately in multiple situations in the classroom without a prescribed, step by step approach to teaching. However, theory and demonstrations alone are not enough to produce internalization of the education, unless teachers also have opportunities to practice in the education setting with feedback from peers and staff developers. Therefore, theory-demonstration-practice-feedback enabled the teachers to reach levels of skill to sustain practice in the classroom. Showers et al. (1987) estimated that teachers had to practice the new teaching strategy at least 25 times in their classrooms before the use of the new skill was automatically part of the teaching practices of the teacher.

Skills developed during training were not sufficient to sustain practice in the classroom. The teachers in their study required social support as they practiced their new skills. Showers et al. (1987) called this fifth component of an in-service program coaching. “The coaching relationship is simply a partnership in which two or more people work together to achieve a goal. Visiting one another as they practice, they learn from observing the other person and

particularly by watching the students' responses to the cognitive and social tasks presented to them" (Joyce & Showers, 1988, p. 94). Coaching occurs after the initial presentation of theory, demonstration, practice and feedback during the practice sessions at the training site. With coaching, a peer support group is formed within the school setting. The main purpose of coaching is to aid with implementation until automatization of the new skills occurs in the classroom. With coaching, the effect size of the degree of transfer of learning to classroom teaching practice increased to 1.68 (Bennett, 1987; Joyce & Showers, 1988).

In Showers (1990) study, teachers learned alternative teaching strategies through theory, demonstration, and practice. Next, half the teachers were assigned to "coaching" follow-up training and the remainder were controls. The coaching program included organizing the teachers into self-help teams. In addition, teachers in the coaching program watched their colleagues teach new strategies so they learned from each other. All participants were observed for one year by the researchers. The coached teachers exceeded their uncoached peers in the implementation of new strategies into classroom practice, since 80% of the coached teachers transferred the new strategies into their active teaching practices compared to 10% of the uncoached teachers.

2. Feedback: Linking Change of Instructional Practice with Successful Student Outcomes

A follow-up support structure of teachers meeting to discuss and to examine the learning outcomes of their students after the implementation of new teaching strategies is thought to be an important component in planning professional development. According to Guskey (1986) changes in teachers' beliefs and attitudes occur after changes in student learning outcomes occur. Thus, it would seem that a key factor in the endurance of any change in teaching practice is successful learning outcomes of a teacher's students. When teachers see their students

attaining higher levels of achievement, they will become more involved in the instruction or more confident in their ability to learn the new skills and then, changes in their beliefs and attitudes will occur.

To support the importance of feedback to sustain teachers' change in teaching practice, Guskey (1986) examined a PD involving 52 teachers who were trained to use mastery learning. Of these 52 teachers, 34 teachers used the new procedures in their classes during the first school semester following the training and they gained evidence of improved learning of their students. Higher grades on examinations were attained in those class sections where mastery learning procedures were used compared to classes where the mastery learning techniques were not employed. Ten teachers used the mastery learning procedures and found no difference in the learning outcomes of their students, and eight trained teachers never attempted to use the mastery learning techniques in their classrooms. When measures of change in different beliefs and attitudes were compared among the various teachers, teachers who used the trained techniques and gained evidence of improvement in learning outcomes expressed more positive attitudes toward teaching and they felt a greater responsibility for their students' learning. They felt their teaching had influenced the learning of their students. The other 18 teachers, who did not perceive changes in the learning outcomes of their students, did not perceive their teaching as influencing the learning of their students.

Therefore, a professional development program which links changes in teaching to changes in student learning will assist in sustaining change in teacher practice.

3. Teacher Support Groups

As previously seen in the Joyce and Shower's (1988) research, collaboration and increased communication between teachers and between teachers and administrators were associated with better student learning outcomes. Research has supported the impact of support groups of teachers on the learning outcomes of their students (Gersten et al., 1995; Guskey, 1986; Sparks, 1988; Veenman et al., 1994).

Teachers need the support of colleagues during the introduction of a new set of teaching skills since learning new skills creates discomfort and anxiety among the teachers. First, teachers are concerned that any change in instructional practice may result in students learning less than they do under the present set of instructional practices. Secondly, during the initial stages of implementing innovative practices, teachers' teaching performances will be more awkward and not always effective.

A qualitative study by Mevarech (1995) reported teachers progressing through five learning stages as they attempted to implement new practices introduced during professional development. She interviewed and observed 145 experienced and novice elementary teachers. Ninety-five of these teachers participated in a staff development program for the implementation of a Computer-Assisted program and 50 others were part of a nontreatment control group. In the initial stage when teachers first began to use the new teaching method, there were feelings of frustration, skepticism and even withdrawal from implementing the innovative method. Experienced teachers felt like novice teachers, and they became more concerned with classroom management, focused on physical changes in the classroom, and showed mechanical use of the innovative skill. Their teaching skills became poorer. Mevarech (1995) described this stage as "Survival," since the teachers required affective support and step-

by-step guidance, in order to be motivated to continue with implementing the innovation despite the encountered difficulties. “As teachers discuss, problem solve, and develop alternative strategies, they are able to merge practice knowledge with research knowledge” (Gersten & Brengelman, 1996, p. 69). These discussions help in the difficult period between the letting go of old practices and the grasping of new practices. The need for support “coupled with pressure at this time is vital for continuation” (Guskey, 1998, p. 123). As one teacher stated, “it is a relief to find we have common problems” (Wilkinson, 1994, p. 70).

Teachers who survived the entry stage, moved into the “exploration and bridging” stage in Mevarech’s (1995) model. At this stage the teachers approached the new method more positively and explored the differences between the new method and their own instructional strategies and pedagogical theories. The teachers were preoccupied with their own theories and concepts, rather than examining the effect of the new method on their students’ learning. There remained a routine use of the new method, with little thought to adapting the innovation to the needs of their students. At this stage of exploration the teachers needed support and ongoing guidance to help bridge their own teaching experience with their new experiences with the instructional innovation.

At the “adaptation” stage, the teachers began to examine the innovation and adapt the new method to the needs of their students (Mevarech, 1995). The teachers started looking outward towards the learning outcomes of their students. At this stage the teachers needed ongoing assistance on ways to make changes in the professional development program but at the same time maintain some fidelity to the innovative instruction.

Mevarech (1995) interpreted the next state as “conceptual change” stage since the teachers assimilated the new method into their existing instructional repertoires and developed more questions and, thus wanted more understanding of the innovation. There was more reflection on their teaching, and speculation on the causes for changes in the students’ learning. Few teachers reached Mevarech’s (1995) next stage of “invention” where teachers used their “reconstructed pedagogical knowledge to experiment with new materials, new assignments, or different ways of teaching” (p. 164). During these last two stages the teachers required ongoing support to make connections between old and new pedagogical knowledge. Therefore, as the teachers progressed through these stages of implementing new knowledge and innovative instructional skills, they required different kinds of ongoing support and collaboration with their peers and the developers of the in-service PD programs.

4. Teacher Self Efficacy

Teachers’ confidence in their ability to make a difference in the learning outcomes of their students affects their motivations to make changes in their instructional skills. A study by Sparks (1988) examined the effects of teacher efficacy and the perceived importance of the new strategy on the use of practices recommended in the education program. Self-efficacy is teachers’ confidence in their own ability to affect learning outcomes of their students.

There were 19 teachers in Sparks’ study (1988) who were trained in Stalling’s Effective-Use-of-Time training, which helps teachers increase students’ time on tasks and improve academic interaction between students and teachers. There were three groups of 6 or 7 teachers who attended four weekly workshops. The teachers were observed in their classrooms for three consecutive days after the workshops and the observer coded every teacher-student interaction.

The teachers also completed questionnaires relating to their attitudes and they were interviewed individually to obtain their views on their self-efficacy.

Sparks (1988) found those teachers with a higher level of self efficacy were more likely to take risks and experiment in their classrooms, and thus implement the professional development program. The “improving” teachers indicated that as a result of the training they had experienced a higher sense of control over their teaching environments. However, the “nonimproving” teachers did not implement the newly taught skills, since they had given up on their ability to teach and help students learn. Thus, a low self efficacy and lower perception of the importance of the new strategy may account for the lack of improvement in implementing the knowledge and skills from professional development.

Sparks (1988) suggested that one way to increase self-efficacy is to provide a safe environment where teachers can discuss their concerns, their victories and learn together. In their interviews several teachers stated that they gained confidence to try new strategies from their support group. These data provide qualitative evidence for teacher support groups. Sparks (1988) commented that “instructional support groups are important as powerful vehicles for change by improving teachers’ self confidence for implementing change in their classrooms” (p. 117).

5. Goals: Scope of Instructional Change and Its Relation to Practicality in the Classroom

The amount or scope of change is another factor for consideration in the planning of a professional development program. Programs and innovations that are dramatically different from teachers’ current practices or require teachers to make many revisions are usually not implemented well (Guskey, 1986). Congruence between teachers’ beliefs and the educational content of the in-service PD led to better acceptance and utilization of the knowledge presented

in the PD. In Sparks' (1988) previously discussed study, one of the questionnaire's completed by the 19 teachers was the Ease and Importance constructs. One of the goals of the study was to relate the importance of using a practice with the observed use of the practice. The post training rating of importance of providing academic time correlated significantly with post training academic interactions ($r = .68, p < .01$). Therefore, teachers who rated the practice as important tended to use the practice in their classrooms.

Effective staff development offers teachers practical ideas to enhance learning outcomes in their classrooms (Guskey, 1986; Grant & Mindell, 1989). Useful research must be translated into teaching strategies and procedures that fit into the details of day-to-day classroom instruction. Gersten and Brengelman's (1996) "Reality Principle" stressed the importance of teachers' concern for the feasibility of any proposed innovations for assisting in the learning of students. "The more practically oriented the trainer is, the more the newly acquired knowledge and skills are applied, the more changes at the classroom and school level are achieved" (Veenman et al., 1994, p. 313). Teachers perceived the most effective sources of learning as characterized by knowledge related to practice or "knowing how" (Smylie, 1989, p. 551).

Thus, a trainer of a professional development program must appear competent in the world of the classroom, and focus on the teachers' specific needs and concerns. Close collaboration between program developers or researchers and teachers ensures an appropriate balance between strict adherence to program fidelity and mutual adaptations to the teachers' working environment.

6. Strong School Leadership

Working closely with peers is not typical of schools due to the physical isolation of teachers. Furthermore, teachers' daily schedules limit the time available for the collegiality of peer support groups. Therefore, social reorganization of the school is important for both education and practice to reside in the school setting. For a school to reorganize, it will need strong leadership. Thus, leadership of the principal in a school is critical for the success of good education. Veenman et al. (1994) state, "the principal plays a crucial role in goal setting and planning for application, for monitoring action plans, for using new knowledge and skills, and for setting up follow-up meetings after the education at periodic intervals for further information sharing, problem solving and support of the implementation effort" (p. 316).

In Veenman et al.'s study (1994) questionnaires and interview schedules were developed for individual teachers who participated in an in-service PD. There were 103 individual-based in-service programs surveyed in this study as well as 119 school-focused in-service programs. When a teacher returned a questionnaire the principal of that teacher was also sent a questionnaire. The total sample for the questionnaire part of the study was comprised of 355 teachers and 169 principals who participated in the individual-based PD, and 55 teachers and 35 principals who participated in the school-focused PD. The total sample of the interview part of the study was comprised of 53 school teams and their principals who had participated in the school-focused PD. To examine the influences of design features of the in-service programs and implementation characteristics on the outcome of the in-service PD, several multiple regression analyses were performed. The dependent variables consisted of three levels of impact – classroom-level, school-level and knowledge use – reported by the questionnaires and interviews. Five sets of independent variables were considered as potentially influencing the impact scores: school characteristics, features of the in-service activities, implementation

characteristics, type of in-service activity, whether it was individual or school focused, and the interactions between the type of in-service and the first three dependent variables.

The functions of the principal were found to be a significant variable in predicting the amount of impact of an in-service program on a school, $p < .001$. The principal provides the leadership with resources and the time for staff development.

7. Effects of School Context

The success of a staff development program is dependent on more than the design or content of the program but also on the social context of the school. The cultural norms of a school can either impede or facilitate change. Many factors affect the culture of the school. As previously discussed, the type of leadership provided by the principal is one important factor. Some principals may invoke cultural traditions to reinforce authority and control, and others may invoke a culture of shared values and encourage change.

Other important components of the culture of a school are teachers' assessment of working conditions, teachers' sense of autonomy, the support of the school culture for change and teacher cooperation. These components will affect the amount of change experienced in the classroom by teachers who are participating in professional development. If teachers have low expectation of their performance and that of their students, then they will be skeptical of innovations. Teacher collaboration and collegiality and the nature of the relationship between the principal and the teachers can nurture instructional innovation in the classroom. Smylie (1989) suggested that teachers were more enthusiastic, with greater disposition to improve practice in their schools where they had interaction with their colleagues and principals about their academic work.

Teacher cooperation involves the sharing of the “norm of analysis, evaluation, and experimentation” (Placier & Hamilton, 1994, p. 144). Thus, if the goals of the school staff are to improve student learning then teachers will be motivated to improve their own teaching, and thus improve the efficacy of the school (Little, 1990). In Veenman et al.’s study (1994) exchange of information among the teachers involved in the PD was a significant variable predicting the level of impact of the PD in the classroom ($p < .01$).

According to Fullan (2000), schools with “high professional community” had higher achievement scores in various academic areas (i.e., mathematics, science and social studies). In a school where teachers exchange ideas on an ongoing basis, and collectively feel responsible for student development, successful student learning occurs. An optimal school learning environment provides teachers with opportunities to work and to learn together in an atmosphere of “shared leadership and participative decision-making” (Smylie, 1995, p.105).

Conclusions on Successful Transfer of Learning from In-Service PD to the Classroom

There is no “right answer” for successful professional development as there are powerful influences from the norms, policies and resources in each school, which are constantly changing from year to year. Thus, different schools at different times will require a different “optimal mix” of the critical components of successful professional development (Guskey, 1995, p. 114). For example, change in classroom practice will be the result of changes in the instruction of the individual teacher as well as changes in the overall organization of the school. A debilitating school environment, with no support of change, will stop any efforts to change classroom practice. Thus, there is a need for an optimal mix of individual and organizational processes that contribute to the success of teacher education in a particular context.

Problematic Conditions for Effective In-Service Professional Development

Most teachers know a lot about teaching but their teaching is not communicated or expanded by others. Joyce and Showers (1988) have suggested five conditions as “mitigating against the reflective and collective teaching practice of teachers: teachers work alone, there are loose connections of different areas of specialization, teaching was originally conceived as reflexive so time to prepare for teaching is not built into the assignment of being a teacher, collective decision making or sharing is not built into the work of the teacher, and time and opportunity for academic study to keep abreast of the progress and development of new knowledge is not easily available” (p. 52).

Therefore, principles of good education are compromised in the real world of education. Little (1993) argues that schools cannot accommodate the classroom consultation and peer support recommended by the education and coaching model. Persons designated as coaches or mentors are far outnumbered by the classroom teachers requiring their consulting time. Little (1993) also reports that the lack of consistency and the lack of coordination of effort often found in school districts will limit the results of these larger school in-service programs. Therefore, according to Little (1993), larger-scale school improvements require political will to implement and to support effective staff development. Educational change is a difficult and a gradual process within school systems.

The principles of effective instruction will be beneficial for all students. In-service professional development as well as pre-service teacher education is important to ensure classroom teachers expand their knowledge and understanding of instructional skills for a varied population of students. There are an increasing number of children identified as “at risk” for failure to acquire

literacy skills. The next section describes the potential impact of classroom instruction on the acquisition of literacy skills by children in the “at risk” population.

4. Teachers’ Impact on Children “At-Risk” for Literacy Failure

There are a number of potential risk factors associated with poor literacy achievement. These risk factors include children living in poverty, weak processing abilities, poor classroom instruction and children whose home language is not English (L2).

Teachers’ Knowledge and Classroom Practice

Children whose level of achievement in literacy skills are within the 30 to 50th percentile range compared to their classroom peers are having mild difficulties acquiring their reading, spelling and writing skills. Levine (1987), Willows (1991) and Shaywitz, Escobar, Shaywitz, Fletcher, and Makuch (1992) suggest that children with mild difficulties learning to read, spell and write are only different in degree rather than in kind when compared to children with more severe learning disabilities. This “normal variation” interpretation of written language difficulties explains learning disorders as developmental variation in basic cognitive and linguistic processes (Willows, 1991, p.75). Processing domains distinguishing people who are reading disabled from normally achieving readers, and distinguishing people who are having milder reading problems from normally achieving readers, are the same. Therefore, students with mild learning problems will have similar instructional needs to those of students with severe learning disabilities. This would suggest that if teachers in general classrooms have a higher level of understanding of the range of literacy needs of all students, one would expect a reduction in the proportion of students in their classrooms who would fall behind or fail to acquire reading and writing skills.

However, research shows that classroom teachers are not trained to instruct children with mild or severe learning problems (Afflerbach, 1998; Nevin & Thousand, 1986; Reynolds, Wang & Walberg, 1992; Simpson, Whelan & Zabel, 1993). Therefore, a risk factor, from the child's point of view, is to be placed in the classroom of a teacher who does not know how to properly design a literacy program to meet his or her needs.

A study by Snow, Barnes, Chandler, Goodman and Hemphill (1991) examined the ways in which both home and school experiences affect the literacy achievement of a group of 32 elementary school children from low-income families. Classroom and home environments were rated depending on their level of support for the development of literacy skills. Thus, a low home support environment was described as having a nonliterate mother and father, no provision of informal reading material or experiences, disorganized, dirty, and with emotional hazards. A high supportive home environment was described as having parents who read a lot, who sought out literacy activities, who had very high expectations for their child, who were "house proud," and the family life appeared stable. A low support classroom environment for literacy was described as chaotic, and rigid with a heavy emphasis on discipline, boring instruction, lots of reprimands, lack of variety of reading materials, teacher instructing at levels not commensurate with the child's level of ability. A mixed supportive classroom environment had some direct instruction but it lacked stimulating activities, few teacher-student exchanges, and standard instruction following basal texts with some variety in reading material. The enriched classroom environment had direct instruction, stable routines, challenging learning experiences, a calm, supportive, encouraging classroom climate with enriched literacy activities and discussions. Table 6 shows the percentages of children making the expected gains in reading comprehension from different levels of support in the home and classroom environments.

Table 6 Percentages of Children Making Expected Gains in Reading Comprehension

Ratings of Classroom	Ratings of Homes	
	High	Low
High	100%	100%
Mixed	100%	25%
Low	60%	0%

Snow et al.'s (1991) study suggests that improved classroom instructions can compensate for various risk factors. The term "at-risk" describes a wide range of factors such as poverty, learning disability, non-supportive home environments, English as a second language (L2). This small-scale study supports the view that effective classroom instruction can compensate for risk factors. However, poor classroom instruction may place children at risk despite supportive home environments. As shown in Table 6 (adapted from Snow al. (1991) by Allington and Cunningham (1994), it appears that classroom instruction is a significant factor in the acquisition of literacy skills for the "at risk" population of children. These authors state, "that excellent classroom environments can compensate for less than ideal home conditions, but that ideal home conditions cannot always compensate for very poor classrooms" (p. 161).

From the broad discussion of the weaknesses in pre-service education, the key features of effective in-service PD, and the impact of classroom instruction on children "at risk" for literacy acquisition, the next section narrows the focus to the impact of professional development on the literacy practices of general classroom teachers.

5. Impact of Professional Development on Literacy Acquisition

There have been relatively few studies that directly address the impact of in-service PD on literacy education. This next section describes the implementation of professional development programs to improve the levels of achievement of student learning outcomes specifically in the area of literacy. The focus of this thesis is the improvement of reading and writing skills for all students in general classrooms. Therefore, only those professional development programs that attended to the improvement of literacy skill for all students in classrooms were reviewed, and those programs that targeted only the lowest achieving groups of students were omitted.

Programs such as Success For All (Slavin, Madden, Karweit, Dolan & Wasik, 1991), Reading Recovery (Clay, 1986a, 1986b), The Early Intervention in Reading (Taylor, Short, Frye & Shearer, 1992), and the Boulder Project (Hiebert, Colt, Catto & Gury, 1992) demonstrated significant improvement in the reading and writing scores of low achieving children. However, this thesis is concerned with the impact of improvements in the instructional skills of general classroom teachers through in-service PD. The few studies undertaken to address the impact of PD on the instructional skills of the general classroom teacher are the Four Blocks Literacy Model (Cunningham, Hall & Defee, 1998), a classroom consultation model for reading instruction (Baker, 1977), an in-service PD for reading comprehension (Conley, 1983), a 2-year PD to improve the effectiveness of reading instruction (Miller & Ellsworth, 1985), a follow through project using The Madeline Hunter Instructional Theory (Stallings & Krasavage, 1986), and The Reading Instruction Study (Richardson & Hamilton, 1994). Details of these studies are provided in Tables 7 - 12.

Description of Six In-Service PD Programs for the Improvement of Literacy Instruction

1. The Four Blocks Literacy Model

The Four Blocks model is an instructional framework for the teaching of literacy skills by classroom teachers. The goal of this model is to meet the needs of children without putting them into ability groupings, and it combines the major approaches to reading instruction (see Table 7). One of the important contributions of Cunningham, Hall and Defee's (1998) research is the longitudinal nature of their studies. Teachers taught the following four specific blocks (40 minutes per block) in the daily program which had several levels to meet the needs of all students: Guided Reading, Self Selected Reading, Writing, and Working with Words. The children were assessed throughout each year through observations and conferences with the teacher and at the end of each academic year the children were administered the Basic Reading Inventory as standardized testing. Approximately 100 - 140 children in each grade were included in the data. Their results demonstrated the continuing maintenance of the program in five schools over 6 years, with accompanying improvement in the literacy skills of the students exposed to the instruction of participating teachers.

In another study, Cunningham, Hall and Defee (1998) compared the scores from the Metropolitan Achievement Test of 557 first graders in Four Block classrooms with cohorts of Grade one students from the previous year who had no exposure to the Four Block model of teaching. The total reading score for the Four Block first graders were significantly better (.001 level), than the Baseline Group of matched students.

It is difficult to evaluate their model with respect to instructional skills acquired through professional development because their research did not describe the nature of the professional

development implemented to educate the teachers in their project. The focus of the research seems to be more on the use of a specific time management framework provided by the Four Blocks and on the use of specific teaching strategies within the blocks. Furthermore, there were no control groups in the longitudinal study to compare the efficacy of this program on the learning outcomes of these students. However, the results of this longitudinal study provides evidence of the sustained use of literacy skills in the classroom practice of teachers after their participation in professional development.

Table 7 Description of Four Blocks Literacy Model

Brief Description of Study	Key Features of the In-service PD	Findings and Conclusions
<p>Study 1 N = 6 groups of 100 -140 students per grade No Control Group Grades: 1 and 2 Duration: longitudinal - 6 years PD of Teachers: unknown Dependent Measures: Students: Basic Reading Inventory</p> <p>Study 2 N = 557 Grade 1 students Control Group: Cohort analysis as each Grade 1 student matched with Grade 1 student of previous year according to scores on Cognitive Skills Assessment Battery Dependent Measure: Students: Metropolitan Achievement Test</p>	<p>1. Goals - to meet the needs of all children without putting them into ability groups so instruction within each of the Four Blocks was multilevel -combine the major approaches to reading instruction</p> <p>2. Theory -not all children learn in the same way -grouping children according to abilities is not best solution in providing opportunities to learn to read and write - teachers taught the Four Blocks in their daily program : Guided Reading, Self Selected Reading, Writing and Working with Words</p> <p>3. Feedback of Student Outcomes -assessment throughout the year by observation, running records, writing samples - end of year assessment</p>	<p>Findings: Study 1: of the 10% to 15% children who did not read at grade level at the end of Grade 1, half were reading at grade level and above by Grade 2 - standardized tests of these same children in grades 3, 4, and 5 indicated that 90% of these children were in the top 2 quartiles and none were in the bottom quartile Study 2: students taught with Four Block Model scored on the average a half a year better than Baseline Group</p> <p>Conclusions: - a longitudinal study over 6 years provides valuable information on the maintenance of this model both for teacher instruction and improvement of student achievement - difficult to evaluate with respect to transfer of instructional skills from the PD as their research did not describe the implementation of the project through PD of the teachers</p>

2. In-service Training/Classroom Consultation Model for Reading Instruction

Baker (1977) described an in-service program to help teachers develop assessment and problem solving approaches for classroom learning difficulties, to improve teachers' awareness of available resources, to improve teachers' confidence in their ability to produce changes in students' abilities, and to help teachers become more collaborative with their problem solving (see Table 8). The in-service reading program included spelling, examining the "code," reading comprehension, and writing. There were 18 teachers and administrators involved in the study. The design of this in-service included lectures, videotapes demonstrating teaching strategies, modeling behaviour of the consultant, observations of the teachers in their classrooms by the consultant and discussions of the videotapes of the participating teachers' classroom instruction. The teachers attended 10 in-depth workshops during an 18-week period. The in-serviced teachers taught in general classrooms as well as small groups of struggling readers.

The key features of specific goals, teaching theory, demonstration-coaching-feedback and assessment of student outcomes were part of this PD. However, the study did not describe ongoing student assessment. The impact of the PD on the instructional skills of participating teachers was assessed by examining the learning outcomes of 18 underachieving or "at risk" students who were taught by only three of the participating teachers. Although the goals of this research focused on the instructional practices of the general classroom teacher, it only examined the outcomes of "at risk" children who were taught in small groups and it did not focus on the successful transfer of instructional skills learned from the PD into classroom practice. Finally, this program was only 18 weeks long, and ongoing support of teacher groups beyond that period was not described. The maintenance of the instructional skills presented in the PD was not assessed, as there was only one analysis of the literacy outcomes of students at the end of this 18-week period.

Table 8 Description of the In-service Training/Consultation Model for Reading Instruction

Brief Description of Study	Key Features of In-service PD	Findings and Conclusions
<p><u>N</u> = 18 Teachers</p> <p>Control Group</p> <p>Grade: 4 Duration: 18 weeks</p> <p>PD of Teachers: - 10 weeks of in-depth workshops (plus 6 previous workshops) - observations of teachers in their classrooms</p> <p>Dependent Measures: <u>Student Outcomes</u>: - scores of 18 underachieving students who were taught by only three of the participating teachers compared with scores of 18 control students who were not exposed to the instruction of in-serviced teachers were assessed on : Gilmore Oral Reading Schonell Graded Reading List Metropolitan Achievement Test: spelling <u>Teacher Outcomes</u>: written evaluations</p>	<ol style="list-style-type: none"> 1. Goals - to help teachers develop assessment approaches and assist students with written language difficulties 2. Theory -in-service training and consultation by school psychologist of classroom teachers for prevention of reading failure - Piagetian and linguistic-based reading program 3. Demonstration -modeling in videotapes and by consultation 4. Coaching and feedback -videotapes of teachers' performance -observation by consultant of teachers in classrooms -discussions and feedback 5. Assessment of Student Learning and Outcomes and Changes in Teachers' Attitudes 	<p>Findings:</p> <ul style="list-style-type: none"> - the students taught by the participating teachers achieved significantly higher scores on the Gilmore Oral Reading subtests of accuracy and comprehension, $p < .001$, and $p < .005$, respectively and on the Schonell Graded Reading List, $p < .05$ -no significant difference was shown on the Metropolitan Achievement spelling test <p>Conclusions:</p> <ul style="list-style-type: none"> - the research was composed of many of the key features of effective in-service - over a short time span - assessment occurred at the end of the PD, and was not described as ongoing - the literacy outcomes from small groups of underachieving students taught by 3 teachers - teacher outcomes were qualitative comments of changes in their practices and attitudes

3. In-service for Reading Comprehension Classroom Instruction

Conley's (1983) in-service taught reading comprehension instruction strategies to Grade 6 teachers (see Table 9). The teachers met weekly with the project researcher who taught the theory of learning, higher order questioning techniques and applications of strategies to classroom practice. From November until June, these weekly meetings were arenas for discussion and concerns about the application of newly taught strategies in classroom practice. The researcher provided coaching and feedback on the teachers' instructional techniques in their classrooms. The Grade 6 students were assessed throughout the year, as well as at the end of the project in June. Therefore, learning outcomes of the students were available on an ongoing basis. The results from the standardized assessment at the end of the year showed that the students taught by the in-serviced teachers made twice the reading comprehension gains when compared to the control Grade 6 students who were taught by nonparticipating teachers. In the 6-month instructional period, the control group gained 7 months while the experimental students gained 1 year and 5 months in reading comprehension achievement on the Gates-Mac Ginitie Test.

This in-service model incorporated the components of goals, theory, coaching, feedback, teacher support groups, and ongoing assessment of student learning outcomes. However, this research focused on one aspect of literacy, reading comprehension. Furthermore, maintenance of the new instructional skills learned by the teachers from the PD was not assessed over a longer-term.

Table 9 Description of an In-service for Reading Comprehension

Brief Description of Study	Key Features of In-service PD	Findings and Conclusions
<p>Teachers: N = 32 Experimental</p> <p>Students: N = 385 Control N = 595 Experimental</p> <p>Grade: 6 Duration: November to June</p> <p>PD of Teachers: - weekly 30 to 40 minute meetings, readings, observations of teachers in their classrooms</p> <p>Dependent Measures: - formative tests at the end of each of the 5 instructional units on reading comprehension - Pre- and Post-testing with Gates-MacGinitie Reading Comprehension Test</p>	<ol style="list-style-type: none"> 1. Goal - to improve the reading comprehension scores of Grade 6 students 2. Theory-Coaching-Feedback - mastery learning theory, higher order questioning techniques, classroom applications - observations of teachers by consultant - regular feedback regarding their classroom performance 3. Teacher Support Groups - weekly meetings 4. Feedback of Student Learning Outcomes - student assessment throughout the program - evaluation after each instructional unit and standardized assessment at the end of the PD 	<p>Findings:</p> <ul style="list-style-type: none"> -experimental students made twice the reading comprehension gains made by control students: <ul style="list-style-type: none"> - 7 month gain by control group -1 year 5 month gain by experimental group - $p < .01$ <p>Conclusions:</p> <ul style="list-style-type: none"> - this research implemented many of the key features of effective in-service - focused on one aspect of literacy - maintenance of the new classroom practices by the experimental teachers was not assessed

4. Two-Year Program to Improve Teacher Effectiveness in Reading Instruction

The goals of Miller and Ellsworth's (1985) professional development program were to produce measurable differences between nonparticipating and participating teachers' knowledge, attitudes and teaching behaviour and to demonstrate higher reading achievement test scores of students in classrooms of participating teachers (see Table 10).

The implementation of this professional development program was divided into four semesters over a two-year period. There were weekly meetings where theory of learning, vocabulary development, reading readiness, word recognition, reading comprehension, recreational reading, reading engagement and other topics in reading were discussed. During the 2 years, the contacts became less formal after the background knowledge was taught, and the teachers met in smaller support groups where demonstration and practice of teaching techniques occurred. Peer observations as well as classroom observations by the researchers provided coaching and feedback to the teachers. The teachers implemented an individual assessment program in their classrooms so there was ongoing assessment of learning outcomes.

The key features of specific goals, teaching of theory, demonstration of instructional practices, coaching, observations and feedback of classroom instruction, teacher support groups, assessment of student learning outcomes were part of this PD (see Table 10). Qualitative assessment of changes in teachers' attitudes and improvements in teacher knowledge, observations of the level of implementation of the teaching practices taught in the PD, and the standardized assessment of the reading achievement of students in the classrooms of participating and nonparticipating teachers were implemented in this study. Thus, a very thorough analysis of the immediate outcome of the two-year PD occurred. The results from

surveys and interviews of the participating teachers showed changes in their attitudes towards teaching reading and the participating teachers became more knowledgeable about reading instruction. Secondly, a higher level of implementation of desired teaching practices by participating teachers was observed, with significant differences between participating and nonparticipating teachers in four of the six teaching areas assessed. Finally, scores on the reading tests from students in Grades 2 to 5 who were taught by participating teachers increased by an average of 10 percentile points. The reading comprehension scores of students taught for two years by participating teachers were significantly higher than the scores of students taught for two years by nonparticipating teachers, $p < .05$. Thus, the PD had a significant impact on the instructional practices and attitudes of the teachers, as well as the learning outcomes of the students. However, student achievement was only measured in one aspect of literacy, reading comprehension. Maintenance of the instructional practices by the teachers over a longer-term was not assessed.

Table 10 Description of a Two-Year PD to Improve Teachers' Instruction of Reading

Brief Description of Study	Key Features of In-service PD	Findings and Conclusions
<p>Teachers: $N = 47$ participating $N = 96$ nonparticipating Students: $N = 137$ Experimental Group $N = 135$ Control Group Grades: 2 - 5 Duration: 2 years</p> <p>PD of Teachers:</p> <p><u>Semester 1</u> - weekly meetings to instruct teachers on background knowledge and theory of reading</p> <p><u>Semester 2</u> - assessment techniques, implementing an assessment plan in their classrooms, demonstration teaching, supervised trials in classrooms, analysis of performance in classrooms</p> <p><u>Semester 3</u> - supervised trials and analysis, implementation of program and instructional techniques</p> <p><u>Semester 4</u> - motivational techniques</p> <p>Dependent Measures: Teachers: evaluation of plan for classroom, classroom observations by instructor and peers, pre and post attitude inventories Students: California Achievement Test (comprehension subtest)</p>	<p>1. Goals</p> <ul style="list-style-type: none"> - to examine if there is measurable differences between participating and nonparticipating teachers knowledge, attitudes and behaviour - increase level of students reading achievement <p>2. Theory-Demonstration-Coaching-Feedback</p> <ul style="list-style-type: none"> -during the four semesters the teachers were taught background theory, teaching was demonstrated by instructors, teachers were observed by peers and researchers in their classrooms and they were given feedback <p>3. Teacher support Groups</p> <ul style="list-style-type: none"> - weekly meetings <p>4. Student Learning Outcomes</p> <ul style="list-style-type: none"> - ongoing assessment of student learning - standardized assessment at end of 2 year period <p>Teacher Change:</p> <ul style="list-style-type: none"> - evaluate change in attitudes and knowledge and observations of instructional practice in classroom 	<p>Findings:</p> <ul style="list-style-type: none"> Teachers: attitudes changed, and knowledge of reading increased for participating teachers - better implementation of desired teaching practices by participating teachers Students: an average increase of 10 percentile points in reading comprehension scores of students taught by participating teachers, with a significant difference in scores between students taught by participating and nonparticipating teachers, $p < .05$ <p>Conclusions:</p> <ul style="list-style-type: none"> - many of the key features of effective PD integrated into this program - assessment of teacher change: surveyed knowledge acquisition and attitude change and observed level of implementation of desired teaching practices - assessment of student change: focused only on achievement levels of reading comprehension - there was no assessment of the maintenance of teaching skills acquired in the PD over the long-term

5. Follow-Through Project: Madeline Hunter Instructional Theory

The in-service PD program of Stallings and Krasavage (1986) was interesting due to its focus on the maintenance of instructional skills after a period of in-service implementation (see Table 11). The first year of the program was a pilot study. In the second year, 13 teachers participated in the in-service PD consisting of 5 workshops followed by 5 classroom visits by the researchers to observe the teaching practices of the teachers. In the workshops the teachers were instructed on the basic elements in the design of a Madeline Hunter lesson: set, instruction, guided practice, closure, and independent practice. Standardized reading measures were administered to the students at the end of the academic year. There was an increase in the reading scores of the students in the second year of the program, but there was a decrease in the achievement of the reading scores after the third year of the program. During the third year there was instruction in the PD on classroom management, effective directions, monitoring of the classroom, effective grouping. In the fourth year of the program there were four workshops and weekly visits to the schools. Regardless of this support, the reading achievement levels of the students declined. Moreover, the amount of student engaged academic time also decreased. One of the reasons posited by the researchers for the lack of maintenance of the reading achievement scores was that the model itself could not sustain interest. The teachers felt constrained by the lesson design as it was prescriptive and diagnostic. The teachers were bored. Some teachers felt it did not foster the generation or exchange of new ideas. Therefore, the teachers did not persevere with maintaining this instructional practice. Hargreaves (1995) would argue that emotion or teacher engagement was lacking in this in-service professional development program.

Table 11 Description of the Four-Year Madeline Hunter Follow-Through Project

Brief Description of Study	Key Features of the In-service PD	Findings and Conclusions
<p>Teachers: <u>N</u> = 13 Students: <u>N</u> = 208 Grades: Kindergarten - Grade 4 Duration: 4 years</p> <p>PD of Teachers: <u>Year 1</u> - Pilot Study <u>Year 2</u> - 5 one day training sessions to learn the principles of instruction - each teacher had 5 classroom observations and feedbacks <u>Year 3</u> - training in classroom management techniques (no information on training sessions) followed by classroom visits <u>Year 4</u> - 4 days of in-service training and observations in classrooms by request</p> <p>Dependent Measures: 1. Instructional Skills Observation Instrument (ISOI) 2. Time-off-Task (TOT) observations 3. questionnaires and interviews of teachers and principals</p> <p>Student Tests: Stanford Achievement Test or California Achievement Test</p>	<p>1. Goals: - to improve the reading and mathematics scores of students</p> <p>2. Theory- Feedback-Coaching - the basic tenets and skills of the Hunter model were taught - classroom management techniques were taught -the teachers were observed in their classrooms with feedback on their progress</p> <p>3. Assessment of Student Learning Outcomes - study did not report ongoing assessment of student outcomes - standardized achievement testing each year of all grade levels - assessment of instructional practices of teachers (ISOI)</p>	<p>Findings: Student Engagement: significant increase in engaged time from 2nd to 3rd year of project and then a decrease in 4th year Achievement: - longitudinal study - the 102 children - reading scores increase in 1st and 2nd years, but decreased in 3rd year - experimental students' reading scores were lower in 3rd year than control students</p> <p>Conclusions: - in-service PD was not ongoing throughout the 4 years - teacher support groups were not reported - description of PD was not clearly outlined over the 4 years - program was not maintained, perhaps due to the lack of ongoing PD, lack of peer support, and prescriptive nature of program so teachers became bored - standardized assessment measures do not account for gains or changes, focuses on level of achievement</p>

6. The Reading Instruction Study

Richardson's (1994) in-service PD focused on teacher change through reflection and thoughtful practice. There was little emphasis on theory building or increasing the background knowledge of the teachers. The project researchers were facilitators, and they were not perceived as experts. This staff development program occurred in five schools, but the data were collected in two schools from 12 teachers (see Table 12). Practical Argument was the device implemented "to assist teachers in examining their beliefs and possibly reconstructing them" (Richardson, 1994, p. 117). Coaching, feedback, and teacher support groups were features of this professional development program but there was little emphasis on the students' learning outcomes. The qualitative findings indicated that there was a change in teaching attitudes and beliefs in the group of teachers who participated in this PD. However, the quantitative measures of student change did not show any differences between the reading comprehension scores of students in the participating schools and control schools on the Iowa Test of Basic Skills. A second assessment measure, the Illinois Goal Assessment Program (IGAP), also did not demonstrate sustained changes in reading comprehension scores into the second year.

The Reading Instruction Study (RIS) contained the features of specific, realistic goals, coaching and feedback, teacher support groups, and assessment of teacher change and assessment of the learning outcomes of students. However, in this PD there was no formal instruction to improve the teachers' background knowledge. The researchers' role was that of a facilitator, not of an expert imparting knowledge of research theories and practices. The lack of improvement in the children's reading comprehension skills in the Reading Instructional Study demonstrated the need for presenting content-based knowledge as well as promoting reflective inquiry in teaching.

Table 12 Description of the Reading Instruction Study

Brief Description of Study	Key Features of In-service PD	Findings and Conclusions
<p>Teachers: $N = 12$ Students: $N = 276$ in first year of study $N = 159$ in second year of study</p> <p>Grades: 4, 5, 6</p> <p>PD of Teachers: - monthly meetings - video taping of teachers' classroom instruction - individual meetings with teachers to discuss tapes</p> <p>Dependent Measures: 1. - qualitative measures of teacher change in thinking and practice through observation, belief interviews, and taping of individual sessions and teaching sessions showed findings of change in teaching practice and beliefs 2. - quantitative measures of student change - Iowa Test of Basic Skills (ITBS) - Illinois Goal Assessment Program</p>	<p>1. Goals - focused on teacher change through thoughtful reflection and integration of research practice on reading comprehension</p> <p>2. Coaching and Feedback - individual sessions with staff developers examining videos of teaching sessions where teacher was to describe rationale for teaching strategy</p> <p>3. Teacher Support Groups - group meetings of participating grades 4, 5, and 6 teachers to talk about instructional practice - staff developers served as models for reflection and providers of information</p> <p>4. Coaching - assessed teacher change through qualitative measures</p> <p>5. Feedback on Student Outcomes - assessed student change through quantitative measures at end of each year</p>	<p>Findings: - change in the participating teachers' attitudes towards teaching reading - ITBS showed no significant differences between the control school and participating schools in the first year or second year of the study - the IGAP showed significant differences, with improved scores of students in the classrooms of participating teachers in the first year but no differences in the subsequent year</p> <p>Conclusions: - included many features of effective in-service - no direct instruction of background knowledge and classroom practices - focused exclusively on reflective inquiry - study examined only one aspect of literacy, reading comprehension</p>

Conclusions Concerning Professional Development to Promote Better Literacy Education

There are few in-service PD programs designed to improve the literacy skills for students in the general classroom. These six studies are representative of professional development approaches in the literature and they demonstrate the strengths and shortcomings of the research on PD.

Among the six studies there is considerable variation. There is variation in types of professional development provided, in the methodologies employed, in the measures used to assess the effectiveness of the PD, and in the focus of change in the research (whether it is the students, the teachers, or both). In examining the PD described in this literature, it was apparent that some PD is very prescriptive, with standard instructional practices. In other PD the teachers individualized the instructional strategies for their classroom, using the components and skills taught in the PD. In this research the role of the “expert” in professional development varied. In some studies the “expert” presented information, led discussions, modeled the instruction strategies, provided feedback, and assessed the outcomes. In another study, the “experts” were facilitators who did not lead the discussions, but provided assistance when necessary.

The implementation of the features of effective PD discussed earlier, varied among the studies. Some of the PD presented little theory or background, whereas other PD presented theory, as well as strategies, with demonstration, modeling, practice and coaching. Collegial support of peers was described in some studies and not in others. The type of feedback to teachers on their implementation of the skills from the PD into their classroom practice varied from little feedback, peer feedback, feedback from researchers using video tape or actual classroom observations. Some studies have little description of the professional development. The skills taught to the students and the management of time for teaching these new strategies within the

classroom were described, but the features for implementing the transfer of the knowledge from the PD into classroom practice were not described.

In methodology, these studies varied in their use of control groups or baseline groups. Some studies had no control groups and used test norms as a basis of assessing program effectiveness, whereas others had control groups, or matching cohort groups. In reporting the effectiveness of PD, some studies measured only student outcomes. Student change in achievement was the focus of these studies. Scores from standardized literacy measures from students who had been taught by participating and nonparticipating teachers were compared. However, there was variation in the measures used in assessing student outcomes. Some studies reported multiple aspects of literacy, such as word recognition, reading comprehension, oral reading, and spelling, whereas other studies reported the results from only one aspect of literacy such as reading comprehension. Some research had ongoing assessment of student outcomes during the professional development, whereas other research assessed student outcomes at the end of the PD. Furthermore, some results were reported after a short implementation of the PD, such as a few weeks, or after a year-long PD, or over several years. For some researchers measuring the maintenance of a PD was an important aspect of its effectiveness.

Teacher change in instructional practice after participating in a PD was measured in a few studies. In some studies the teachers were asked to evaluate the PD program, not changes in their classroom practice. In other studies, teacher change was assessed through checklists examining their knowledge and attitudes before and after their participation in the PD.

Observations by the researchers of the implementation of the skills presented in the PD in the classrooms of participating and nonparticipating teachers occurred in some studies. Therefore, there are variations in the type of results reported, student change in the acquisition of literacy

skills, teacher change in instructional skills, or both, and the types of measures used for assessing these changes as well as the time over which these changes were assessed.

There is a need for much more research on the effectiveness of professional development in literacy education for teachers in the general classroom. With the large variation among the few existing studies of PD, more research is required to provide converging evidence for the effectiveness of PD in improving the classroom practice of teachers with the associated improvement in the literacy achievement levels of students.

In the literature on effective professional development, there is more research on teaching small groups of low achieving students, or one-to-one tutoring of individual students such as described in Reading Recovery (Clay, 1986a, 1986b). This research focuses more on improving the literacy skills of “at risk” children, and not on improving the literacy achievement for all students in the general classroom.

This current research examines the improvement of literacy skills of primary students in the general classroom so the more expensive special education programs are avoided. Improving classroom practice of teachers may prevent children from failing to acquire competent literacy skills. This study examines the effect of PD for both teacher change and student change as reflected in their achievement levels in literacy in the short-term, immediately after the PD, and over the longer-term, two and three years after the PD. Furthermore, the study examines the effect of PD on the instruction of the classroom teacher for improving the reading and writing success of children considered to be “at risk” for literacy failure.

The next chapter describes the context, rationale, content and strategies of the current professional development in this study, and the objectives of this research.

CHAPTER 3

*1. The Context, Rationale, Content and Strategies for the Current PD**

The Context for the Professional Development Initiative

The initiative to implement school-wide PD occurred when testing of the literacy skills of the grade 3 students in the target school showed approximately 40% of these students to be performing below the grade/age expected level in reading and spelling (Willows & Burgess, 1996).

Although there was a temptation to explain these results away by pointing to a variety of potential “risk factors” – English as a second language, single parents, poverty, few books in the home – Willows and Burgess (1996) raised the possibility that many of the children, who were not reaching grade-appropriate literacy levels by the end of primary education, might be “curriculum disabled.” Rather than having intrinsic learning problems, it was speculated that the students’ poor literacy skills might be a product of a weak literacy program in their school combined with inadequate classroom instruction. In light of the low literacy achievement and the possibility that inadequate teacher knowledge might be a contributing factor, it was decided to attempt to raise the literacy level of the primary grade children through the use of the professional development model described in the next sections.

Rationale for Use of *The Literacy Diet* Metaphor

The Balanced and Flexible Literacy Diet: Putting Theory into Practice is a professional development system designed to put theory into practice in primary and junior classrooms (Willows, 1994, 2000). The content of the system is based on current research-based wisdom about how literacy skills are acquired and about the most effective instructional practices for

* Note: With the permission of the author, this section was drawn directly from an unpublished manuscript.

facilitating literacy acquisition. The key components of the system were drawn from a review of the research literature, summarized earlier in Table 3. The professional development strategies followed in *The Literacy Diet* system were based on what theory and research suggest are key features contributing to effective professional development and to the implementation of change in schools. These key components for effective PD were summarized earlier in Table 5.

In the school participating in the current project (i.e., the target school), as in nearly every other school in the English-speaking world, there has long been an ongoing dispute among educators about how children should be taught to read in the beginning stages. Some have advocated starting with a phonics emphasis while others have argued for a whole language approach. Disagreement has centered on whether teaching should begin with explicit instruction in phonemic awareness and symbol-sound correspondences, or whether initial instruction should be meaning-centered, with letter-sound correspondences taught incidentally, in context, as needed. Unfortunately, this phonics-*versus*-whole-language debate has long been characterized by intolerance and polarization, with extremists advocating either a rigid teacher-centered "back-to-basics" approach or an uncompromising child-centered "developmental" approach. Such dogmatism and inflexibility is often a consequence of limited knowledge of a subject. Teachers who are either relatively new and inexperienced or who have been in the profession for many years and have not recently updated their knowledge are particularly vulnerable to embracing educational fads and ideologies uncritically.

Pre-service and in-service education have a key role in providing teachers with a level of theoretical and practical knowledge that should prepare them to approach early literacy acquisition as a complex challenge to which they must address all of their critical and creative

faculties, formulating their own working models based on the assimilation of new information and accommodation of their own schemas. Unfortunately, as discussed earlier, pre-service teacher education is all too brief and rarely provides teachers with the foundation of understanding they need in order to be critical consumers of educational theories and new methodologies. Moreover, in-service teacher education – which often involves single inspirational sessions – may briefly energize teachers but rarely results in significant positive changes in classrooms. Thus, as discussed in the literature review, most pre-service and in-service teacher education has failed to provide the necessary theoretical understanding of the processes involved in literacy acquisition (American Federation of Teachers in *Teaching Reading IS Rocket Science*, 1999; Richardson, 1998; Snow et al., 1998).

To be effective, teachers' personal working models must be open to change as new and conflicting data about reading and writing become available. Those who lack sufficient breadth and depth of knowledge can "get stuck" in narrow and simplistic interpretations, resorting to rhetoric and slogans to support their beliefs. One of the greatest barriers to progress in the field of literacy education has been the development of a vocabulary of emotion-laden language. Words and phrases such as: *literacy, mechanistic, authentic, drill, whole language, phonics, learning by osmosis, child-centered, direct instruction, skills, natural, standardized testing, holistic, word calling, developmental*, etc. have come to carry much more meaning than the lay observer might imagine. What was once called the "great debate" in the 1960s and '70s (Chall, 1967, 1983, 1996) has become the "reading wars" in the 1980s and '90s (Stanovich, 1993/94). Colleague has been pitted against colleague and children have been the victims of the conflict.

The Balanced and Flexible Literacy Diet framework was designed by Willows (1994) to allow educators to approach the topic of literacy acquisition without the baggage of old wars. In *The Literacy Diet*, the overused and emotion-laden vocabulary of the "great debate" and the "reading

wars" has been replaced with an intuitively-appealing diet metaphor that educators easily understand, allowing them to move away from slogans and rhetoric to logic and common sense. Within this framework research findings are presented in an organized and comprehensive system and classroom practices are structured in a meaningful way.

Content of the Professional Development

The simple notion underlying *The Balanced and Flexible Literacy Diet* is that in order to promote growth in literacy we must provide the right amount and type of "food for literacy" and we must insure that every student consumes enough of the right literacy foods on a daily basis.

Balance. Literacy diet "components" represent the equivalent of the food groups (e.g., grains, fruit and vegetables, meat and alternatives, dairy products and alternatives), and obviously no diet is balanced if it includes only one or two food groups. The key "food groups" of *The Literacy Diet*, which are based on both research and practice, are listed in Table 13. These components are required in appropriate proportions, complementing each other in fulfilling all nutritional requirements for literacy growth. Classroom teachers must understand why, when and how these components should be offered to insure the literacy success of their students.

Flexibility. As in any other diet, not everyone enjoys all foods for literacy. In *The Literacy Diet* framework, it is "OK" if you don't like broccoli! – for both teachers and children. There are many different "nutritious" activities to provide each of the literacy diet components. Teachers need not throw out all previous classroom practices to create a more effective literacy program; they simply need to do a "nutritional analysis" and choose/create a literacy diet for their students that is balanced and appealing.

Table 13 Key Components of A Balanced Literacy Diet

COMPONENT	BRIEF DESCRIPTION
Motivation for literacy	promoting enjoyment of books and appreciation of the usefulness of reading and writing
Concepts of print	demonstrating for children the purpose of letters and printed words and how they are used in books and other texts
Word/world knowledge	extending children's experience, understanding and use of a large repertoire of words and concepts
Language development	modeling and eliciting from children language structures that are more elaborated and more varied than the ones they use outside of school
Listening/thinking skills	promoting active listening, reasoning, and relating of new information to what children already know
Sight words	systematically teaching and promoting practice in recognizing the commonest irregularly spelled words, beginning with the most useful ones for reading and gradually, as these are learned, moving on to the less common ones
Phonemic awareness	teaching children to focus on and manipulate the sounds in spoken words, to blend them together into words (for reading) and to break words down into their constituent sounds (for spelling)
Letter-sound associations	explicit, systematic teaching of the letter-sound connections (40+) and practice in using them for decoding (reading) and encoding (spelling/writing)
Letter formation	providing explicit teaching and practice of the correct formation of letters, and setting reasonable quality expectations for printing and writing
Spelling	systematic teaching and practice of common spelling patterns and word families, providing mnemonics for the spelling of irregular/"tricky" words; promoting correct spellings of high-frequency grade-appropriate words and inventive spellings of less-common ones
Schema development	teaching, modeling and eliciting the use of frameworks to assist children in developing and organizing what they already know and to facilitate their understanding, memory and use of new information in their comprehension and production of oral and written language
Real reading	engaging children in reading real things for authentic purposes
Fluency	promoting speed, accuracy, and proper expression in reading and writing through extensive appropriate practice with guidance and feedback
Text types	providing opportunities, explanations, structures and purposes to promote children's reading of a wide range of stories and informational texts
Comprehension strategies	teaching, modeling and promoting children's use of a variety of effective strategies to improve their comprehension of what they read
Real writing	engaging children in writing real things for authentic purposes
Writing conventions	teaching, modeling, and promoting the proper use of capitalization, punctuation, paragraphing, etc.
Composition strategies	teaching, modeling and promoting children's use of a variety of effective strategies to increase the length, accuracy, coherency, and complexity of their written compositions
Written language structures	providing opportunities, explanations, structures and purposes to promote children's writing of a wide range of narrative and expository texts

Developmental Stages. Another useful lesson based on “food rules” is that human dietary requirements change at different stages of maturation. For example, when children’s bones are growing they require more foods from the dairy group because these foods contain calcium. Similarly there are changing literacy diet requirements as children grow in literacy. Although learning to read and write do not reflect “natural” developmental processes comparable to human maturation (Lyon, 1999), it is helpful to consider literacy acquisition within a stage framework (Chall, 1983, 1996; Frith, 1986; Juel, 1987; Marsh et al., 1981). Students at different stages of literacy development have different “literacy nutritional needs.” As students progress through the stages, the components and activities in their literacy diet must change in order to promote growth. To be effective, teachers must understand the nature of the stages of literacy development and provide their students with stage-appropriate foods for literacy. Table 1, presented earlier, provides a rough guide to features of the stages of literacy development through which teachers need to lead their students in the elementary school grades.

Special Needs. Some students come to the classroom with special “literacy nutritional requirements” because of their linguistic, cognitive, and experiential backgrounds. For these, *The Literacy Diet* approach is designed to assist the teacher in assessing needs and in designing appropriate “special diets” to ensure literacy growth. Teachers in both regular and special education need to share an understanding of *The Literacy Diet* framework to facilitate program planning for all of their students. Classroom teachers must work closely with the specialists to insure that the literacy diet they provide in the classroom is appropriate and consistent with that being provided by the special educators.

Effectiveness and Engagement. Even if a classroom literacy program is well designed to satisfy the basic “dietary requirements” of all children in the class, not everyone will grow in their

literacy unless they consume and enjoy what is being offered to them. The appeal of the diet – to both “the cook” and “the consumer” – are central to its success. All the right “literacy food groups” can be put together into a literacy meal by a “hospital dietitian” and nobody will ask for more (and the cook probably won’t get much enjoyment out of preparing the meal) or, essentially the same ingredients can be put together by a “gourmet chef” – with a little spice – and everyone will ask for more (and the cook will feel pride and personal satisfaction). In *The Literacy Diet* approach the inspired gourmet chef has a much better chance of motivating the children to consume the essential foods for growth in literacy.

Nutrition and Growth. Concepts of planning, classroom organization and time management are central to *The Literacy Diet* system. In order to assist teachers in organizing and managing classroom time and activities, *The Literacy Diet* metaphor suggests that they reflect on how much and what type of literacy nutrition each of the children in their classrooms is receiving. At any given point in time, the teacher learns to ask: What, if any, literacy nutrition is provided by this particular classroom practice? Who is getting the nutrition? How much are they getting?, etc. From such analyses, teachers come to distinguish between “literacy junk food” and highly nutritious literacy activities. They also recognize that activities which focus attention on one or two children while the others watch may, in many cases, be providing nutrition only to the children who are directly involved in the activities while other children are “starving!” Teachers develop detailed day plans based on their recognition that if every child is to receive as much as possible of the right foods for literacy every day, then the teacher needs to plan their day in the same way as a dietitian would plan a menu to insure that all children are receiving enough appropriate nutrition each day. Teachers’ day plans also must accommodate children’s special needs and preferences to ensure that all children will, in fact, consume the literacy foods that they require for growth.

Implementation of the Professional Development

Whereas *The Literacy Diet* metaphor provided the framework for the content for the professional development, the research literature on the change process in schools provided the implementation strategies for the professional development. The strategies used were designed to promote transfer of new skills to the classrooms by using aspects of the seven key features outlined in Table 5, earlier. That is, the approach involved theory-demonstration-practice, feedback linking change of instructional practice to successful student outcomes, a teacher support group, promotion of teacher self-efficacy, setting realistic goals, strong school leadership, and a supportive school environment.

As Guskey (1986) recommended, the implementation involved a systematic attempt to bring about change in the beliefs, attitudes and classroom practices of teachers, and a change in the learning outcomes of students through ongoing professional development over an entire school year. Twice-monthly project meetings were held during which the participants met with the university "expert" to develop and implement balanced and flexible programs to meet the needs of the children in their classrooms. All of the primary teachers and support staff in the target school were given the opportunity to opt into the project. All but one teacher did. In order to have her classroom included in the project (all participants were female), each primary teacher had to be willing to make a commitment to participate in professional development throughout the year. This entailed reading books and articles, attending bi-weekly after-school meetings, attending in-school literacy meetings on the alternate weeks, preparing and sharing program materials, implementing program components, individualizing programs for special needs students, documenting classroom time-allocation, and participating in assessment to evaluate program effectiveness.

The bi-weekly, after-school meetings involved presentations by the university “expert/facilitator,” group discussion of professional development materials from books and videotapes, demonstration of research-based teaching strategies, and sharing by participating teachers of approaches and materials that they had tried and found effective. Copies of the four books listed below were provided from school budget for each teacher, as required reading. A variety of supplementary theoretical and practical materials were also provided as optional reading.

Phonics They Use: Words for Reading and Writing (Second Edition) (Cunningham, 1995)
Classrooms That Work: They Can All Read & Write (Cunningham & Allington, 1994,1999)
Early Literacy in the Classroom (Depree & Iverson, 1994)
Learning to Read and Write: Adapting Programs to Children's Needs (Willows, 1993)

Planning. The beginning step in program planning was to examine the outcomes desired by the end of each of the literacy stages. Based on these, *The Literacy Diet* framework was used to generate the types of activities needed to achieve the desired outcomes. An integrated program approach was developed collaboratively with the participants. This program included all of the components (listed in Table 13) deemed essential to achieve the desired outcomes, and made use of a range of teaching materials and approaches. The program incorporated the best aspects of phonics approaches, sight approaches, and the whole language philosophy (based on *Learning to Read and Write*, Willows 1993, and *A Normal Variation View of Written Language Difficulties and Disabilities: Implications for Whole Language Programs*, Willows 1991). Balance and flexibility built into the program allowed individual teachers to select from a range of classroom activities, as long as essential program components were included and sufficient time was allotted to activities to insure satisfactory program outcomes. Two simple principles

guided all teaching practices: (1) all program activities *had to make sense* in terms of the desired outcomes of the program and (2) all program activities had to be *motivating to children*.

Time management. A key problem in many classrooms is that the actual amount of time children spend reading and writing is uncontrolled. Although there are certain times of day that are considered "literacy time," for any given child there is often little monitoring of "on task" time. Whereas some children may spend most of the language arts periods actively engaged in reading and writing activities, others (usually those most in need of assistance) spend much of their time engaged in talking to other children, waiting in line for help, doing artwork or other non-literacy activities. Many of the children who are "off task" (because reading and writing are hard for them) are those most in need of in-school reading and writing practice because they may be involved in very little reading and writing outside of school. In order to ensure success for all children, classroom activities were developed that promoted a high degree of "individual engagement time"; each child was encouraged to do their "personal best" and not depend too much on others to do their thinking. This did not preclude working in groups or pairs, but required careful planning on the teacher's part to ensure that all children were engaged in activities appropriate to their stage of literacy development every day.

Program implementation. Under the guidance of the university "expert/facilitator", the school vice-principal, the support staff (special educators and ESL teachers), and classroom teachers at each of the primary grades (junior kindergarten through grade 3) worked together to develop program plans based on the model and materials discussed in the professional development sessions. Within *The Literacy Diet* framework, teachers (guided by professional development readings and discussion) determined the essential components required to achieve the desired

* Note: School budget was also allocated to provide delicious and nutritious food for these after-school meetings!

outcomes. A range of alternative classroom activities were generated, consistent with each program component, and each teacher developed an individual program that included all of the essential program components. Teachers developed, regularly revised, and implemented 5-day cycle plans that included detailed child-by-child, activity-by-activity, day plans.

Monitoring and adjustment. Even the best planned programs do not always work out as expected. Sometimes the children may not be interested in a particular activity or may find it too easy or too hard. Ongoing monitoring of the effectiveness of the program and a willingness to modify aspects of a program that did not seem to be achieving the desired ends were both essential to success. Informal assessment and regular collection of writing samples provided valuable information to assist in monitoring the progress of the group as a whole, as well as of individual children. Bi-weekly site visits by the university "expert" to observe and conference with teachers provided an opportunity for feedback on individual classroom programs. Specific children and small groups of children who had special programming needs received customized programs, with the support of the special education and/or ESL teacher. Standardized evaluation of the effectiveness of the professional development on student outcomes was based on the assessment at the end of the school year, to be described in the Method section.

2. Current Research Objectives

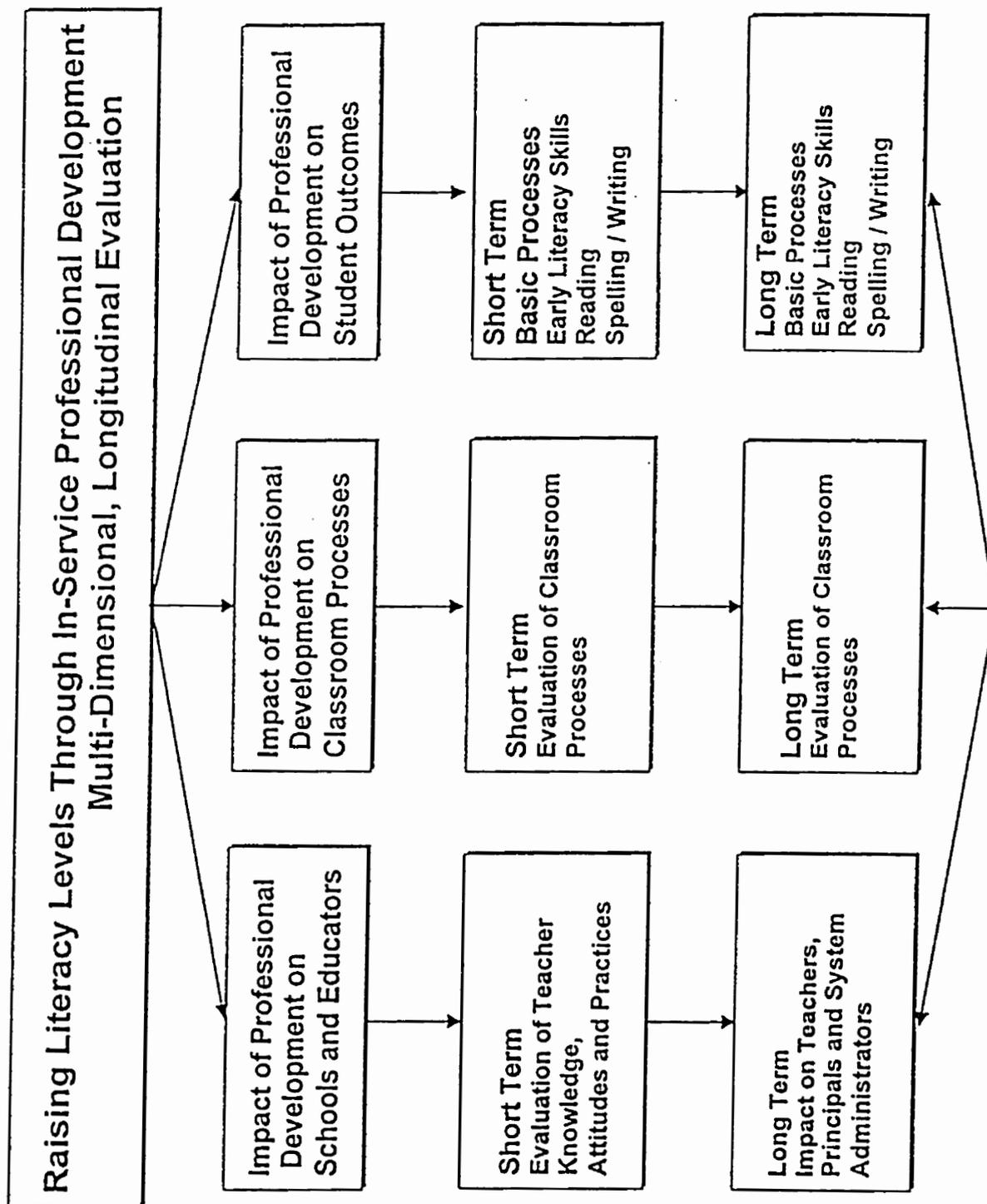
The present thesis was designed to provide empirical data about the effectiveness of this particular in-service PD system – the *Balanced and Flexible Literacy Diet: Putting Theory into Practice* (Willows, 1994) – on the acquisition of literacy skills by primary grade students.

The overall goal of the present research was to determine whether the participating teachers internalized and applied the knowledge acquired through the in-service PD in their classroom practice by examining whether the reading (word recognition and reading comprehension) and spelling of children in their classrooms reflected benefits of their teachers' raised knowledge levels.

Research Context

The research reported in this thesis is part of a much larger project in which the impact of *The Literacy Diet* professional development approach on teachers' and school administrators' knowledge, attitudes and practices, as well as students' literacy outcomes, have been assessed (Barnes-Haig & Willows, 1999; Haas & Willows, 1998; Jackett & Willows, 1998; Siegelman & Willows, 2000; Willows & Burgess, 1996). For example, one component of the larger study involved detailed time sampling to examine how classroom literacy time was spent compared to controlled classrooms, where the teachers had not participated in the in-service PD. The results clearly demonstrated positive changes in literacy programming after teachers' participation in the in-service PD (Sumbler & Willows, 1998). This larger study involved many more schools and teachers. Figure 1 shows the various areas studied within the larger project. The present research focuses on short-term and long-term learning outcomes of students in literacy.

Figure 1. Overall Framework for Evaluation of Professional Development



Research Questions

To assess the effects of the *Balanced and Flexible Literacy Diet* in-service PD on the reading and spelling performance of students in the primary grades the following questions were addressed:

1(a). Immediately following the year-long in-service PD (June 1995), did student outcomes in word recognition, spelling and reading comprehension in the classrooms of Grade 1 and Grade 2 teachers who participated in the professional development improve in relation to the pretreatment baseline obtained one year earlier (June 1994), and were there longer-term effects two years subsequent to the in-service (June 1996 and June 1997)?

1 (b). For the three participating teachers who remained in the same grade across all four years, did the student outcomes in word recognition, spelling and reading comprehension in the year of the in-service PD (June 1995) in these Grade 1 and Grade 2 classrooms improve in relation to the pretreatment baseline obtained one year earlier (June 1994), and were there longer-term effects in the two years subsequent to the in-service (June 1996 and June 1997)?

2. Was the performance of Grade 1 and Grade 2 children who had one, two, three or four years of ongoing instruction by teachers who had participated in the professional development significantly better on measures of word recognition, spelling and reading comprehension than that of children who had no exposure to in-serviced teachers ?

3. Did teachers' participation in professional development effectively reduce the degree of literacy failure among their students?

CHAPTER 4

Overview of Research Methodology

This study was designed to evaluate the effectiveness of in-service PD by examining the literacy outcomes of the children. Before the initiation of in-service program, baseline testing of children in Grades 1 and 2 was undertaken in June 1994.* Between September 1994 and June 1995 the treatment period occurred with the in-service PD of teachers. The *Balanced and Flexible Literacy Diet* system was introduced to Junior and Senior Kindergarten teachers, Grades 1, 2 and 3 teachers, Special Education and ESL teachers. The immediate impact of the professional development was assessed at the end of the in-service year, in June 1995. The longer-term impact of the effectiveness of the in-service on the participating teachers was assessed with the same literacy measures in June of subsequent years. Each year the content and timing of the main assessment battery was the same, involving measures of word recognition, spelling and reading comprehension.** Table 14 presents a brief overview of the study in the form of a time line.

Table 14 Overview of Data Collection on a Time Line

June 1994	Sept.1994-May 1995 Implementation Year	June 1995	June 1996	June 1997	June 1998
Baseline data collection	In-service PD of JK, SK, Grade 1, 2 & 3 teachers	Immediate follow-up evaluation	One-year follow-up evaluation	Two-year follow-up evaluation	Three-year follow-up evaluation

*Note: Although some testing was undertaken in Grade 3 classrooms, due to the stress on the teachers created by the newly implemented Province-wide Grade 3 testing (EQA0), it was not possible to undertake complete longitudinal testing in the Grade 3 classrooms.

**Note that although a written composition measure was also administered, it was not included in the present research, but has been examined elsewhere (Jackett & Willows, 1998).

Test Instruments

Table 15 presents the study's overall design, along with the measures administered in each testing period.

Table 15 Literacy Measures Administered During Each of the Five Testing Periods

June 1994 Baseline Year	June 1995 Intervention Year	June 1996 Follow-up Year 1	June 1997 Follow-up Year 2	June 1998 Follow-up Year 3
Burns & Roe	Burns & Roe	Burns & Roe	Burns & Roe	Burns & Roe
WRAT- R: Spell	WRAT-R: Spell	WRAT-R: Spell	WRAT-R: Spell	WRAT-R: Spell
Gates-M.	Gates-M.	Gates-M.	Gates-M.	Gates-M.

Key to Abbreviations

Burns & Roe	Word Recognition	Burns and Roe Informal Reading Inventory
WRAT-R: Spell	Spelling	WRAT-R Spelling Subtest
Gates-M.	Reading Comprehension	Gates MacGinitie Reading Tests - Second Edition

In each of the five testing periods in June the students in Grades 1 and 2 were administered the main assessment battery. For purposes of analyses, raw scores were used for all the outcome measures.

Word Recognition

Burns & Roe Word Recognition List (B&R; Burns & Roe, 1985)

The B&R word lists (from the Burns & Roe Informal Reading Inventory, 1985) are graded from a preprimer to 12th grade level. Each level consists of 20 words. Two alternative sets of lists (List 1 and List 2) are available. For this study, the List 1 was used. Words were presented on 4.25" x 12" cards (one card per level) listed in a column in the same font as in the inventory (Times, 18 point). Each child read as many lists as required until the discontinuation criterion of 7 errors out of 20 words at a particular level was met.

Spelling

WRAT-R Spelling subtests (WRAT Sp.; Jastak & Wilkinson, 1984)

The spelling subtest of the WRAT - R (Wide Range Achievement Test - Revision) is a standardized measure assessing spelling. The WRAT-R provides absolute scores, standard scores and grade scores. It was normed on a stratified sample of individuals from aged 5 to 75. The test-retest reliability for the spelling subtest is .97. It has two levels. The first level for ages 5 years to 11 years 11 months was used for this study. The spelling test requires the child to write words pronounced orally by the tester. The maximum number of words to be spelled, depending on performance, is 45. After 10 consecutive spelling errors the testing ended. The total number of correctly spelled words was scored. The copying of the symbols and correct spelling of the child's name were not included in the spelling raw score, as the focus of the study was the spelling of words. However, in the calculations of percentiles, the total score, including the children's names and symbols, was used.

Reading Comprehension

Gates-MacGinitie Reading Tests-Second Can. Ed. (Gates-M.; MacGinitie & MacGinitie, 1992):

The reading comprehension subtest of the Gates-MacGinitie Reading Tests was administered to assess reading comprehension of story passages. The first passages are simple sentences, and the later passages involve longer sentences and more complex verbal relationships. Four pictures accompany each passage. The child's task is to choose the picture that illustrates the passage or that answers a question about the passage. Form 3 was administered to all the students, but Level A was given to the Grade 1 children and Level B was given to the Grade 2 children. Raw and standard scores were obtained for each child. There was a time limit of 35

minutes for the completion of the test questions. The test-retest reliability is .93 for Level A (Grade 1) and .94 for Level B (Grade 2). This test was normed on a stratified sample of students from kindergarten to Grade 12.

Procedures: Testing and Scoring

Throughout the years of the study the children were assessed in the school by Special Education teachers and Graduate students. The assessment of word recognition skills occurred in one-to-one testing between the examiners and each child. Spelling and reading comprehension assessments occurred in group testing. All assessments were conducted with standard procedures under controlled conditions.

Word Recognition

The Burns & Roe Word Recognition was administered individually to each child. Each child began with a card showing the 20 word pre-primer list and continued reading into the higher levels until seven words were read incorrectly within a level (not necessarily consecutive). In keeping with the Burns and Roe format, the words were listed vertically and were shown one-at-a-time by moving a cover-card down the list. Five seconds were allowed for a response before moving on to the next word. The number of correctly read words was recorded for each child.

Spelling

After all the classes had been administered the reading comprehension task, each Grade 1 and Grade 2 class were administered the spelling subtest of the WRAT-R in the standardized manner, in a room without letters, words, or writing samples on display. The children first copied the symbols at the top of the page, and then printed their name. Next, the examiner

orally pronounced a word, repeated a sentence with the word in it, and pronounced the word again. The classroom teacher circulated around the room, and the testing was stopped when all the children had made 10 consecutive errors. Therefore, for many children, there were more than 10 spelling errors as the testing ceased when the child who was the best speller had 10 spelling errors. The test was scored in the standardized manner.

Reading Comprehension

The reading comprehension measure was administered to each class of Grade 1 and Grade 2 students in a classroom without letters and writing samples. The examiner administered the test in the standardized manner. The classroom teacher assisted with any clarification of the instructions and the monitoring of the behaviour of the children during the testing session. At the end of 35 minutes, the children were asked to close their booklets. The number of correct answers was recorded.

Participant Characteristics

The children attended one large, neighbourhood elementary school. The school was situated in a suburban area of small houses, duplexes, and apartments in Metropolitan Toronto. Parental permission was obtained. (Sample of parental consent for children's participation in the project is in Appendix A.) Teachers volunteered for the in-service professional development. In view of the complexity of the different samples associated with each research question, the participant characteristics are described in the design, results and discussion section of the chapter addressing each question.

CHAPTER 5

Tracking the Teachers: Study of Teacher Effectiveness

In this chapter the impact of a particular approach to in-service PD, *The Balanced and Flexible Literacy Diet: Putting Theory into Practice*, is assessed with respect to growth in teacher effectiveness. Within this context one main experimental question and one subsidiary question are addressed. The overall goal of the first research question was to draw inferences about teacher effectiveness on the basis of the literacy outcomes of the students in their classrooms. It was assumed that if teachers gained new knowledge and skills through their participation in professional development sessions then they would become better literacy teachers, and their students would achieve greater literacy success. Ideally, to study teacher effectiveness the mean learning outcomes from the classrooms of a large number of teachers would be the best approach to this research question. However, the number of teachers in this study was particularly small so the examination of teachers as a variable was not possible. Therefore, student outcomes were examined as a reflection of teacher effectiveness.

The focus was on the outcomes of children in the classrooms of Grade 1 and Grade 2 teachers who participated in the in-service. To determine whether teachers' acquisition of new knowledge in the instruction of literacy was associated with corresponding student improvements, word recognition, spelling and reading comprehension abilities of the students in their classrooms were assessed both immediately after the in-service (June 1995) and later, one and two years after the in-service (June 1996 and 1997). The main focus of the first research question was on the effectiveness of the instruction of the teachers who participated in the in-service PD. A subsidiary research question examined the effectiveness of the instruction of three particular teachers who continued over the years of the study to teach at the same grade level.

The chapter will begin with a description of the sample of teachers who were involved in this study and the characteristics of the children in their classrooms. The design, results and discussion of the first part of Research Question 1 will follow. Next, the characteristics of the three continuing teachers and the children in their classrooms will be described for the subsidiary question, followed by the design, results and discussion.

Tracking the Teachers: Part I

The main focus of this chapter is on Research Question 1(a), which can be stated as follows:

1(a). Immediately following the year-long in-service PD (June 1995), did student outcomes in word recognition, spelling, and reading comprehension in the classrooms of Grade 1 and Grade 2 teachers who participated in the professional development improve in relation to the pretreatment baseline obtained one year earlier (June 1994), and were there longer-term effects in the two years subsequent to the in-service year (June 1996 and June 1997)?

Participant Characteristics

Teachers

Experimental teachers were those teachers who had participated in the in-service PD during the school year from September 1994 to June 1995. Control year teachers had not participated in the in-service since they had taught in the school year before the in-service PD from September 1993 to June 1994. Those control teachers who remained in the school during 1994/95 became part of the experimental group of teachers as they participated in the in-service PD during the Intervention year of September 1994 to June 1995.

Each teacher who participated in the PD was coded using an alphabet letter and the teachers retained their letter throughout the study for identification purposes. Table 16 provides a listing

of the teachers participating in the study and shows the number of teachers who remained over the four years in the study. Table 17 presents the characteristics of the participating teachers, including: the number of years of previous teaching experience and the grades the teachers had taught prior to participating in the in-service teacher education.

Table 16 Overview of Experimental Teachers

	Year of Study			
	1993/1994 Baseline Year	1994/1995 Intervention Year	1995/1996 Follow-up 1 Year	1996/1997 Follow-up 2 Year
Grade 1	TA	TA	TA	TA
	TB	TB		TD
	TC	TH		
	TD	TI		
Grade 2	TE	TE	TE	TE
	TF	TF	TF	TF
	TG	TG	TH	TH
	TH	TD		

- Notes:
- T = Teacher and A...H = Teacher identification number
 - There were similar number of classrooms in each grade for each year, but only those teachers who participated in the in-service are included in the Table.
 - In all grades there were multiple classes and assignment of children to classrooms was not based on children's performance. Children did not remain with the same teacher in the following year if a teacher changed to a different grade level.
 - Data from the classes of new teachers coming into the school after the PD year (1994/95) were not included in any analyses.

Table 17 Characteristics of the Teachers

Teachers	Characteristics	
	Years of Teaching Experience	Grades Taught
TA	9	1, 2
TB	2	1, 2
TC	1	Kindergarten
TD	6	1, 3, Jr. Remedial
TE	18	1, 2, 3
TF	28	Kind., 1, 2, 3, 4, 5
TG	1	2
TH	11	1, 2, 4
TI	18	1, 2, 3, 4, 5

Note: These characteristics describe the teachers as they entered the PD on September 1994.

Children

The Experimental Groups were the Grade 1 and Grade 2 children who were in the classes of teachers who had participated in the professional development. The Baseline Group were Grade 1 and Grade 2 children who were in the school during the baseline year, before the initiation of the in-service PD (tested in June 1994). The effect of the in-service PD on the instruction of Grades 1 and 2 teachers who participated in the in-service was examined by analyzing the impact on the outcome literacy measures (word recognition, spelling and reading comprehension) of children in their classrooms immediately following the PD year (June 1995) and then examining the longer-term effects of the in-service by analyzing the same measures of the children in their classrooms one and two years after the in-service (June 1996 and 1997). Analyses of the scores of these groups of children – across years within grade – provide an indication of the impact of being in the classroom of in-serviced teachers on the literacy outcomes of children.

Each of the years of participation in the study will be described as follows:

In the *Baseline Year* the children were in the classrooms of teachers before the teachers participated in the in-service professional development and the children were assessed on the three dependent measures in June 1994. In the *Intervention Year* the children were in the classrooms of teachers who had participated in the in-service PD from September 1994 to June 1995, and were subsequently assessed in June 1995. In *Follow-up Year 1* the children were in the classrooms of teachers who had participated in the in-service PD the previous year and these children were assessed on the same dependent measures in June 1996. In *Follow-up Year 2* the children were in classrooms of teachers who had participated in the in-service PD two years previously and these children were assessed on the same three dependent measures in June 1997. The Baseline Group were children tested at the end of the Baseline Year (June 1994). The Experimental Groups were children tested at the end of the Intervention Year (1995), Follow-up Year 1 (1996) and Follow-up Year 2 (1997).

The first Research Question was addressed through analyses comparing the scores of children in Grade 1 and Grade 2 in each of the four successive years, where the first year served as a pre-treatment baseline. These analyses were concerned with the effectiveness of the classroom teachers in promoting literacy skills in the children in their classrooms.

Table 18 presents the overall number of children assessed each year in the Grade 1 and Grade 2 classrooms of participating teachers and the characteristics of this population of children. There was no difference between the mean ages of Grade 1 children over the four years of the study, $F(3, 188) = .232, p < .874$. However, there was a significant difference between the mean ages of the Grade 2 children over the four years, $F(3, 254) = 4.52, p < .004$. As Table 18 shows, this

effect resulted from the fact that the Grade 2 children in 1995 and 1997 were slightly younger than those in 1994 and 1996. Therefore, the main analyses included ANCOVAs to determine if age was a significant factor affecting the results.

Analyses

In order to evaluate the immediate and longer-term effects of the in-service teacher education on student outcomes, MANOVAs, orthogonal comparisons and trend analyses were performed. The results of these analyses will be presented for Grades 1 and 2 children for each of the dependent measures: word recognition, spelling and reading comprehension.

Effect sizes were used to assess the effectiveness of the PD on the outcome measures by measuring how much the means of the Intervention Year, Follow-up Year 1 and Follow-up Year 2 exceeded the mean of the Baseline Year in standard deviation units. An effect size of 1.0 indicates that the means of the Experimental years were one standard deviation higher than the mean of the Baseline Year, revealing a strong effect of the PD. An effect size of 0 indicates that the means of the Baseline Year and Experimental years were identical and there was no effect of the PD. To judge the strength of an effect size, values as suggested by Cohen (1988) are commonly used. An effect size of 0.20 is considered small; a moderate effect size is 0.50; an effect size of 0.80 or above is large.

Table 18 Summary Table of Participant Characteristics

	<u>Baseline Year</u> June 1994	<u>Intervention Year</u> June 1995	<u>Grade 1</u>			
			<u>Follow-up Year 1</u> June 1996	<u>Follow-up Year 2</u> June 1997		
Sex	n = 83 47 boys 36 girls	n = 52 21 boys 31 girls	n = 25 15 boys 10 girls	n = 33 15 boys 18 girls		
Age (mon.)	<u>M</u> 83.1 (3.3)	<u>M</u> 83.3 (4.3)	<u>M</u> 82.6 (4.9)	<u>M</u> 83.1 (3.0)	<u>F</u> .232	<u>p <</u> .874

	<u>Baseline Year</u> June 1994	<u>Intervention Year</u> June 1995	<u>Grade 2</u>			
			<u>Follow-up Year 1</u> June 1996	<u>Follow-up Year 2</u> June 1997		
Sex	n = 68 42 boys 26 girls	n = 58 29 boys 29 girls	n = 79 34 boys 45 girls	n = 54 28 boys 26 girls		
Age (mon.)	<u>M</u> 95.2 (3.7)	<u>M</u> 94.5 (3.4)	<u>M</u> 96.7 (5.3)	<u>M</u> 94.2 (4.3)	<u>F</u> 4.52	<u>p <</u> .004*

Notes: - * $p < .01$.

- All scores are raw scores.

- Over the four years of study, some of the in-serviced teachers retired, moved to other schools, or into the junior grades, so the pool of children in the classrooms of in-serviced teachers has dropped yearly in the Grade 1 classrooms.

- Due to the relatively high mobility of children in the school these analyses include both children who entered or left the school over the four years of testing as well as those who remained in the school for the full period.

One of the assumptions of analyses of variance is homogeneity of variance. Analyses were run on both parametric and nonparametric statistics because the Levene Test of Equality of Error Variance was significant for the scores of the Grade 2 children on the word recognition and spelling measures, and thus there was a concern about the degree of variability among the scores on these measures. Therefore, nonparametric statistics were also run on those analyses for which Levene's Test was significant. The nonparametric Kruskal-Wallis Test using Chi Square analyses produced negligible differences from the ANOVAs. Both analyses showed the word recognition and spelling scores of the Grade 2 children as significant. The Chi Squares for word recognition and spelling were $\chi^2(3, N = 237) = 30.17, p < .001$, and $\chi^2(3, N = 240) = 22.74, p < .001$, respectively. For simplicity, the parametric statistics only will be discussed further.

The data were examined for outliers with the scores transformed into z scores. Any scores beyond 2.5 standard deviations from the mean were considered outliers. In Grade 1 there were 2 outliers in Follow-up Year 2 and 1 outlier in the Baseline Year. In Grade 2 there were 2 outliers in Follow-up Year 1 and 2 outliers in Follow-up Year 2. There was only a negligible difference with and without outliers in the multivariate level of significance. Therefore, the outliers were included in the analyses, since these scores represented students who performed very well or very poorly on the literacy measures, likely reflecting the realistic variation in children's abilities in Grades 1 and 2 classrooms. In all cases, the analyses reported in the text were performed on raw scores.

Results and Discussion

Table 19 presents the Means and Standard Deviations for the Grade 1 and Grade 2 children for the four test periods.

Table 19 Assessment of Teacher Effectiveness: Means and Standard Deviations of Scores for Grade One and Grade Two Children on Three Literacy Measures

GRADE 1

	<u>Baseline</u> <u>Year</u> June 1994	<u>Intervention</u> <u>Year</u> June 1995	<u>Follow-up</u> <u>Year 1</u> June 1996	<u>Follow-up</u> <u>Year 2</u> June 1997
Word Recognition	$\underline{n} = 81$ \underline{M} 42.96 (38.47)	$\underline{n} = 51$ \underline{M} 58.47 (32.58)	$\underline{n} = 22$ \underline{M} 74.5 (44.95)	$\underline{n} = 33$ \underline{M} 84.64 (37.45)
Spelling	$\underline{n} = 74$ \underline{M} 8.18 (5.30)	$\underline{n} = 51$ \underline{M} 11.06 (4.19)	$\underline{n} = 24$ \underline{M} 10.04 (5.34)	$\underline{n} = 33$ \underline{M} 13.10 (5.74)
Reading Comp.	$\underline{n} = 74$ \underline{M} 25.31 (12.70)	$\underline{n} = 49$ \underline{M} 28.08 (11.59)	$\underline{n} = 23$ \underline{M} 26.52 (11.76)	$\underline{n} = 33$ \underline{M} 30.36 (10.96)

GRADE 2

	<u>Baseline</u> <u>Year</u> June 1994	<u>Intervention</u> <u>Year</u> June 1995	<u>Follow-up</u> <u>Year 1</u> June 1996	<u>Follow-up</u> <u>Year 2</u> June 1997
Word Recognition	$\underline{n} = 63$ \underline{M} 79.35 (36.28)	$\underline{n} = 56$ \underline{M} 90.88 (36.1)	$\underline{n} = 64$ \underline{M} 112.03 (28.64)	$\underline{n} = 54$ \underline{M} 103.56 (30.73)
Spelling	$\underline{n} = 60$ \underline{M} 12.60 (5.94)	$\underline{n} = 54$ \underline{M} 15.15 (5.69)	$\underline{n} = 72$ \underline{M} 16.71 (5.29)	$\underline{n} = 54$ \underline{M} 17.22 (5.52)
Reading Comp.	$\underline{n} = 66$ \underline{M} 28.41 (10.27)	$\underline{n} = 64$ \underline{M} 30.96 (11.66)	$\underline{n} = 72$ \underline{M} 32.58 (9.01)	$\underline{n} = 54$ \underline{M} 32.76 (9.33)

Notes: - All scores are raw scores.

- Not all children were present for all testing days so there is some minor variation in the sample size for the analyses.

As shown in Table 19, pre-treatment baseline mean scores were significantly lower on all three measures of word recognition, spelling and reading comprehension for both Grades 1 and 2 compared to the years of treatment. In order to avoid an inflated Type I error rate, omnibus MANOVA analyses were first performed for each grade (Stevens, 1996). The MANOVAs were all significant, thus treatment – the in-service PD of the teachers – significantly affected the Grade 1 and Grade 2 children’s scores on the literacy measures, $F(9, 389) = 8.18, p < .001$, $F(9, 523) = 5.2, p < .001$, respectively.

The ANCOVAs were run to determine if age was a factor affecting the outcomes of the scores of the Grade 2 children. There were negligible differences between the results of the ANCOVAs and the original analyses for the scores on the literacy measures of the Grade 2 children. The overall ANCOVA was $F(9,520) = 5.10, p < .001$.

Based on the significant MANOVAs, further analyses of the dependent variables were performed to determine if there was a treatment effect (1994 vs. 1995+1996+1997). The Baseline scores (June 1994) of each of the dependent variables from students in Grades 1 and 2 were compared to the scores of Grade 1 and Grade 2 children from the three experimental years (June 1995, 1996 and 1997).

Subsequently, more detailed analyses with planned comparisons of the dependent variables were performed to determine whether there was an immediate effect of the in-service education by comparing the scores on literacy measures of Grade 1 and Grade 2 children in classes of Baseline Year teachers who had no in-service PD (June 1994) with the scores on the literacy measures of children who were in the classrooms of Grade 1 and Grade 2 teachers who had participated in the one year long in-service PD (June 1995).

Table 20 Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness:

Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the Teachers' Professional Development and in Follow-up Years

Measure	t	Significance
Immediate Effects		
Grade 1		
<u>Baseline vs. Intervention Year: June 1994 vs. June 1995</u>		
Word Recognition	2.31	.02†
Spelling	3.10	.002 *
Reading Comprehension	1.26	.21
Longer-Term Effects		
Grade 1		
<u>Baseline vs. Follow-up Year 1: June 1994 vs. June 1996</u>		
Word Recognition	3.49	.001**
Spelling	1.56	.12
Reading Comprehension	0.42	.67
<u>Baseline vs. Follow-up Year 2: June 1994 vs. June 1997</u>		
Word Recognition	5.37	.001**
Spelling	4.80	.001**
Reading Comprehension	2.06	.05†
Immediate Effects		
Grade 2		
<u>Baseline vs. Intervention Year: June 1994 vs. June 1995</u>		
Word Recognition	1.90	.06
Spelling	2.43	.02†
Reading Comprehension	1.40	.17
Longer-Term Effects		
Grade 2		
<u>Baseline vs. Follow-up Year 1: June 1994 vs. June 1996</u>		
Word Recognition	5.57	.001**
Spelling	4.20	.001**
Reading Comprehension	2.44	.02†
<u>Baseline vs. Follow-up Year 2: June 1994 vs. June 1997</u>		
Word Recognition	3.95	.001**
Spelling	4.40	.001**
Reading Comprehension	2.36	.02†

Note: **p < .001, * p < .01, †.01 < p < .05

To find out if there were longer-term effects of the in-service PD, the scores from the literacy measures of children in the classrooms of experimental teachers one and two years after the in-service PD (June 1996 and 1997) were compared to the scores of children in the classrooms of teachers who taught during the year prior to the in-service PD (Baseline Year).

Since there were more than 3 comparisons ($k-1$) with the inclusion of an overall comparison for treatment effect, the level of significance was set at .0125 (.05/4) using the Bonferroni procedure. This was done in order to reduce the likelihood of a Type I error (Tabachnick & Fidell, 1996). Table 20 presents the t value and significance level of three comparisons.

Word Recognition

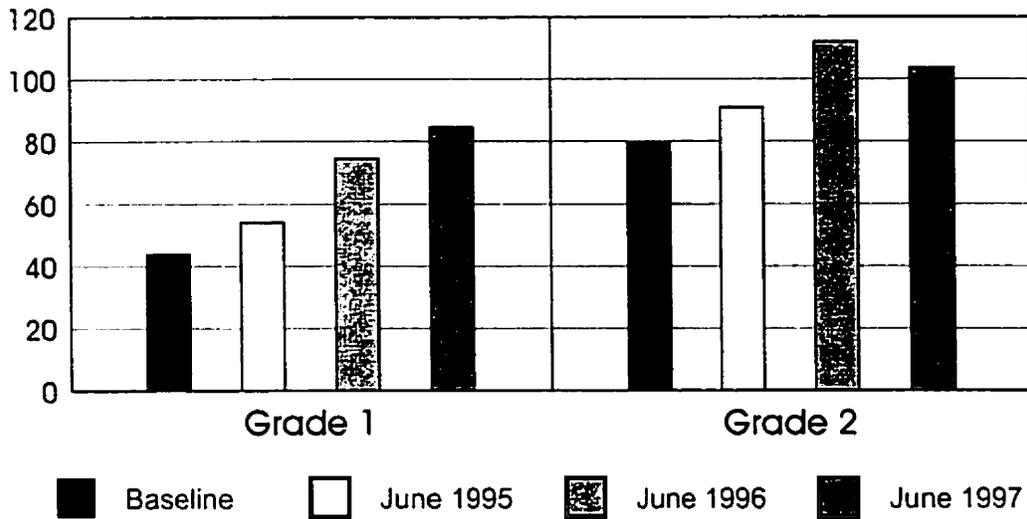
To determine whether there was an overall treatment effect of the professional development on the instructional skills of the teachers, the Baseline mean was compared with the averaged means of the combined Treatment years for the word recognition scores. The results showed an overall treatment effect of the professional development on the teaching skills of the teachers with associated improvement in the word recognition scores of the Grade 1 and Grade 2 children, $F(1, 164) = 14.59, p < .001$ and $F(1, 219) = 14.56, p < .001$, respectively.

As shown in Figure 2, the means on word recognition scores in the Treatment years were better than the Baseline Year for both Grade 1 and 2 children. There was a progressive growth in the number of words read from the Baseline Year to Follow-up Year 2 for Grade 1 children. In particular, the mean word recognition raw score of 84 (number of words recognized) from the *Grade 1* Experimental children in Follow-up Year 2 (June 1997) was above the mean word recognition score of 79 from the *Grade 2* children in the Baseline Year (June 1994).

Figure 2. Assessment of Teacher Effectiveness

Mean Number of Words Recognized

(Burns & Roe Informal Reading Inventory)



Grade 1. As seen in Figure 2, the mean number of correctly read words by Grade 1 children on the word recognition task of the Burns and Roe Informal Reading Inventory increased from Primer level almost to a Grade 1 level in the Intervention Year. The difference between the mean scores approached significance so the children's word recognition skills were improving, $t(183) = 2.31, p < .02$ (effect size, $d' = .44$). Subsequently, the word recognition scores of the Grade 1 children improved significantly to a mid-Grade 2 level in Follow-up Year 1 (effect size, $d' = .76$) and to an early Grade 3 level in Follow-up Year 2 (effect size, $d' = 1.10$), $t(183) = 3.49, p < .001$ and $t(183) = 5.37, p < .001$. The approximate number of words expected to be read by Grade 1 children on this word recognition task is 40 to 60 words, by Grade 2 children is 61 to 80 words and by Grade 3 children is 81 to 100 words. A linear trend analysis supported a significant improvement of word recognition scores for Grade 1 children from the Baseline year in 1994 across the years to the Follow-up Year 2 in 1997, $F = 31.35, p < .001$.

Grade 2. As seen in Figure 2, the word recognition scores of Grade 2 children progressively increased from the Baseline Year to Follow-up Year 1 (effect size, $d' = 1.0$). Subsequently, a slight decrease in the mean number of words read occurred in Follow-up Year 2, but the mean word recognition score remained at an early Grade 4 level (effect size, $d' = .72$). The mean number of correctly read words by Grade 2 children increased from a Grade 2 level in the Baseline Year to a mid-Grade 3 level in the Intervention Year, a difference that approached significance (effect size, $d' = .32$). Table 19 presents the mean number of words read by Grade 2 students and it shows significant increases in both Follow-up Years, to a mid-Grade 4 level in Follow-up Year 1, and an early Grade 4 level, in Follow-up Year 2, $t(233) = 5.57, p < .001$, and $t(233) = 3.95, p < .001$, respectively.

Summary. Both Grade 1 and Grade 2 teachers appear to have consolidated the knowledge they had acquired from the PD in the two years following the in-service. During the year of in-service, the teachers were learning new ideas and strategies each week, so there would be a gradual process of learning the new content and skill material. As discussed earlier, Mevarech (1995) and Guskey (1998) reported teachers feeling frustrated and more inexperienced at the start of implementing new skills learned from in-service PD. Probably most teachers did not make significant changes in their classroom practice for at least half of the year. Therefore, the lack of significant improvement in word recognition scores was not unexpected at the end of the year of in-service PD. However, in the years following the in-service the teachers would possess the new knowledge and strategies gained from the PD and they would be able to implement these new skills at the start of the school year. Thus, there would be an entire year with better instruction.

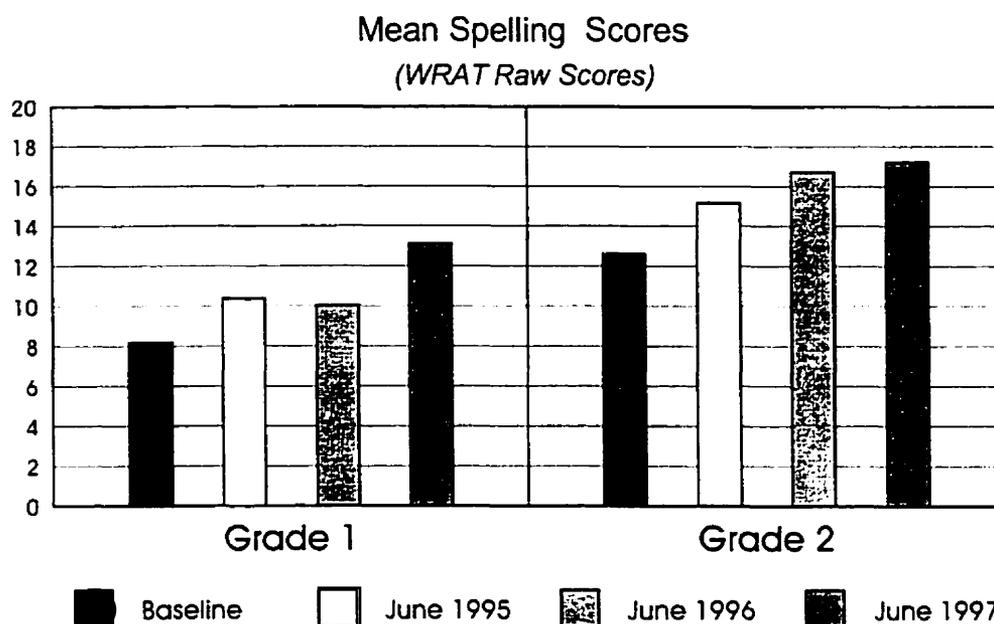
Overall, there was improvement in the word recognition scores of both Grade 1 and Grade 2 students. The mean number of words read by Grade 1 students improved to a mid-Grade 3 level by Follow-up Year 2. Similarly, the mean number of words read by Grade 2 students improved to the Grade 4 level by Follow-up Years 1 and 2. It seems clear that improved instruction by teachers who had participated in the professional development had a direct impact on the word recognition skills of children in their classrooms.

Spelling

A comparison of the Baseline mean with the averaged means of the combined Treatment years for the word spelling scores showed that there was an overall effect of teacher professional development on spelling outcomes for both Grade 1 and Grade 2 children, $F(1,164) = 19.05$, $p < .001$ and $F(1,219) = 17.61$, $p < .001$, respectively.

Figure 3 shows progressive improvement in the mean spelling scores of both Grade 1 and 2 children in the classrooms of Experimental teachers across the four years of the study. It is particularly noteworthy that the mean score of 13.10 correctly spelled words from the *Grade 1* Experimental children in the classrooms of in-serviced teachers two years following the in-service PD (June 1997) was above the mean score of 12.60 correctly spelled words from the *Grade 2* children in the Baseline Year (June 1994), before teachers participated in the PD.

Figure 3. Assessment of Teacher Effectiveness



Grade 1. There was significant increase from the Baseline Year to the Intervention Year, when the Grade 1 students' mean spelling scores reached the 79th percentile, $t(183) = 3.10$, $p < .002$. The teachers appeared to use the new knowledge and skills to improve their teaching of spelling immediately in their classroom practice during the in-service year (effect size, $d' = .61$).

Although the actual mean spelling score was nearly the same in Follow-up Year 1 as in the Intervention Year, there was only a small non-significant effect when Follow-up Year 1 scores were compared to the scores of Grade 1 children in the Baseline Year (effect size, $d' = .35$). This finding may be due to the smaller sample size of Grade 1 children in June 1996 ($n = 24$) – about one half of the size of the sample in the Intervention Year, since there was only one Grade 1 teacher who had participated in the in-service PD who taught Grade 1 students during Follow-up Year 1. However, in Follow-up Year 2 (June 1997), there was again a significant difference between the scores of Grade 1 children in the Experimental Group and the Baseline Group, $t(178) = 4.80$, $p < .001$ (effect size, $d' = .89$).

Grade 2. The mean spelling scores of the Grade 2 children showed consistent increases from the 32nd percentile in the Baseline Year to the 66th percentile in Follow-up Year 2 (see Figure 3). There were immediate effects (approaching significance) of teacher participation in the in-service on the spelling scores of Grade 2 children in the Intervention Year, with an improvement to the 45th percentile, $t(236) = 2.43, p < .02$ (effect size, $d' = .44$). The orthogonal comparisons also showed long-term effects of the PD on the teaching skills of the participating teachers since the Grade 2 children's mean spelling scores in Follow-up Years 1 and 2 were significantly better compared with the mean spelling scores of Grade 2 children in the Baseline Year, $t(236) = 4.20, p < .001$ and $t(236) = 4.40, p < .001$, respectively. The effect sizes were strong for Follow-up Year 1, $d' = .73$ and for Follow-up year 2, $d' = .81$. The sample sizes of the groups of Grade 2 children remained more stable over the years of the study. The progressive improvement in the spelling scores of the Grade 2 children over the four years of the study was reflected in a significant linear trend, $F = 21.76, p < .001$.

Summary. There was progressive significant improvement in the spelling scores of students in the classrooms of in-serviced teachers. The mean number of words spelled correctly, out of a total of 45 increased from 8 words (53rd percentile) in the classrooms of the Grade 1 Baseline Group to 13 words (91st percentile) in the classrooms of Grade 1 Experimental Groups by Follow-up Year 2 (June 1997). For the Grade 2 children, the mean number of correctly spelled words increased from 12 words (32nd percentile) in the classrooms of the Baseline Group to 17 words (66th percentile) in the classrooms of the Experimental Groups by Follow-up 2 Year (June 1997). Thus, there was continued improvement in the spelling skills of children in the classrooms of in-serviced teachers over the longer-term. The teachers apparently consolidated

the strategies learned from the professional development with the resulting improvement in the spelling scores of children they taught in their classrooms.

Reading Comprehension

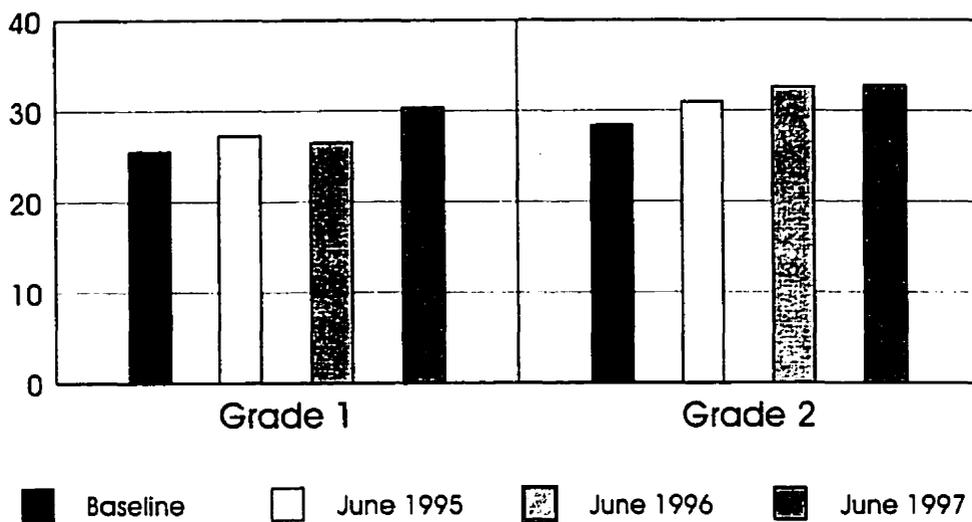
An overall analysis of the impact of the professional development on children's reading comprehension compared the means of the Baseline Years with the averaged means for the Treatment Years. The results showed that there was no significant improvement in the reading comprehension scores of children in the Grade 1 and Grade 2 classrooms of Experimental teachers, $F(1,164) = 2.29, p < .132$ and $F(1,219) = 2.78, p < .10$, respectively.

As shown in Figure 4 the levels of the reading comprehension scores of the children in both Grade 1 and Grade 2 Experimental Groups were higher than the Baseline Groups. For the measures of word recognition and spelling, a direct comparison of the raw scores was made between the levels attained by each grade during the four years of the study. This comparison was possible since the Grade 1 and Grade 2 children were administered the same test. However, for the measure of reading comprehension, the standardization required that Grade 1 and Grade 2 children complete different test forms. Thus, it was not possible to compare the raw scores attained for each year of the study between the Grade 1 and Grade 2 children.

Figure 4. Assessment of Teacher Effectiveness

Mean Reading Comprehension

(Gates-MacGinitie Raw Scores)



Grade 1. The mean reading comprehension scores of the Grade 1 children increased from the 48th percentile in the Baseline Year to the 55th percentile in the Intervention Year, declined to the 53rd percentile in Follow-up Year 1 and increased to the 61st percentile in Follow-up Year 2 (see Table 19). These variations in the reading comprehension scores of the Grade 1 children were not significant, although the scores approached significance in Follow-up Year 2, $t(175) = 2.02$, $p < .05$. The trend analysis was not significant. This result is not too surprising, as the focus of Grade 1 reading instruction is word analysis and recognition, not reading comprehension. Moreover, the type of reading comprehension measure employed was normed to show only large changes, as there are relatively few items.

Grade 2. The reading comprehension scores of the Grade 2 children, as shown in Figure 4, reflected a small but consistent pattern of improvement. The mean reading comprehension scores of Grade 2 children improved from the 30th percentile in the Baseline Year to only the 35th percentile in the Intervention Year, a small non-significant effect. Subsequently, however,

there were increases in Follow-up Years 1 and 2 to the 40th percentile (effect sizes, $d' = .43$ and $.44$, respectively). These improvements in reading comprehension scores both approached significance, $t(243) = 2.44, p < .02$ and $t(243) = 2.36, p < .02$ respectively.

In support of a pattern of consistent improvement in the reading comprehension scores of the Grade 2 children, a linear trend analysis showed significant improvement in the reading comprehension scores over the years from June 1994 to June 1997, $F = 6.35, p < .01$.

Summary. There was no significant improvement in the reading comprehension scores of the Grade 1 children. This result is not too surprising since the focus of Grade 1 instruction is word analysis. The reading comprehension scores of the Grade 2 children showed consistent improvement with instruction from teachers who had participated in the PD with longer-term benefits one and two years later.

Summary of Findings Concerning Teachers Overall

On the measures of word recognition, and reading comprehension the children in both Grade 1 and Grade 2 Experimental Groups showed only small insignificant improvement in their scores during the year their teachers participated in the in-service PD. It appeared that the teachers were gradually learning new skills and knowledge and may not have implemented changes in classroom practice until later in the school year. Thus, the impact of the in-service professional development was not immediately evident in the word recognition and reading comprehension scores of the children taught by the Experimental teachers. In contrast, the spelling scores of the Grade 1 children were significantly improved in the Intervention year compared to the Baseline year, and the spelling scores of the Grade 2 children during the year of the PD approached significant improvement. The participating teachers appeared to more easily apply new spelling strategies in their classroom instruction during the Intervention Year.

Children in both Grade 1 and Grade 2 Experimental groups did show longer-term effects of instruction by in-serviced teachers on their word recognition skills. There were significant differences in the word recognition scores between the Baseline Groups and the Experimental Groups in both Grades 1 and 2 in the two years following the year of the in-service PD period, in Follow-up Year 1 and Follow-up Year 2. The effect sizes were strong in both Follow-up Years for Grade 1 and Grade 2 children's word recognition skills. A significant linear trend supports the consolidation of skills from the PD as there was consistent improvement in the word recognition scores of Grade 1 children one and two years after the in-service. The Grade 1 children improved their word recognition scores from an early Grade 1 level to an early Grade 3 level. The scores of the Grade 2 children reflects an increase in the mean scores from a late Grade 2 level with a subsequent decrease, having reached a plateau at the Grade 4 level.

During Follow-up Year 1, the Grade 1 Experimental children's spelling scores showed only small non-significant improvement, compared to the group of children assessed in the Baseline year. As noted earlier, the sample size was reduced dramatically in that year since there was only one Grade 1 teacher who participated in the in-service, so the level of significance would have been affected. In Follow-up Year 2, the spelling scores of the Grade 1 Experimental Group – at the 91st percentile – were significantly better than the spelling scores of the Grade 1 Baseline Group, with a strong effect size of .89.

The effects on spelling at the Grade 2 level were clear and strong. There was a significant linear trend supporting the consolidation of skills from the PD with the accompanying improvement in classroom instruction over the longer-term. The spelling scores improved from the 32nd percentile in the Baseline Year to the 66th percentile in Follow-up Year 2 and showed a strong effect size of .81.

Furthermore, it is noted that the spelling measure is a very conservative test since it does not tap into the full extent of the children's knowledge. The child must get the word completely right in order to receive a point. There was no analysis of the progression toward more accurate spellings as the quality of classroom teaching improved. In an analysis of the spelling development of some of the children in this same sample, however, consistent improvements in their spelling skills, from random letter strings through to phonetically approximations to correct spellings, were shown from the Baseline Year to Follow-up Year 1 (Haas & Willows, 1998).

In reading comprehension, the linear trend of the reading comprehension scores of the Grade 2 children was significant. Therefore, there were progressive improvements in the reading comprehension skills of children in the Grade 2 Experimental Groups over the years of the study. The number of questions correctly answered by the Grade 2 children increased from the 30th percentile to the 40th percentile.

Thus teachers' participation in in-service professional development did affect the learning outcomes of their Grade 1 and Grade 2 students on the literacy measures of word recognition and spelling. The instruction of Grade 2 in-serviced teachers also impacted the reading comprehension achievement outcomes of their students. It is particularly noteworthy that the significant improvements in the literacy achievement levels of the Grade 1 and Grade 2 students in the classrooms of in-serviced teachers were progressive over the years. It appears that as teachers consolidated their knowledge, they improved their classroom literacy practice, and there was improvement in the acquisition of literacy skills by their students.

Tracking the Teachers: Part II

The subsidiary part of the first Research Question focused on the performance of children in the classrooms of teachers who remained in the same grade across all four years, Baseline Year (June 1994), Intervention Year (June 1995), Follow-up Year 1 (June 1996) and Follow-up Year 2 (June 1997). Thus, this research question is a longitudinal study following the same teachers across the years of the study.

This subsidiary Research Question can be stated as follows:

- 1 (b). For the three participating teachers who remained in the same grade across all four years, did the student outcomes in word recognition, spelling and reading comprehension in the year of the in-service PD (June 1995) in these Grade 1 and Grade 2 classrooms improve in relation to the pretreatment baseline obtained one year earlier (June 1994), and were there longer-term effects in the two years subsequent to the in-service (June 1996 and June 1997)?

Participant Characteristics

Teachers

The experimental teachers TA, TE and TF had participated in the in-service PD. These three teachers were part of the previous analysis on teacher change and they were chosen for further analyses as case studies since they had remained in the same grade across all four years of the study. Teacher A was a Grade 1 teacher, and Teachers E and F were Grade 2 teachers.

Because this question involved within teacher longitudinal analyses variation due to individual differences among teachers was eliminated. The analyses of successive groups of children – across years, within grade – provide an indication of the impact of being in the classrooms of particular teachers. Comparisons across the 4 years provide an indication of whether there was an immediate impact of being in the classrooms of these teachers (TA, TE and TF) during the year of in-service (1995 vs. 1994 Baseline Year) and a longer-term impact in the two years subsequent to the professional development (1996 vs. 1994 and 1997 vs. 1994).

Children

The characteristics and overall numbers of the sample of Experimental children assessed in each of the three participating teachers' classrooms across the four years of the longitudinal study are presented in Table 21.

There were no significant differences between the mean ages of the Grade 1 and Grade 2 children in the classrooms of these three teachers over the four years of the study.

Table 21

Summary Table of Participant Characteristics in the Within-Teachers Longitudinal Study

			Teacher A Grade 1			
	June 1994 $\underline{n} = 18$	June 1995 $\underline{n} = 14$	June 1996 $\underline{n} = 25$	June 1997 $\underline{n} = 17$		
Age (months)	\underline{M} 83.3 (3.33)	\underline{M} 81.86 (2.8)	\underline{M} 82.58 (4.89)	\underline{M} 83.12 (3.26)	\underline{F} .549	$p <$.65
			Teacher E Grade 2			
	June 1994 $\underline{n} = 24$	June 1995 $\underline{n} = 20$	June 1996 $\underline{n} = 26$	June 1997 $\underline{n} = 15$		
Age (months)	\underline{M} 96.58 (3.99)	\underline{M} 94.10 (2.65)	\underline{M} 95.64 (4.66)	\underline{M} 96.04 (4.32)	\underline{F} 1.49	$p <$.225
			Teacher F Grade 2			
	June 1994 $\underline{n} = 23$	June 1995 $\underline{n} = 19$	June 1996 $\underline{n} = 26$	June 1997 $\underline{n} = 19$		
Age (months)	\underline{M} 94.78 (3.54)	\underline{M} 94.21 (3.66)	\underline{M} 95.92 (4.39)	\underline{M} 92.99 (2.86)	\underline{F} 2.68	$p <$.052

Analyses

In order to evaluate the immediate and longer-term effects of the in-service PD on the outcomes of students in the classrooms of these three specific teachers, orthogonal comparisons and trend analyses were performed. Effect sizes were calculated. The results of these analyses will be presented for the children in each of teacher's classroom over the four years of the study for each of the dependent measures: word recognition, spelling and reading comprehension.

The overall effect of the professional development on the teaching skills of each teacher was examined by comparing the scores of the three dependent measures from the Baseline Year with the scores from the combined Treatment Years: Intervention Year, Follow-up Year 1 and Follow-up Year 2.

Subsequent planned comparisons examined the immediate effects of the in-service PD on the learning outcomes of the children in the classrooms of these three teachers by looking at the differences in the scores of children they taught in the pretreatment Baseline Year (1994) with the scores of the children they taught during the Intervention Year of in-service education (1995). The longer-term effects of the children's scores on the literacy measures were examined by comparing the scores on the literacy measures for children in their classrooms one and two years after the in-service PD with the scores of the children in the Baseline Year. There were 3 subsequent orthogonal comparisons (K-1). The total number of comparisons was four. Therefore, using the Bonferroni method to reduce the likelihood of a Type I error, each comparison was measured at a significance level of .0125 (.05/4). By following the scores of children in the classrooms of the same in-serviced teachers, there is less error due to individual differences between teachers.

Results and Discussion

Grade 1 Teacher A

Table 22 presents the Means and Standard deviations for the scores of the Grade 1 children in the classroom of Teacher A over the four test periods.

Table 22

Teacher A: Assessment of Teacher Effectiveness: Part II Means and Standard Deviations of Three Literacy Measures

Teacher A: Grade One

	June 1994	June 1995	June 1996	June 1997
Word Recognition	$\bar{n} = 18$ <u>M</u> 55.11 (41.51)	$\bar{n} = 14$ <u>M</u> 61.21 (33.71)	$\bar{n} = 22$ <u>M</u> 74.5 (44.95)	$\bar{n} = 17$ <u>M</u> 91.24 (34.33)
Spelling	$\bar{n} = 15$ <u>M</u> 8.40 (5.57)	$\bar{n} = 13$ <u>M</u> 10.38 (4.89)	$\bar{n} = 17$ <u>M</u> 10.04 (5.34)	$\bar{n} = 17$ <u>M</u> 14.59 (5.13)
Reading Comp.	$\bar{n} = 15$ <u>M</u> 27.27 (14.61)	$\bar{n} = 12$ <u>M</u> 28.08 (11.61)	$\bar{n} = 23$ <u>M</u> 26.52 (11.76)	$\bar{n} = 17$ <u>M</u> 32.29 (10.55)

Note: All scores are raw scores.

Table 23 presents the significance level for the orthogonal comparisons. The overall comparisons (Baseline vs. Treatment years combined) showed significant improvement in the mean spelling scores of the children in the classroom of Teacher A during the Treatment Years, $F(1,60) = 8.38, p < .005$.

Table 23

Teacher A: Immediate and Longer-Term Effects of In-Service PD
on Teacher Effectiveness

Orthogonal Comparisons between the Baseline and Experimental Groups of Children
Immediately after the In-Service PD and in the Follow-up Years

Measure	t	Significance
Immediate Effects		
<u>Baseline Year vs. Intervention Year: June 1994 vs. June 1995</u>		
Word Recognition	.43	.67
Spelling	1.00	.32
Reading Comprehension	.17	.86
Longer-Term Effects		
<u>Baseline Year vs. Follow-up Year 1: June 1994 vs. June 1996</u>		
Word Recognition	1.54	.13
Spelling	.95	.35
Reading Comprehension	.19	.85
<u>Baseline Year vs. Follow-up Year 2: June 1994 vs. June 1997</u>		
Word Recognition	2.75	.008*
Spelling	3.32	.001**
Reading Comprehension	1.17	.25

Note: ** $p < .001$, * $p < .01$.

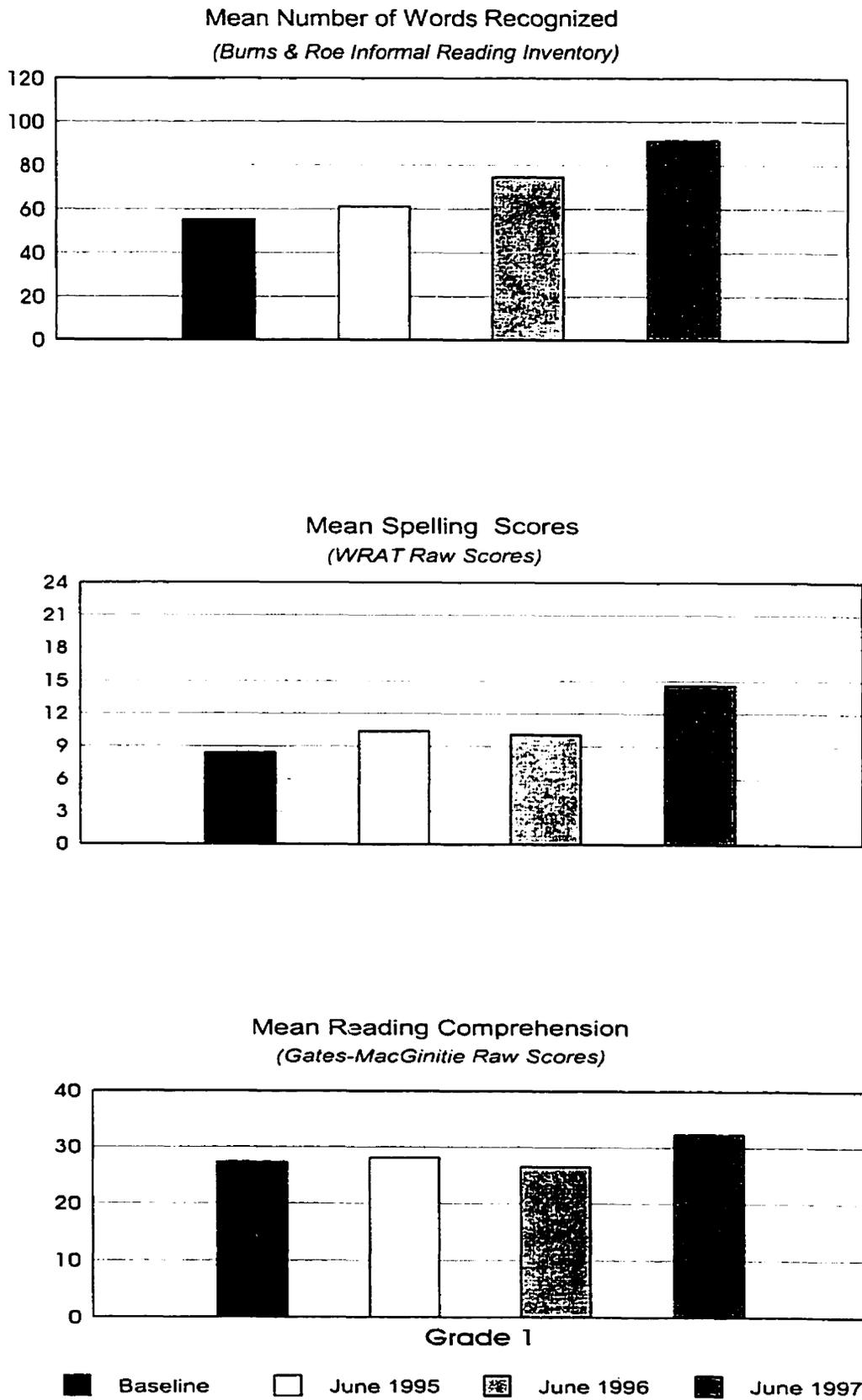
Word Recognition. As shown in Table 23, in comparison with word recognition scores in the Baseline Year, only Follow-up Year 2 produced results that were statistically significant (effect size, $d' = .95$). However, as shown in Figure 5, there was a progressive increase in the mean word recognition scores of children in the class of Teacher A across the 4 years of the study from a Preprimer level to mid-Grade 3 level. A trend analysis confirmed that there was a significant linear improvement over the years of the study, $F(1, 67) = 8.50$, $p < .005$.

Spelling. Similarly, in spelling, as shown in Table 23, only Follow-up Year 2 produced results that were statistically significant (effect size, $d' = 1.16$). However, Figure 5 shows a trend of an increasing number of correctly spelled words from the 45th percentile in the Baseline Year to the 96th percentile in Follow-up Year 2. Again, the trend analysis confirmed that there was a significant linear improvement over the years of the study, $F(1, 65) = 9.60, p < .003$.

Reading Comprehension. The orthogonal comparisons of the reading comprehension scores showed no significant differences in the reading comprehension scores of children in the classroom of Teacher A in the Intervention Year, Follow-up Year 1 and Follow-up Year 2 compared with the Baseline year. However, as seen in Figure 5 there was a small non-significant improvement in the mean number of correctly answered questions, from the 55th percentile in the Baseline Year (June 1994) to the 65th percentile in Follow-up Year 2 (June 1997). There were no significant trends.

Summary. The results show that Teacher A improved the effectiveness of her classroom instruction after participating in the in-service PD. The improvement was progressive, resulting in increasingly better performance in the word recognition and spelling of students in her classroom across the years. Their reading comprehension scores did not, however, show significant improvement.

Figure 5. Assessment of Teacher Effectiveness: Teacher A



Grade 2: Teacher E

Table 24 presents the outcomes on the literacy measures for the children in the Grade 2 classrooms of Teacher E across the years she participated in this study.

Table 24. Teacher E: Assessment of Teacher Effectiveness: Part II
Means and Standard Deviations of the Three Literacy Measures

Teacher E: Grade Two				
	June 1994	June 1995	June 1996	June 1997
Word Recognition	$\underline{n} = 22$ \underline{M} 72.18 (39.22)	$\underline{n} = 20$ \underline{M} 84.25 (40.18)	$\underline{n} = 21$ \underline{M} 103.1 (28.4)	$\underline{n} = 15$ \underline{M} 112.2 (33.72)
Spelling	$\underline{n} = 22$ \underline{M} 11.23 (5.94)	$\underline{n} = 18$ \underline{M} 14.11 (5.91)	$\underline{n} = 22$ \underline{M} 16.31 (5.24)	$\underline{n} = 15$ \underline{M} 20.73 (6.21)
Reading Comp.	$\underline{n} = 23$ \underline{M} 25.43 (10.81)	$\underline{n} = 20$ \underline{M} 32.15 (13.60)	$\underline{n} = 22$ \underline{M} 27.05 (9.58)	$\underline{n} = 15$ \underline{M} 36.47 (8.6)

Note: All scores are raw scores.

Table 25 presents the significance level for the orthogonal comparisons. The overall comparisons (1994 vs. 1995+1996+1997) showed significant improvement in the mean word recognition and mean spelling scores of children in the classroom of Teacher E in the Intervention year and subsequent Follow-up Years compared to the Baseline Year, $F(1,70) = 5.02$, $p < .03$ and $F(1,70) = 10.06$, $p < .002$, respectively.

Word Recognition. As shown in Figure 6, the word recognition levels of the children in the Experimental Groups taught by Teacher E showed consistent improvement from the Baseline Year through to Follow-up Year 2. In the first year, immediately after the professional development (June 1995) the number of words read improved from a mid-Grade 2 level in the Baseline Year to a beginning Grade 3 level in the Intervention Year (1995). The trend of the number of words read continued to significantly improve to a beginning Grade 4 level, in Follow-up Year 1 (effect size, $d' = .91$) to a significant improvement of 112 words read in Follow-up Year 2, a mid Grade 4 level (effect size, $d' = 1.10$). Thus it appears that Teacher E consolidated the knowledge she had learned about teaching word recognition from the in-service PD during the two years following the instruction.

This is reflected in the significant orthogonal comparisons between the word recognition scores in the Baseline year (1994) and the two follow-up years (1996 and 1997), $t(74) = 2.83, p < .006$ and $t(74) = 3.34, p < .001$, respectively. The linear trend analysis confirms the consistent improvement in the word recognition scores of her students over the period of time from the Baseline Year (1994) to Follow-up Year 2 (1997), $F(1,74) = 13.58, p < .001$.

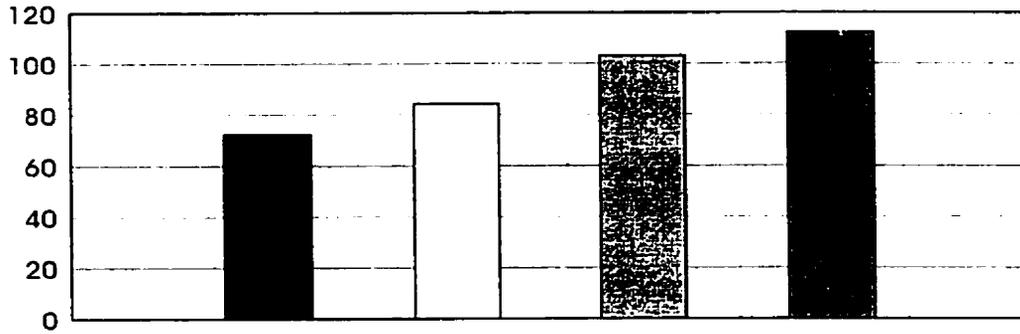
Table 25

Teacher E: Immediate and Longer-Term Effects of In-Service PD
on Teacher Effectiveness

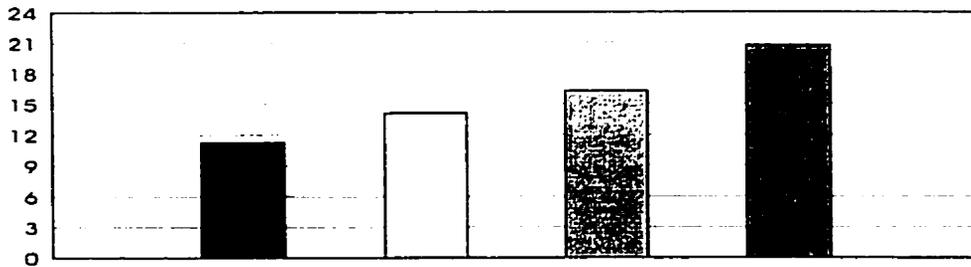
Orthogonal Comparisons between the Baseline and Experimental Groups of Children
Immediately after the In-Service PD and in the Follow-up Years

Measure	t	Significance
Immediate Effects		
<u>Baseline Year vs. Intervention Year: June 1994 vs. June 1995</u>		
Word Recognition	1.09	.28
Spelling	1.57	.12
Reading Comprehension	2.01	.05
Longer-Term Effects		
<u>Baseline Year vs. Follow-up Year 1: June 1994 vs. June 1996</u>		
Word Recognition	2.83	.006*
Spelling	2.91	.005*
Reading Comprehension	.50	.622
<u>Baseline Year vs. Follow-up Year 2: June 1994 vs. June 1997</u>		
Word Recognition	3.34	.001**
Spelling	4.90	.001**
Reading Comprehension	3.05	.003*
<u>Note: *p < .01, **p < .001.</u>		

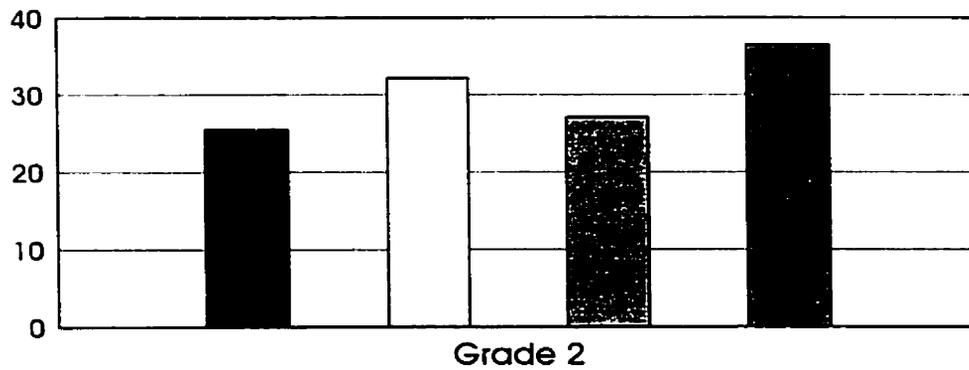
Figure 6. Assessment of Teacher Effectiveness: Teacher E
 Mean Number of Words Recognized
(Burns & Roe Informal Reading Inventory)



Mean Spelling Scores
(WRAT Raw Scores)



Mean Reading Comprehension
(Gates-MacGinitie Raw Scores)



Baseline
 June 1995
 June 1996
 June 1997

Spelling. As shown in Figure 6 there was also a progressive increase in the mean spelling scores of children in the class of Teacher E across the 4 years of the study. There was no significant immediate improvement in the spelling scores of children in the Intervention Year, although their mean scores increased from the 25th percentile in the Baseline Year to the 45th percentile in the Intervention Year. There were, however, significant longer term effects of the in-service on the spelling outcomes of the children in Teacher's E classroom during Follow-up Year 1 and Follow-up Year 2, $t(73) = 2.91, p < .005$ (effect size, $d' = .91$), and $t(73) = 4.90, p < .001$ (effect size, $d' = 1.56$), respectively as the mean spelling scores increased to the 86th percentile by Follow-up Year 2. This progressive improvement was reflected in a significant linear trend, $F(1,73) = 25.33, p < .001$.

Reading Comprehension. As seen in Figure 6, the scores on the reading comprehension measure were more scattered, and did not fit a trend of consistent improvement, although the number of correctly answered questions significantly improved from the 24th percentile in the Baseline Year to the 50th percentile in Follow-up Year 2 (effect size, $d' = 1.14$). In the Intervention Year (40th percentile), the improvement over the Baseline year was only marginally significant ($p < .05$) and in Follow-up Year 1 (28th percentile) there was no significant gain over the Baseline Year. In Follow-up Year 2, however, the mean score was substantially above the Baseline Year, $t(76) = 3.05, p < .003$.

Summary. The results show that Teacher E improved the effectiveness of her classroom practice after participation in in-service PD. The improvement was progressive, with substantial growth in the word recognition and spelling scores of children in the classroom of Teacher E during the two Follow-up years. There were small gains in the mean reading

comprehension scores of her students over the years of the study, with a statistically significant gain by Follow-up Year 2.

Grade 2: Teacher F

Table 26 presents the Means and Standard Deviations of the scores on the literacy measures of the children in the Grade 2 classrooms of Teacher F across the years she participated in the study.

Table 26

Teacher F: Assessment of Teacher Effectiveness: Part II
Means and Standard Deviations of the Three Literacy Measures

Teacher F: Grade Two

	June 1994	June 1995	June 1996	June 1997
Word Recognition	$\underline{n} = 22$ \underline{M} 84.36 (39.45)	$\underline{n} = 17$ \underline{M} 94.65 (27.82)	$\underline{n} = 21$ \underline{M} 109.76 (18.63)	$\underline{n} = 19$ \underline{M} 92.95 (32.0)
Spelling	$\underline{n} = 20$ \underline{M} 13.5 (5.86)	$\underline{n} = 17$ \underline{M} 14.12 (5.25)	$\underline{n} = 26$ \underline{M} 16.27 (4.92)	$\underline{n} = 19$ \underline{M} 15.11 (4.57)
Reading Comp.	$\underline{n} = 23$ \underline{M} 30.57 (10.23)	$\underline{n} = 18$ \underline{M} 30.22 (9.81)	$\underline{n} = 26$ \underline{M} 35.88 (7.13)	$\underline{n} = 19$ \underline{M} 30.26 (10.39)

Note. All scores are raw scores.

Table 27 presents the significance level of the orthogonal comparisons. The overall comparisons (1994 vs. 1995+1996+1997) did not show any significant improvement in the mean word recognition, spelling and reading comprehension scores of children in the classroom of Teacher F in the Intervention Year and subsequent Follow-up years compared to the Baseline Year.

Table 27

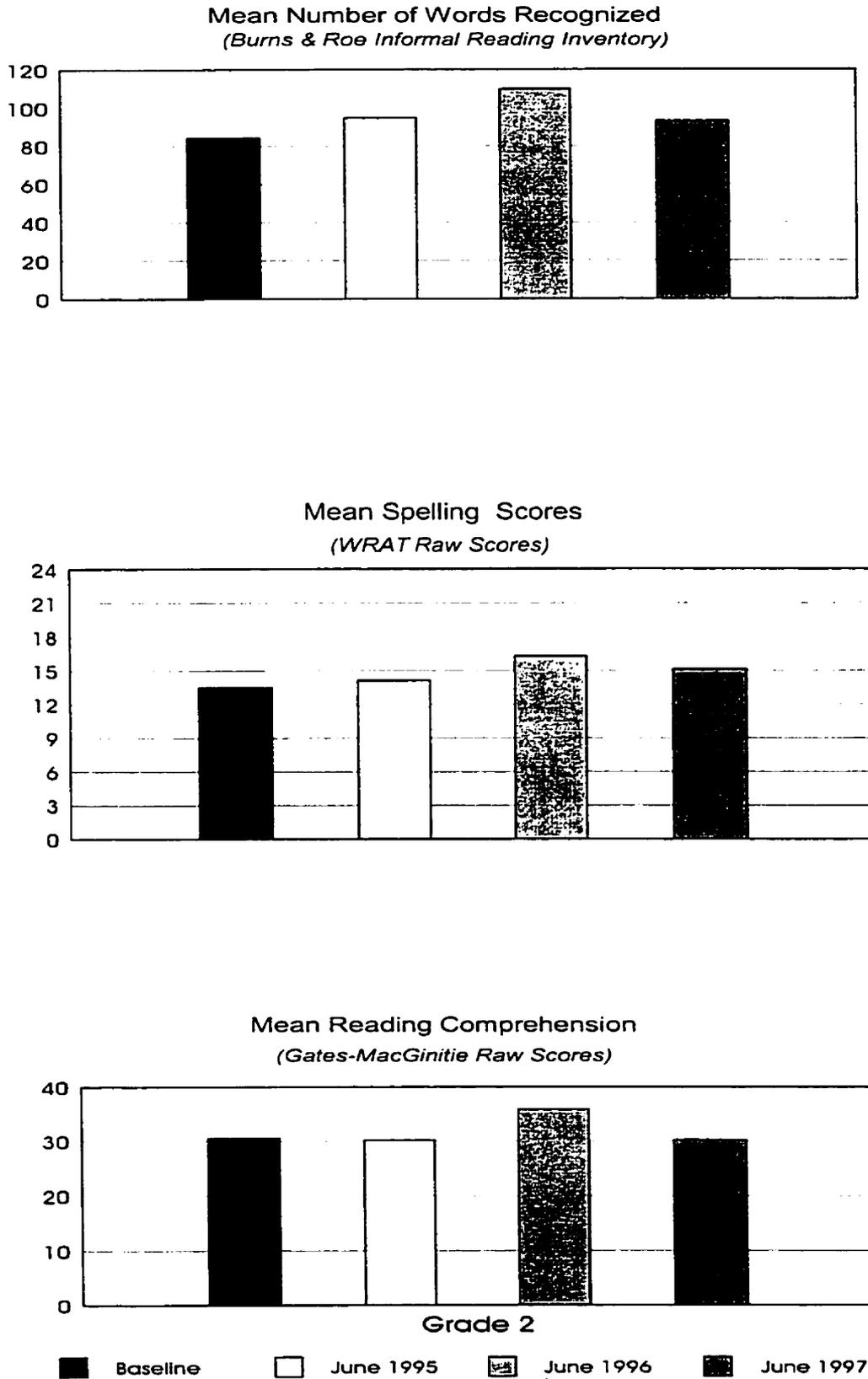
Teacher F: Immediate and Longer-Term Effects of In-Service PD on Teacher Effectiveness

Orthogonal Comparisons between the Baseline and Experimental Groups of Children Immediately after the In-Service PD and in the Follow-up Years

Measure	t	Significance
Immediate Effects		
<u>Baseline Year vs. Intervention Year: June 1994 vs. June 1995</u>		
Word Recognition	1.07	.29
Spelling	.36	.72
Reading Comprehension	.12	.91
Longer-Term Effects		
<u>Baseline Year vs. Follow-up Year 1: June 1994 vs. June 1996</u>		
Word Recognition	2.81	.006*
Spelling	1.81	.08
Reading Comprehension	1.99	.05
<u>Baseline Year vs. Follow-up Year 2: June 1994 vs. June 1997</u>		
Word Recognition	.93	.36
Spelling	.97	.33
Reading Comprehension	.10	.92

Note. * $p < .01$, ** $p < .001$.

Figure 7. Assessment of Teacher Effectiveness: Teacher F



Word Recognition. As shown in Figure 7, the word recognition scores of the children in the Experimental Groups taught by Teacher F showed consistent improvement from the Baseline Year through to Follow-up Year 1, with a subsequent decrease in the mean scores in Follow-up Year 2. Although as shown in Table 27, in comparison with word recognition scores in the Baseline Year, only Follow-up Year 1 produced results that were statistically significant (effect size, $d' = .44$). However, the number of words read improved from a beginning Grade 3 level in the Baseline Year to a mid-Grade 3 level in the Intervention Year. The trend of the number of words read by the children in her classroom continued to significantly improve to a beginning Grade 4 level, in Follow-up Year 1. Therefore, the longer-term effects of the in-service PD on the instruction of Teacher F was demonstrated with significant improvement in the scores of the children she taught one year following the in-service, Follow-up Year 1, $t(75) = 2.81, p < .006$. Although, the number of mean words read by the Grade 2 children in her classroom in Follow-up Year 2 decreased to a mean of 92, this is still a mid-Grade 3 level.

Spelling. Although as shown in Figure 7, there were small increases in the number of correctly spelled words in the Intervention Year and Follow-up Years compared to the Baseline Year, these were not significant increases. There was a non-significant increase in the mean number of correctly spelled words by the children taught by Teacher F from the Baseline Year at the 32nd percentile to Follow-up Year 1 at the 58th percentile. However, similar to the trend of scores on the word recognition task, there was a subsequent decrease in Follow-up Year 2 to the 53rd percentile. There were no significant trends.

Reading Comprehension. Figure 7 shows the scores on the reading comprehension measure as more scattered, and these scores did not fit a trend of consistent improvement. In comparison to the Baseline Year, there were no improvements in the mean reading comprehension scores in the Intervention Year, which remained at the 35th percentile. There was a small non-significant gain in Follow-up Year 1 to the 48th percentile, with a decrease to the same level as in the Baseline Year, the 35th percentile in Follow-up Year 2. It was noted that the mean reading comprehension score of the Grade 2 children in the classroom of Teacher F in the Baseline Year was higher at the 35th percentile compared to the mean reading comprehension score of the Grade 2 children in the classroom of Teacher E at the 24th percentile. Therefore, there was more room for improvement by the children in the classroom of Teacher E. There were no significant trends for the reading comprehension scores of children taught by Teacher F.

Summary. These results do not demonstrate significant immediate effect or longer-term effects of the in-service PD on the classroom practice of Teacher F for teaching word recognition, spelling and reading comprehension skills. There were trends of improvement in word recognition, spelling and reading comprehension scores up to Follow-up Year 1, and a significant difference in the word recognition scores of the children in her classroom one year after the in-service compared to the children in her classroom during the Baseline Year. However, this longer-term effect did not persist into the second follow-up year, where the scores decreased to grade level in spelling and a mid-grade 3 level in word recognition. During the Baseline year, the scores of the Grade 2 children on the literacy measures in the classroom of Teacher F were higher than the scores of the Grade 2 children in the classroom of Teacher E. It appears that Teacher F's instructional practices were already improving the literacy skills of her students before she participated in the in-service, and her classroom instruction did not change

very much. Furthermore, a limitation of this study is the confounding of teachers instructional abilities with the children's learning abilities. It is not known if Teacher F had a particularly difficult class of children to teach in Follow-up year 2 or whether she lacked the motivation to continue to apply the knowledge and skills gained from the professional development.

Summary of Findings Concerning Individual Teachers

The subsidiary question involving tracking teachers, examined how well three teachers applied the knowledge from the in-service PD of the *Balanced and Flexible Literacy Diet* system. The results demonstrated the variability of the effect of in-service PD on the classroom practice of teachers. It seems likely that there were differences in the instructional practices of these three teachers, which resulted in the different outcomes of their students on the three literacy measures. It appears that some teachers profited more than others from the in-service, and transferred the benefits to their students. Teacher E's students clearly made substantial gains in all three literacy measures of word recognition, spelling and reading comprehension over the years of the study, with significant improvement in Follow-up Years 1 and 2. Teacher A's students showed trends toward improvement in their word recognition and spelling skills with significant improvement in Follow-up Year 2. Teacher F's students showed some gains, with significant improvement only in word recognition in Follow-up Year 1, but these gains did not continue into Follow-up Year 2. To some extent, the in-service PD seems to have influenced the classroom practices of all three teachers, but the strength of the impact varied among these three teachers, as shown by the learning outcomes of their students. This may suggest that it could be important to examine further the impact of in-service PD on individual teachers to determine why some benefit more than others.

The Balanced and Flexible Literacy Diet PD had more impact on word recognition and spelling instructional practices than reading comprehension practices. All three teachers showed some gains in the word recognition and spelling scores of their students. There was no significant improvement in the reading comprehension skills of students taught by Teacher A and Teacher F. It is clear, however, that Teacher E's improved teaching practices resulted in higher achievement levels among her students in reading comprehension, as well. Thus, it is apparent that teachers can improve the reading comprehension skills of their students, but some teachers are more able to do this than others. It would be useful to have classroom observations to examine the teaching practices of teachers who are more effective in teaching reading comprehension skills.

Whereas Research Questions 1(a) and 1 (b) were designed to track the group and individual teachers who participated in the in-service to determine the impact of the PD on them as teachers, Research Question 2, was designed to track children to determine if there were cumulative benefits of having more exposure to the instruction of teachers who had received in-service PD. The sample of children for Research Question 2 was smaller because it consisted of only those children who were consistently taught by teachers who participated in the PD.

CHAPTER 6

Tracking the Children: Study of Student Change

In this chapter the impact of the in-service PD, *The Balanced and Flexible Literacy Diet*, is assessed with respect to improvement in the literacy scores of the children as they pass through the classrooms of in-serviced teachers. Within this context the sample of children is reduced to only those who continued in the classrooms of in-serviced teachers for multiple years. There is more control of error due to differences between children, since only those children who were consistently taught by teachers who participated in the in-service PD are included in the samples.

There are seven cohorts of children examined in this longitudinal study who are grouped together according to their Grade level (Grade 1 or Grade 2) and the number of years of exposure to the instruction of in-serviced teachers (one, two, three or four years).

The focus of this chapter is the effect of in-service teacher professional development on the literacy outcomes of children on word recognition, spelling and reading comprehension as they progress through grades from Junior Kindergarten to Grade 1 or Grade 2. This chapter examines the cumulative learning effects of increasing exposure to the instruction of teachers who participated in the in-service PD using *The Balanced and Flexible Literacy Diet* system.

This research examined the literacy outcomes of children in Grade 1 who were exposed for 1, 2 or 3 years to the balanced literacy instruction of in-serviced teachers and compared their results to the scores of Grade 1 children in the Baseline Year (0 years of exposure). Similarly, the literacy outcomes of children in Grade 2 who were exposed for 1, 2, 3 and 4 years to the

balanced literacy instruction of in-serviced teacher were compared to the results of Grade 2 children in the baseline (0 years of exposure).

This chapter will begin with a description of the children tracked through the classrooms of Experimental teachers. The analyses, results and discussion will follow.

Tracking the Children

The overall Research Question addressed in this chapter can be stated as follows:

2. Was the performance of Grade 1 and Grade 2 children who had one, two, three or four years of ongoing instruction by teachers who had participated in the professional development significantly better on measures of word recognition, spelling, and reading comprehension than that of children who had no exposure to in-serviced teachers?

Participant Characteristics

Children

As shown in Tables 28 and 29, the children tracked were those who were taught by in-serviced teachers over 3 (Grade 1) or 4 years (Grade 2). Cohort 1A were children in Grade 1 who were taught for one year by Grade 1 teachers who had participated in the in-service during the 1994/1995 school year. Cohort 1B were Grade 1 children who were taught for two years by in-serviced teachers, Senior Kindergarten teachers in 1995, and Grade 1 teachers in 1996. Cohort 1C were Grade 1 children who were taught for three years by in-serviced teachers, Junior Kindergarten in 1995, Senior Kindergarten teachers in 1996, and Grade 1 teachers in 1997. Similarly, for the Grade 2 children, Cohort 2A were children taught for one year by Grade 2 teachers who participated in the in-service during the 1994/1995 school year. Cohort 2B were

Grade 2 children who were taught for two years by in-serviced teachers, Grade 1 teachers in 1995 and Grade 2 teachers in 1996. Cohort 2C were Grade 2 children who were taught for three years by in-serviced teachers, Senior Kindergarten teachers in 1995, Grade 1 teachers in 1996 and Grade 2 teachers in 1997. Cohort 2D were Grade 2 children who were taught for four years by teachers who had participated in the in-service education, Junior Kindergarten teachers in 1995, Senior Kindergarten teachers in 1996, Grade 1 teachers in 1997 and Grade 2 teachers in 1998.

As shown in Table 28, Cohorts 1A and 2B are the same children as are Cohorts 1B and 2C, and Cohorts 1C and 2D, since these Grade 1 children became Grade 2 children in the following year.

Table 28 Cohorts of Children in Grade 1 and Grade 2 Exposed to the Instruction of In-serviced Teachers for 1, 2, 3 and 4 Years

Grades	Years of Exposure			
	1 year (1995)	2 years (1996)	3 years (1997)	4 years (1998)
Grade 1	Cohort 1A	Cohort 1B	Cohort 1C	
Grade 2	Cohort 2A	Cohort 2B	Cohort 2C	Cohort 2D

Table 29 Description of the Cohorts in the Children's Longitudinal Study:
Children Taught by In-serviced Teachers

Years of Exposure		Assessed in Grade 1
1		Cohort 1A: Grade 1 (1995-treatment year)
2		Cohort 1B: Grade 1 (1996) and SK (1995)
3		Cohort 1C: Grade 1 (1997), SK (1996), JK (1995)
		Assessed in Grade 2
1		Cohort 2A: Grade 2 (1995-treatment year)
2		Cohort 2B: Grade 2 (1996) and Grade 1 (1995)
3		Cohort 2C: Grade 2 (1997), Grade 1 (1996), SK (1995)
4		Cohort 2D: Grade 2 (1998), Grade 1 (1997), SK (1996), JK (1995)

Note: JK is Junior Kindergarten and SK is Senior Kindergarten

Due to the transience of this population, the number of children in the sample decreased across the years of exposure. Furthermore, as shown in Table 30, there was considerable change and movement among the teachers who had participated in the in-service PD. In order to track the children who were exposed to the instruction of in-serviced teachers, the sample could only include children who had been consistently taught in the classrooms of in-serviced teachers. For example, there was only one Grade 1 teacher in 1996 who had participated in the in-service. Therefore, any children who were taught by other Grade 1 teachers in 1996 could not be included in the longitudinal sample of children. Thus, there is a marked reduction in the number of children in the Grade 2 sample of 1997, since only those children who had been taught by this one in-serviced Grade 1 teacher were tracked into the Cohort 2C, the 1997 Grade 2 sample.

Table 30 Grade 1 and Grade 2 Teachers from 1994 to 1998

Grades	Years of Exposure				
	0 years (1994)	1 year (1995)	2 years (1996)	3 years (1997)	4 years (1998)
Grade 1	TA	TA	TA	TA	
	TB	TB	NT	TD	
	TC	TI	NT	NT	
	TD				
Grade 2	TE	TE	TE	TE	TA
	TF	TF	TF	TF	TF
	TG	TG	TH	TH	TH
	TH				

Note: *NT* is new teacher who did not participate in the in-service PD.

This research question was addressed by comparing each of the cohorts with the pre-treatment baseline (1994) to assess the impact of successive years of instruction by in-serviced teachers for the groups of children in Grade 1 and Grade 2.

The characteristics of the children in this longitudinal study who were instructed by in-serviced teachers are presented in Table 31.

There was no difference between the mean ages of the control group (Baseline) of Grade 1 children and Cohorts 1A, 1B and 1C who were assessed June 1995, June 1996 and June 1997 respectively, $F(3, 152) = .288, p < .83$. However, there were significant differences between the mean ages of the baseline group of Grade 2 children and Cohorts 2A, 2B, 2C and 2D who were assessed June 1995, June 1996, June 1997 and June 1998 respectively, $F(4, 191) = 3.16, p < .02$. As Table 31 shows, this age effect resulted from the fact that the Grade 2 children in 1995 and 1997 were slightly younger than those in 1994, 1996 and 1998. Therefore, the main analyses included ANCOVAs to determine if age was a significant factor affecting the results.

Table 31 Summary of Participant Characteristics: Tracking the Children

Grade 1						
	Baseline (1994) (<u>n</u> = 67)	Cohort 1A (<u>n</u> = 52)	Cohort 1B (<u>n</u> = 17)	Cohort 1C (<u>n</u> = 20)		
Years of Exposure	0	1	2	3		
Sex	29 girls 38 boys	31 girls 21 boys	6 girls 11 boys	10 girls 10 boys		
Age (months)	<u>M</u> 83.07 (3.31)	<u>M</u> 83.4 (4.55)	<u>M</u> 82.41 (4.6)	<u>M</u> 83.25 (3.23)	<u>F</u> .288	<u>p</u> < .83

Grade 2							
	Baseline (1994) (<u>n</u> = 68)	Cohort 2A (<u>n</u> = 58)	Cohort 2B (<u>n</u> = 45)	Cohort 2C (<u>n</u> = 14)	Cohort 2D (<u>n</u> = 16)		
Years of Expos.	0	1	2	3	4		
Sex	26 girls 42 boys	29 girls 29 boys	28 girls 17 boys	6 girls 8 boys	7 girls 9 boys		
Age (month)	<u>M</u> 95.74 (3.66)	<u>M</u> 94.55 (3.88)	<u>M</u> 96.13 (4.73)	<u>M</u> 93.21 (3.45)	<u>M</u> 97.54 (4.36)	<u>F</u> 3.16	<u>p</u> < .02

Note: Years of Expos. = Years of Exposure to In-serviced Teachers.

Analyses

In order to evaluate the effect of length of exposure to the instruction of Experimental teachers on student literacy outcomes, analyses of variance (MANOVAs), orthogonal comparisons and trend analyses were performed. Effect sizes were calculated. The results of the scores on the three dependent measures of word recognition, spelling and reading comprehension will be presented for Grade 1 and Grade 2 children.

One of the assumptions of analysis of variance is homogeneity of variance. One of the common test for homogeneity of variance is the Levene's Test of Equality of Error Variances. The data from the Grade 1 students showed Levene's Test significant for the word recognition task, and the reading comprehension task. The Levene's Test of Equality of Error Variance was not significant for the spelling task. The Levene's Test was also significant for the scores of the Grade 2 children on word recognition and reading comprehension measures. Therefore, due to the significance level of Levene's Test, analyses were performed using both parametric and nonparametric statistics. Based on the results of the Kruskal-Wallis Test using Chi-Square analyses, there was very little difference between the parametric and non-parametric results on the word recognition and reading comprehension scores of the Grade 1 children: both statistics showed the word recognition scores as significant and the reading comprehension scores as not significant, the Chi-Squares being $\chi^2(3, N = 153) = 30.98, p < .001$ and $\chi^2(3, N = 145) = 3.81, p < .282$, respectively. Similarly, for the word recognition and reading comprehension scores of the Grade 2 children, the results of the parametric and nonparametric statistics showed negligible differences as both statistics showed the scores of the word recognition and reading comprehension measures as significant, the Chi-Squares being $\chi^2(4, N = 153) = 32.35, p < .001$, and $\chi^2(4, N = 145) = 13.71, p < .008$, respectively. Therefore, for simplicity only, the parametric statistics will be discussed further.

The data were examined for outliers with the scores transformed into z scores. The Grade 1 data showed 1 outlier in the Baseline year and in the Grade 2 data there were 2 outliers in the Baseline year and 1 outlier in Follow-up Year 2. Analyses of variance were performed with and without outliers. When these outliers were removed there was only a negligible difference in the multivariate level of significance. Therefore, the outliers were included in the analyses since these scores represented students who performed very well and students who performed very

poorly on the literacy measures, which likely reflect the variation of children's abilities in Grades 1 and 2 classrooms. In all cases, the analyses reported in the text were performed on raw scores.

Results and Discussion

Table 32 presents the Means and Standard Deviations of the scores of Grade 1 and Grade 2 children on the word recognition, spelling and reading comprehension tasks over the years of exposure (0, 1, 2, 3 or 4) to teachers who had participated in the in-service PD.

These data show a substantial decrease in the sample sizes of the Cohorts of children over the 4 years of this longitudinal study. It is difficult to retain a reasonable sample size in a longitudinal study over this period of time. There are two factors impacting the size of the sample in this study. First, there is the natural attrition of the children as families move to other neighbourhoods and children leave the school, and thus are no longer participating in the study. Secondly, children are eliminated from the study if they are placed in classrooms of teachers who did not participate in the in-service PD. The focus of this study is the literacy outcomes of children who are exposed to the consistent instruction of in-serviced teachers. Therefore, the rate of lost children in this sample is especially high due to these two compounding factors.

In order to avoid an inflated Type I error rate (Stevens, 1996) omnibus MANOVA analyses were first performed. Two multivariate analyses of variances (MANOVAs), one for Grade 1 and another for Grade 2, were performed to determine the effect of four different levels (0, 1, 2 or 3 years) of exposure to the instruction of in-serviced teachers on the literacy scores of Grade 1 children and the effect of five different levels (0, 1, 2, 3, or 4 years) of exposure to the instruction of in-serviced teachers on the literacy scores of Grade 2 children. Overall significant differences were found across the groups of children exposed to the instruction of in-serviced

teachers. The MANOVA results for Grade 1 and Grade 2 students were $F(9, 314) = 8.02, p < .001$, and $F(12, 423) = 3.65, p < .001$, respectively.

Table 32

The Means and Standard Deviations of the Scores on the Achievement Measures of Word Recognition, Spelling and Reading Comprehension for the Grades 1 and 2 Children Across Years of Exposure to Experimental Teachers

<u>Achievement Measures</u>		<u>Grade 1</u>			<u>Grade 2</u>		
		<u>M</u>	<u>SD</u>	<u>n</u>	<u>M</u>	<u>SD</u>	<u>n</u>
Word Recognition							
(Baseline)	0 year	43.9	(38.6)	66	79.3	(36.3)	63
(Cohort A)	1 year	58.5	(32.6)	51	90.9	(36.1)	56
(Cohort B)	2 years	93.8	(36.0)	16	117.2	(28.6)	38
(Cohort C)	3 years	82.7	(32.6)	20	107.6	(27.5)	14
(Cohort D)	4 years				107.5	(42.6)	11
Spelling							
(Baseline)	0 years	8.2	(5.1)	59	12.6	(5.9)	60
(Cohort A)	1 years	11.1	(4.2)	51	15.2	(5.7)	54
(Cohort B)	2 years	12.3	(4.7)	17	17.5	(5.0)	42
(Cohort C)	3 years	12.9	(5.3)	20	17.9	(4.8)	14
(Cohort D)	4 years				19.2	(7.7)	11
Reading Comprehension							
(Baseline)	0 years	25.6	(12.7)	59	28.4	(10.3)	66
(Cohort A)	1 year	28.1	(11.6)	49	31.0	(11.7)	55
(Cohort B)	2 years	30.3	(9.9)	17	34.6	(9.0)	41
(Cohort C)	3 years	31.2	(11.3)	20	34.2	(8.7)	14
(Cohort D)	4 years				36.9	(5.1)	15

Note. All scores are raw scores.

The ANCOVAs were run to determine if age was a factor affecting the outcomes of the scores of the Grade 2 children. There were negligible differences between the results of the ANCOVAs and the original analyses for the scores on the literacy measures of the Grade 2 children. The overall ANCOVA was $F(12, 420) = 3.62, p < .001$.

Based on the significant MANOVAs, overall comparisons were performed. These analyses compared the scores on the literacy measures of the Baseline group of children (0 years of exposure) with the Experimental groups of children (1, 2, 3, 4 years of exposure) to see if there was an effect of amount of exposure to the instruction of teachers who had participated in the in-service PD.

Further analyses with orthogonal comparisons and trend analysis of the dependent variables were performed to determine if there was an effect on the literacy outcomes of the children after exposure to the instruction of in-serviced teachers by comparing the scores of the Grade 1 and Grade 2 children who had no exposure to the instruction of in-serviced teachers in the Baseline year with those of children who had one, two, three or four years of exposure to the instruction of in-serviced teachers. There were 4 comparisons of the Grade 1 children and there were 5 comparisons of the Grade 2, when the overall orthogonal comparison examining treatment effect is included. Therefore, in order to reduce the likelihood of a Type I error the level of significance was set at .0125 for the orthogonal comparisons. Table 33 presents the t value and significance level of these comparisons.

Table 33 Assessment of Student Change:
Orthogonal Comparisons between the Baseline and Experimental Groups of Children who
had 1, 2, 3, or 4 years of Exposure to In-serviced Teachers

Grade 1		
Measure	t	Significance
1 Year of Exposure: June 1994 vs. June 1995 (Cohort 1A)		
Word Recognition	2.20	.03†
Spelling	3.13	.002*
Reading Comprehension	1.12	.263
2 Years of Exposure: June 1994 vs. June 1996 (Cohort 1B)		
Word Recognition	5.02	.001**
Spelling	3.11	.002*
Reading Comprehension	1.47	.144
3 Years of Exposure: June 1994 vs. June 1997 (Cohort 1C)		
Word Recognition	4.26	.001**
Spelling	3.80	.001**
Reading Comprehension	1.84	.07
Grade 2		
1 Year of Exposure: June 1994 vs. June 1995 (Cohort 2A)		
Word Recognition	1.81	.07
Spelling	2.38	.02†
Reading Comprehension	1.39	.17
2 Years of Exposure: June 1994 vs. June 1996 (Cohort 2B)		
Word Recognition	5.32	.001**
Spelling	4.31	.001**
Reading Comprehension	3.09	.002*
3 Years of Exposure: June 1994 vs. June 1997 (Cohort 2C)		
Word Recognition	2.77	.006*
Spelling	3.10	.002*
Reading Comprehension	1.97	.051
4 Years of Exposure: June 1994 vs. June 1998 (Cohort 2D)		
Word Recognition	2.48	.01*
Spelling	3.51	.001**
Reading Comprehension	2.96	.003*

Note: ** $p < .001$, * $p < .01$, † $.01 < p < .05$

Word Recognition

A comparison was performed between word recognition scores of children with no exposure to the instruction of teachers who had participated in the in-service PD in the Baseline year (0 years of exposure) and the word recognition scores of children who had been exposed to the instruction of in-serviced teachers for 1, 2, 3, or 4 years. The results showed an overall effect of exposure to the instruction of teachers who had participated in the in-service PD on the word recognition scores of Grade 1 and Grade 2 children, $F(1,133) = 11.80, p < .001$ and $F(1,165) = 13.15, p < .001$, respectively.

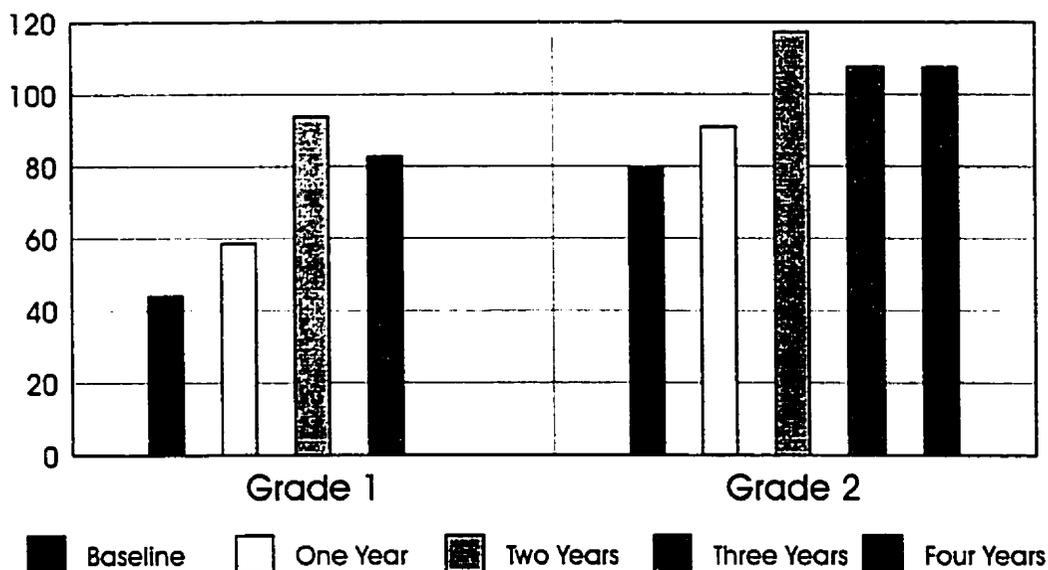
Grade 1. As seen in Table 32, the mean number of correctly read words by Grade 1 children in the Baseline Group increased from the Primer level to a late Grade 1 level after one year of exposure (Cohort 1A) to the instruction of in-serviced teachers. The difference between mean scores approached significance, $t(149) = 2.20, p < .03$ (effect size, $d' = .41$). Relative to the Baseline Group, the word recognition scores of the Grade 1 children improved significantly after two years of exposure (Cohort 1B) (effect size, $d' = 1.34$) and, after 3 years of exposure (Cohort 1C) (effect size, $d' = 1.09$) when instructed by in-serviced teachers, $t(149) = 5.02, p < .001$ and $t(149) = 4.26, p < .001$, respectively.

Grade 2. The mean word recognition scores of Cohort 2A increased to a mid Grade 3 level after one year of exposure to the instruction of a teacher who had participated in the in-service PD. Table 32 presents the mean number of words read by Grade 2 students. There was a significant gain, compared to the Baseline Group, after two years, three and four years of exposure to in-serviced teachers, $t(177) = 5.32, p < .001$, $t(177) = 2.77, p < .006$, and $t(177) = 2.45, p < .01$, respectively (see Table 33). The size of the effect of exposure to the instruction of teachers who had participated in the PD was strong for Cohort B (effect size, $d' = 1.17$) and

Cohort D (effect size, $d' = .72$), and moderate for Cohort C (effect size, $d' = .44$). Although there was a decrease in the absolute number of words read by Cohorts 2C and 2D, compared to the mean scores of Cohort 2B, their scores remained at a level of achievement expected by mid-Grade 4 students.

Figure 8 shows a graph of the mean word recognition scores for both Grade 1 and Grade 2 children. Especially noteworthy is that the mean word recognition score of 82 from the Grade 1 children after three years of exposure to the instruction of in-serviced teachers (Cohort 1C) was above the mean recognition score of 79 from the Grade 2 children in the Baseline year (0 years of exposure).

Figure 8. Mean Number of Words Recognized
(Burns & Roe Informal Reading Inventory)



As shown in the figure, as well, there was progressive growth in the number of words read from the Baseline Group to the groups of Experimental children exposed to one or two years of instruction by in-serviced teachers (Cohorts 1A, 1B and 2A, 2B), after which students in both

grades reached a plateau at score levels more than a year above those expected for their grade, above Grade 2 for Grade 1 children and above Grade 3 for Grade 2 children.

For the Grade 1 word recognition scores, a trend analysis confirmed that there was a significant linear component, $F(1, 149) = 27.05, p < .001$. For the Grade 2 sample, trend analysis showed both a significant linear component, $F(1, 177) = 8.63, p < .004$, and a significant quadratic component, $F(1, 177) = 4.72, p < .03$.

A note of caution is required about the data from the Grade 1 children. Due to the substantial reduction in the number of Grade 1 children in the study after 1995, Cohorts 1B and 1C were only one-third of the original sample size. In 1996 there was only one Grade 1 teacher who had participated in the in-service PD. Therefore, there would be more variation in the data from Cohorts 1B and 1C, with more likelihood of spurious results.

Spelling

The overall comparison for the effect of exposure to the instruction of in-serviced teachers showed significant improvement in the mean spelling scores of both Grade 1 and Grade 2 students who were consistently exposed to the instruction of in-serviced teachers compared to the mean spelling scores of Grade 1 and Grade 2 children with no exposure to the instruction of in-serviced teachers, $F(1, 133) = 20.46, p < .001$ and $F(1, 165) = 19.23, p < .001$, respectively.

Grade 1. As indicated in Table 32 the mean spelling scores of Cohort 1A who had one year of exposure to the instruction of an in-serviced teacher showed significant gains compared to the children with no exposure to the instruction of in-serviced teachers, $t(143) = 3.13, p < .002$ (effect size, $d' = .62$). In addition, two and three years of consistent exposure to the instruction

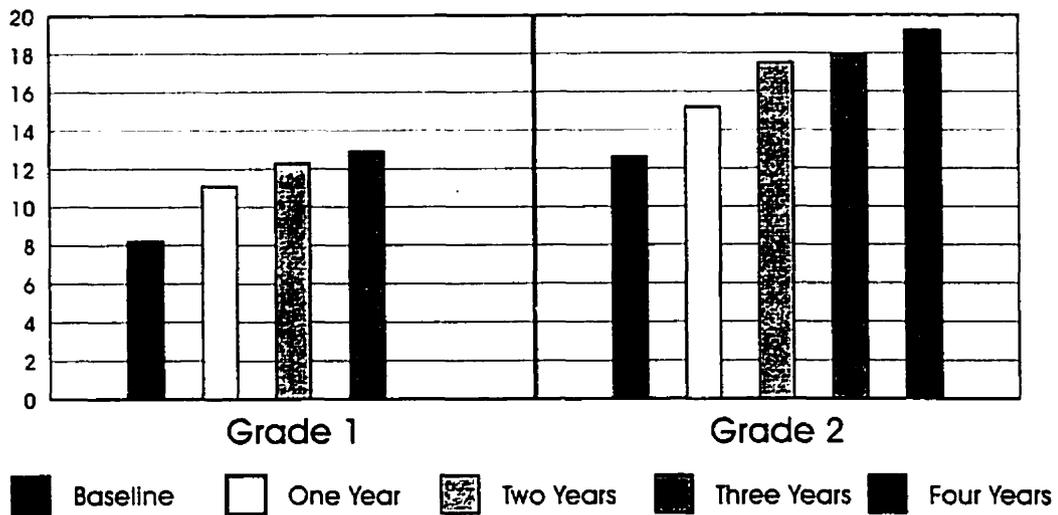
of teachers who had participated in the in-service PD (Cohorts 1B and 1C) significantly improved the spelling skills of Grade 1 students, $t(143) = 3.11, p < .002$ and $t(143) = 3.80, p < .001$, respectively (effect sizes, $d' = .84$ and $d' = .90$).

Grade 2. Similarly the spelling scores of the Grade 2 children showed improvement with exposure to the instruction of teachers who had participated in the in-service PD. After one year of exposure to a teacher who had participated in the in-service, Cohort 2A's increase in the mean number of correctly spelled words approached significance, $t(176) = 2.38, p < .02$ (effect size, $d' = .45$). Moreover, relative to the Baseline Group, Cohorts 2B, 2C and 2D showed significant gains in their mean spelling scores after two, three and four years of exposure to the instruction of in-serviced teachers, when compared to the Baseline Group who had no exposure to the instruction of in-serviced teachers, $t(176) = 4.31, p < .001$, $t(176) = 3.10, p < .002$, and $t(176) = 3.51, p < .001$, respectively (effect sizes, $d' = .90$, $d' = .50$ and $d' = .97$).

Figure 9 shows that there was a progressive improvement in the mean spelling scores of both Grade 1 and Grade 2 children from 0 years of exposure to 3 and 4 years of exposure to the instruction of teachers who had participated in the in-service PD. As can be seen in Figure 9, the mean spelling score of students in Grade 1 who were in the classrooms of in-serviced teachers for three years (Cohort 1C) was higher than the mean spelling score of the Grade 2 students in the Baseline Group.

Significant linear trends for both the Grade 1 and Grade 2 students indicated a progressive, consistent improvement in spelling scores with more exposure to the instruction of teachers who participated in the in-service PD, $F(1, 143) = 19.26, p < .001$ and $F(1, 176) = 25.82, p < .001$, respectively.

Figure 9. Mean Spelling Scores
(WRAT Raw Scores)



Thus, the spelling skills of children in Grades 1 and 2 showed cumulative benefits as a result of more years of exposure to the instruction of teachers who had participated in the in-service PD.

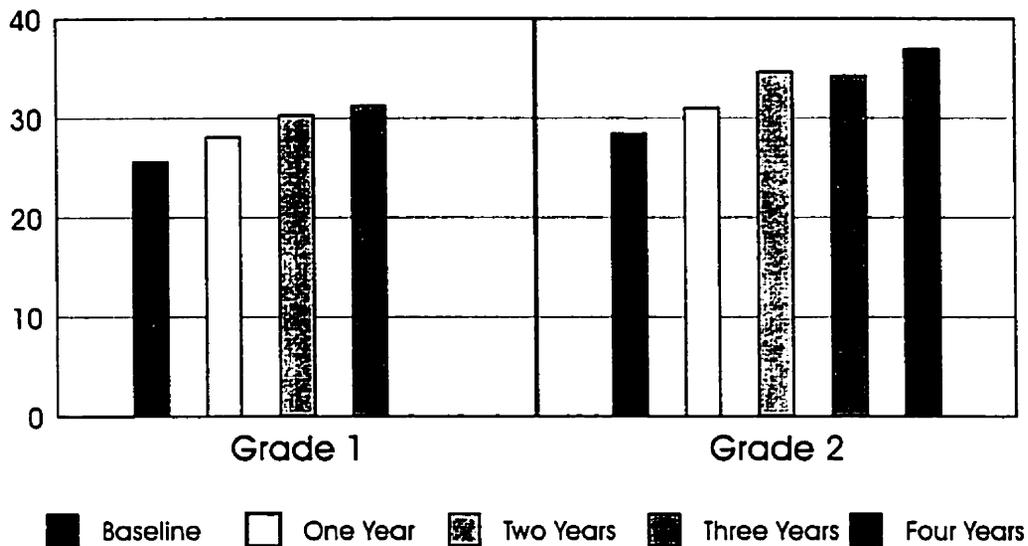
Reading Comprehension

The overall comparisons (1994 vs. 1995+1996+1997 for the Grade 1s and 1994 vs. 1995+1996+1997+1998 for the Grade 2s) showed significant improvement in the reading comprehension scores of Grade 2 students who were exposed to the instruction of in-serviced teachers compared to that of children who had no exposure to the instruction of in-serviced teachers, $F(1,165) = 4.32, p < .05$. There was no significant overall improvement in the reading comprehension scores of Grade 1 students.

Grade 1. Although, there were no significant gains in the reading comprehension scores of Grade 1 students in the year-by-year comparisons, there were small improvements in the scores. As seen in Table 32 the mean reading comprehension scores showed some consistent

improvement from 25 correctly answered questions for the Baseline Group to 31 correctly answered questions by Cohort 1C, after three years of exposure to the instruction of in-serviced teachers.

Figure 10. Mean Reading Comprehension
(Gates-MacGinitie Raw Scores)



Grade 2. In contrast, after 2 and 4 years of exposure to the instruction of in-serviced teachers, the mean reading comprehension scores of Cohorts 2B and 2D showed significant gains compared to the children with no exposure to the instruction of in-serviced teachers, $t(186) = 3.09, p < .002$ and $t(186) = 2.96, p < .003$, respectively (effect sizes, $d' = .64$ and $d' = 1.10$).

Figure 10 shows that there was a progressive improvement in the mean reading comprehension scores of both Grade 1 and Grade 2 children in the classrooms of in-service teachers from 0 years of exposure to 3 and 4 years of exposure to the instruction of teachers who had participated in the in-service PD. The linear trend analysis for the Grade 1 students' reading comprehension scores produced an approaching-significance linear component, $F(1,141) = 3.83$,

$p < .05$. Although there were only small year-by-year improvements, the children appeared to show some cumulative benefits, but these were not as striking as those for word recognition and spelling. This finding seems to reflect the fact that in Grade 1 teaching reading comprehension skills is not a priority. Until children are able to read with some fluency, based on good word recognition skills, they will not have the cognitive resources to apply to reading comprehension strategies (Stanovich, 1986, 2000).

For the Grade 2 students the trend analysis results were very strong and clear. A highly significant linear component indicated a progressive, consistent improvement in reading comprehension scores with more exposure to the instruction of teachers who participated in the in-service PD, $F(1, 186) = 14.52, p < .001$.

Conclusions Concerning Tracking Student Change Across Years with In-serviced Teachers

The word recognition and spelling scores of Grade 1 and 2 children and the reading comprehension scores of Grade 2 children showed significant progressive improvement as these students had the opportunity for increased exposure to the instruction of teachers who had participated in the in-service PD.

It should be noted that the decrease in scores in 1996 and 1997 may be due to the substantial drop in the numbers of children in the samples (Cohorts 1B, 2C and 2D). In 1996, during the third year of the study, there was only one Grade 1 teacher who had participated in the in-service. Therefore, there was only one class of Grade 1 students who could be included in the subsequent Grade 2 cohort data. By fourth year there was also the natural attrition of longitudinal subjects. Therefore, due to the smaller sample sizes the results of the fourth and fifth years of the study may be less stable.

These analyses confirmed a progressive improvement in the literacy scores of Grade 1 and Grade 2 children with more years of exposure to the instruction of teachers who participated in the in-service PD. Therefore, the results support the value of this type of in-service PD for Junior and Senior Kindergarten teachers. Grade 1 and Grade 2 children who had instruction in word recognition and spelling skills from in-serviced Junior and Senior Kindergarten teachers performed better on these assessment measures than children who had no exposure or who were exposed to only one year of instruction by in-serviced Grade 1 and Grade 2 teachers. This research appears to support the learning benefits for children who have consistent exposure to high quality teaching during their school years.

CHAPTER 7

Children “At Risk” of Literacy Failure

This chapter examines the influence of the in-service PD on the effectiveness of teachers’ literacy instruction for those students who are most at risk of literacy failure. This was done by examining the proportion of students scoring in the lowest percentile ranges on the literacy measures across the years of the study from Baseline Year to Follow-up Year 2. Thus, it was determined whether the professional development of teachers was effective in reducing the rate of literacy failure in students.

A potential risk factor was the home language of students. This school was in an area of a high proportion of students who spoke English as their second language (ESL or L2). Therefore, the impact of ESL on the acquisition of literacy skills was felt to be an important factor to explore. To do this, the literacy acquisition of children who spoke English as their first language (L1) and children who spoke English as their second language (L2) assessed at the end of the Baseline Year and the Treatment years (combining the Intervention Year and the two Follow-up Years) were compared. However, the sample of students with English as L2 was relatively small and the designation as L2 was not entirely consistent. Therefore, any comparisons between the literacy outcomes of L1 and L2 children are very tentative. In Appendix B the discussion and results of this exploratory analysis comparing the effect of PD on the acquisition of literacy skills between L1 and L2 children is provided.

Teachers’ Knowledge and Classroom Practice

The Research Question examining the effect of professional development on the acquisition of literacy skills by the “at risk” population can be stated as follows:

3. Can teachers’ participation in professional development effectively reduce the degree of

literacy failure among their students?

Even with the best teaching, some children will continue to be find literacy learning challenging. The literacy scores of such “at risk” children will remain relatively low compared with their classmates. However, if teacher knowledge and instructional skills are improved, the question is: Will there be fewer children scoring in the lowest absolute score ranges on literacy measures? If there are fewer children whose absolute score levels fall into very low score ranges after their teachers have participated in in-service PD, then it suggests that improved instructional practice can directly affect the number of children who are “at risk” of literacy failure.

Participant Characteristics

The Experimental groups were the Grade 1 and Grade 2 children who were in the classes of teachers who had participated in the professional development. These are the same children who were in the analyses examining the impact of the professional development with respect to growth in teacher effectiveness. The Baseline group were the Grade 1 and Grade 2 children who were in the school during the Baseline Year, before the initiation of the in-service PD (tested in June 1994). The Intervention Year group were children in the classrooms of teachers who had participated in the in-service PD from September 1994 to June 1995. These children had been tested on the three dependent literacy measures in June 1995. The Follow-up Year 1 and Follow-up Year 2 groups were children in the classrooms of teachers who had participated in the professional development one and two years previously. These children were assessed on the literacy measures in June 1996 and June 1997, respectively. Table 21 (abbrev.) is abbreviated from the full version of Table 21 presented earlier and it is presented here for the reader’s convenience.

Table 21 (abbrev.)

Summary Table of Participant Characteristics

	Grade 1			
	<u>Baseline Year</u> June 1994	<u>Intervention Year</u> June 1995	<u>Follow-up 1 Year</u> June 1996	<u>Follow-up 2 Year</u> June 1997
	n = 83	n = 52	n = 25	n = 33
Age (mon.)	<u>M</u> 83.1 (3.3)	<u>M</u> 83.3 (4.3)	<u>M</u> 82.6 (4.9)	<u>M</u> 83.1 (3.0)

	Grade 2			
	<u>Baseline Year</u> June 1994	<u>Intervention Year</u> June 1995	<u>Follow-up 1 Year</u> June 1996	<u>Follow-up 2 Year</u> June 1997
	n = 68	n = 58	n = 79	n = 54
Age (mon.)	<u>M</u> 95.2 (3.7)	<u>M</u> 94.5 (3.4)	<u>M</u> 96.7 (5.3)	<u>M</u> 94.2 (4.3)

Notes: - All scores are raw scores.

- Over the four years of study, some of the in-serviced teachers have retired, moved to other schools, or to the junior grades so the pool of children in the classrooms of in-serviced teachers has been dropping yearly in the grade one classrooms.

- Due to the relatively high mobility of children in the school these analyses include both children who entered or left the school over the four years of testing as well as those who remained in the school for the full period.

Analyses

In order to evaluate the effect of professional development on reducing the literacy failure of students, descriptive statistics were used. The proportion of word recognition, spelling and reading comprehension scores in each quartile of the score ranges was calculated for all of the

Grade 1 and Grade 2 children across the four years of the study. First, the scores on each of the three tests (separately) for all four years (combined) were sorted from highest to lowest, and the range of scores in each quartile was calculated across the years within each grade separately. Then, the percentage of scores within each quartile was calculated year by year. The lowest scores would be within the 1st Quartile and the highest scores would be within the 4th Quartile. The number of children scoring in the lowest range, or the 1st Quartile, would probably reflect the approximate number of children in the sample who would need Special Education during each year of the study.

Results and Discussion

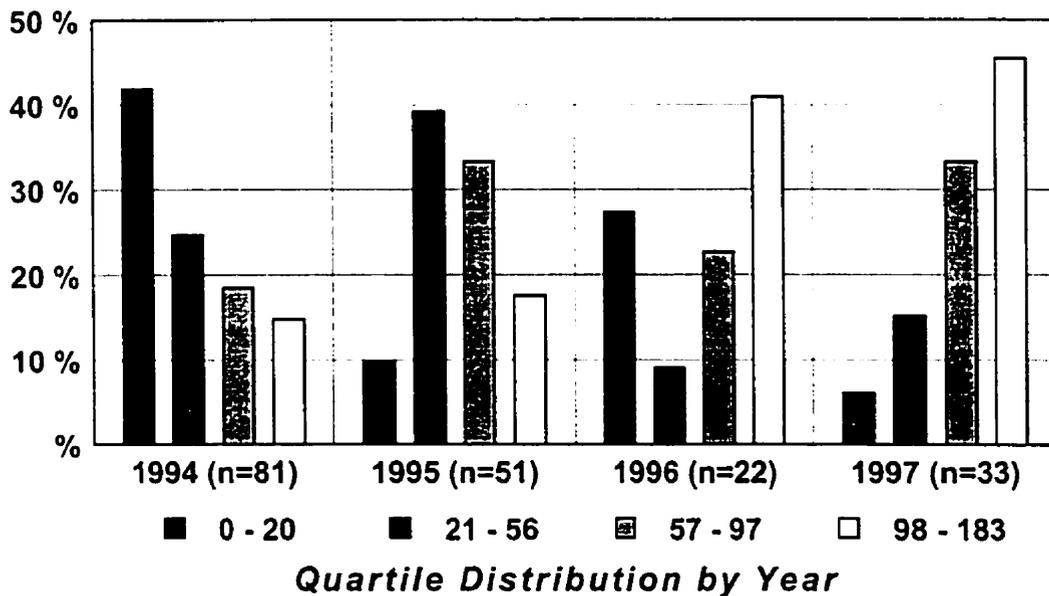
Tables 34 and 35 present the proportion of scores as percentages in each quartile for the three literacy measures for each grade across the four years of the study. Descriptive statistics and chi-square analyses (comparing each Treatment Year with the Baseline Year) were used to determine whether in-service PD reduced the percentages of children who would be considered “at risk” of literacy failure.

Word Recognition

Grade 1: As shown in Figure 11, there was a dramatic decrease in the percentage of Grade 1 children’s scores in the lowest quartile from 41.9 percent in the Baseline Year to 6.0 percent in Follow-up Year 2, $\chi^2 (1, N = 114) = 11.26, p < .005$. There was an immediate effect of the in-service PD on the percentage of Grade 1 children’s scores in the lowest quartile, with a decrease in the Intervention Year to only 9.8 percent of the total word recognition scores in the 1st Quartile $\chi^2 (1, N = 132) = 14.04, p < .005$. Although there was a subsequent increase in the number of scores in the 1st Quartile to 27.3 percent in Follow-up Year 1, this percent was still only half the number of scores in the 1st quartile during the Baseline Year $\chi^2 (1, N = 103) =$

9.37, $p < .005$, before the initiation of the in-service PD. Due to the much smaller samples of grade one children in 1996 and 1997, the use of percentage scores can result in a somewhat unreliable picture, since small absolute differences can produce relatively large percentage differences. Thus, the rises and falls in the percentages should not be over interpreted. The percentage scores in 1994 and 1995 for all the Grade 1 children and for all four years for the Grade 2 children should be considered more reliable because of the larger sample sizes.

Figure 11. Word Recognition (Burns & Roe Raw Score)
Grade One



Examining the absolute numbers of Grade 1 children whose scores were in the lowest quartile (to avoid the distorting effect of percentages when n 's are low), before the in-service PD in the Baseline year, there were 34 children whose word recognition scores were very poor. However, after the Grade 1 teachers participated in the PD, there were only 5 children in the Intervention Year in 1995, 6 children in Follow-up Year 1, and 2 children in Follow-up Year 2 who were in the lowest range of scores.

Table 34 Percentage of Grade 1 Students' Word Recognition, Spelling, and Reading Comprehension Scores in each Quartile

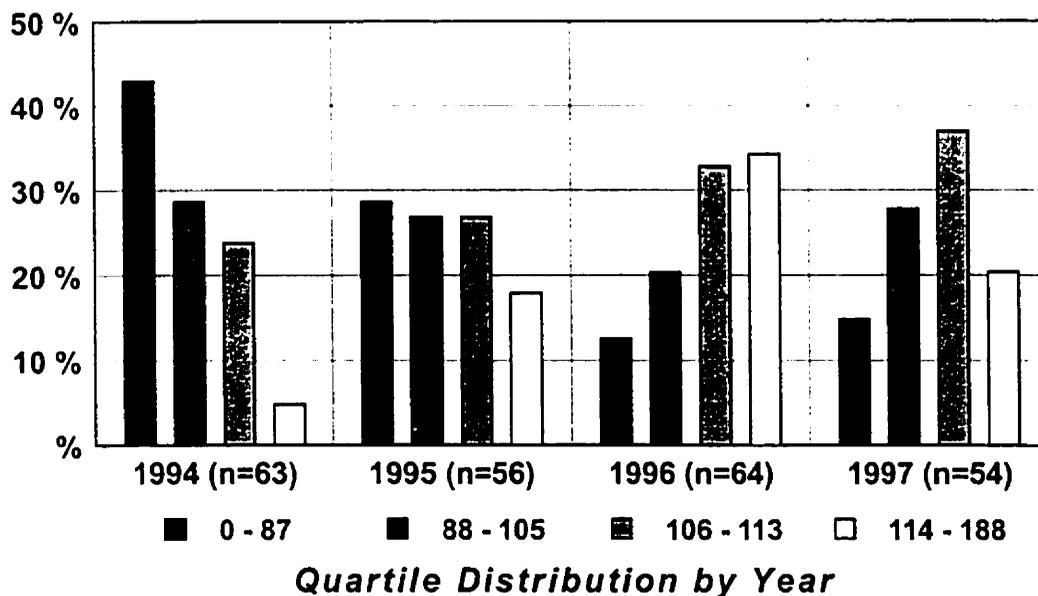
<u>Year</u>	<u>Word Recognition Quartiles</u>			
	1 st Quartile (0 - 20)	2 nd Quartile (21 - 56)	3 rd Quartile (57 - 97)	4 th Quartile (98 - 183)
Baseline Year (June 1994)	41.9 % (34/81)	24.7 % (20/81)	18.5 % (15/81)	14.8 % (12/81)
Intervention Year (June 1995)	9.8 % (5/51)	39.2 % (20/51)	33.3 % (17/51)	17.6 % (9/51)
Follow-up Year 1 (June 1996)	27.3 % (6/22)	9.0% (2/22)	22.7 % (5/22)	40.9 % (9/22)
Follow-up Year 2 (June 1997)	6.0 % (2/33)	15.1 % (5/33)	33.3 % (11/33)	45.5% (15/33)

<u>Year</u>	<u>Spelling Quartiles</u>			
	1 st Quartile (0 - 6)	2 nd Quartile (7 - 10)	3 rd Quartile (11 - 14)	4 th Quartile (15 - 27)
Baseline Year (June 1994)	41.8 % (31/74)	32.4 % (24/74)	14.8 % (11/74)	10.8 % (8/74)
Intervention Year (June 1995)	13.7 % (7/51)	29.4 % (15/51)	35.3 % (18/51)	21.6 % (11/51)
Follow-up Year 1 (June 1996)	33.3 % (8/24)	20.8% (5/24)	20.8 % (5/24)	25.0 % (6/24)
Follow-up Year 2 (June 1997)	6.0 % (2/33)	27.3 % (9/33)	36.4 % (12/33)	30.3% (10/33)

<u>Year</u>	<u>Reading Comprehension Quartiles</u>			
	1 st Quartile (0 - 16)	2 nd Quartile (17 - 27)	3 rd Quartile (28 - 40)	4 th Quartile (41 - 46)
Baseline Year (June 1994)	36.5 % (27/74)	22.9 % (17/74)	21.6 % (16/74)	18.9 % (14/74)
Intervention Year (June 1995)	22.4 % (11/49)	28.5 % (14/49)	20.4 % (10/49)	28.6 % (14/49)
Follow-up Year 1 (June 1996)	21.7 % (5/23)	30.4% (7/23)	34.8 % (8/23)	13.0 % (3/23)
Follow-up Year 2 (June 1997)	15.1 % (5/33)	24.2 % (8/33)	33.3 % (11/33)	27.3% (9/33)

Grade 2. For the Grade 2 students' word recognition scores, Figure 12 shows a remarkable progressive decrease in the percentage of total scores in the lowest quartile, from 42.9 percent in the Baseline Year to 14.8 percent in Follow-up Year 2. In absolute numbers there were 27 Grade 2 children whose word recognition was exceedingly poor in the Baseline Year. There were only 16 children scoring poorly in the Intervention Year in 1995, $\chi^2(1, N = 119) = 9.17$, $p < .005$, 8 children scoring poorly in Follow-up Year 1, $\chi^2(1, N = 127) = 12.03$, $p < .005$, and 8 children scoring poorly in Follow-up Year 2, $\chi^2(1, N = 117) = 10.27$, $p < .005$.

**Figure 12. Word Recognition (Burns & Roe Raw Score)
*Grade Two***



Clearly, the need for Special Education to improve children's word recognition skills were considerably reduced as a result of classroom teachers' participation in professional development.

Table 35 Percentage of Grade 2 Students' Word Recognition, Spelling, and Reading Comprehension Scores in each Quartile

<u>Year</u>	<u>Word Recognition Quartiles</u>			
	1 st Quartile (0 - 87)	2 nd Quartile (88 - 105)	3 rd Quartile (106 - 113)	4 th Quartile (114 - 188)
Baseline Year (June 1994)	42.9 % (27/63)	28.6 % (18/63)	23.8 % (15/63)	4.8 % (3/63)
Intervention Year (June 1995)	28.6 % (16/56)	26.8 % (15/56)	26.8 % (15/56)	17.9 % (10/56)
Follow-up Year 1 (June 1996)	12.5 % (8/64)	20.3% (13/64)	32.8 % (21/64)	34.3 % (22/64)
Follow-up Year 2 (June 1997)	14.8 % (8/54)	27.8 % (15/54)	37.0 % (20/54)	20.4% (11/54)

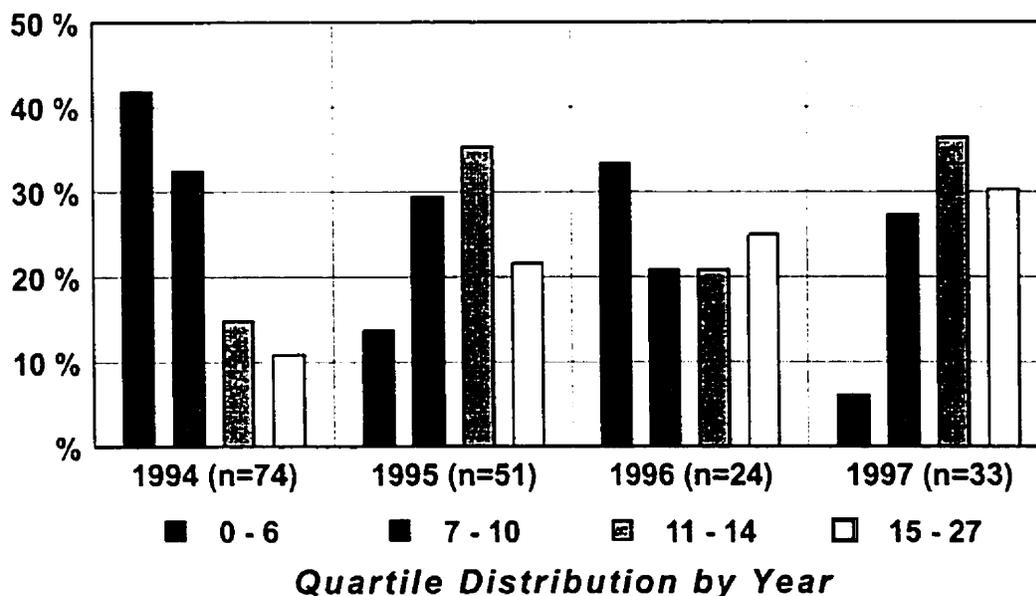
<u>Year</u>	<u>Spelling Quartiles</u>			
	1 st Quartile (0 - 11)	2 nd Quartile (12 - 15)	3 rd Quartile (16 - 19)	4 th Quartile (20 - 35)
Baseline Year (June 1994)	48.3 % (29/60)	25.0 % (15/60)	13.3 % (8/60)	13.3 % (8/60)
Intervention Year (June 1995)	29.6 % (16/54)	25.9 % (14/54)	20.4 % (11/54)	24.0 % (13/54)
Follow-up Year 1 (June 1996)	16.7 % (12/72)	25.0% (18/72)	29.2 % (21/72)	29.2 % (21/72)
Follow-up Year 2 (June 1997)	16.7 % (9/54)	25.9 % (14/54)	24.1 % (13/54)	33.3% (18/54)

<u>Year</u>	<u>Reading Comprehension Quartiles</u>			
	1 st Quartile (8 - 24)	2 nd Quartile (25 - 33)	3 rd Quartile (34 - 40)	4 th Quartile (41 - 46)
Baseline Year (June 1994)	34.8 % (23/66)	25.7 % (17/66)	28.8 % (19/66)	10.6 % (7/66)
Intervention Year (June 1995)	30.9 % (17/55)	18.2 % (10/55)	23.6 % (13/55)	27.3 % (15/55)
Follow-up Year 1 (June 1996)	18.0 % (13/72)	31.9 % (23/72)	25.0 % (18/72)	25.0 % (18/72)
Follow-up Year 2 (June 1997)	18.5 % (10/54)	33.3 % (18/54)	22.2 % (12/54)	25.9% (14/54)

Spelling

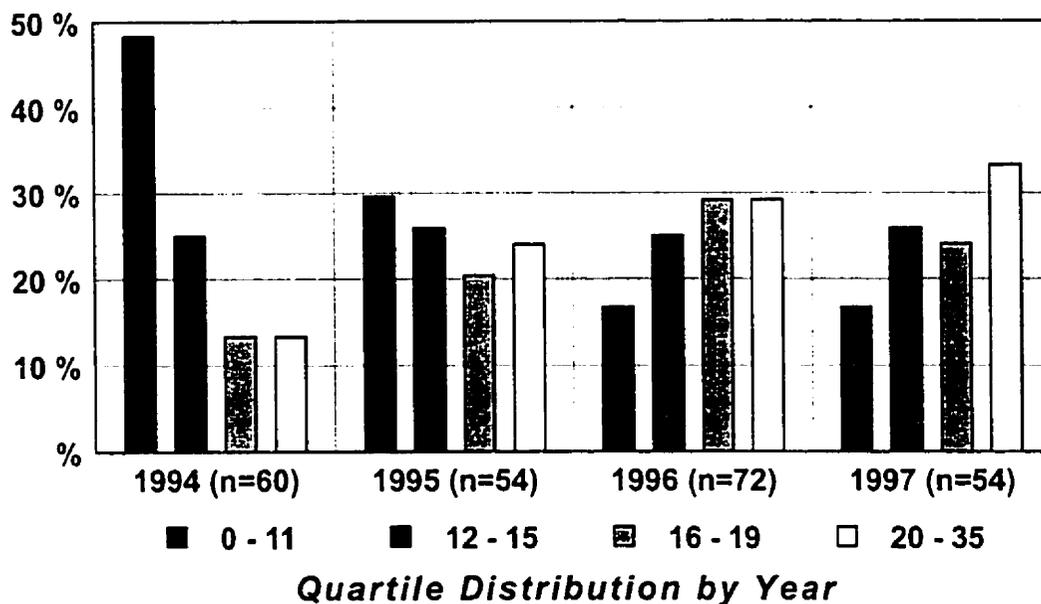
Grade 1. There was a decrease in the actual numbers of Grade 1 spelling scores within the first quartile during the in-service PD and two years after the PD. Again, however, due to the small sample sizes in the two Follow-up Years for the Grade 1s, there was a distorting effect of percentage scores that is reflected in the peaks and valleys in Figure 13. In absolute numbers, there were 31 children whose scores on the spelling measure were very poor in the Baseline Year, 7 children scoring poorly in the Intervention Year (1995), $\chi^2 (1, N = 125) = 11.04, p < .005$, 8 children scoring poorly in Follow-up Year 1, $\chi^2 (1, N = 98) = 9.11, p < .005$, and only 2 children in Follow-up Year 2, $\chi^2 (1, N = 107) = 13.18, p < .005$, who scored poorly on the spelling task. Thus, there appeared to be a decrease in the children with significant problems in spelling when their teachers had participated in the in-service PD, but the small sample sizes in the two Follow-up Years preclude drawing any firm conclusions.

**Figure 13. Spelling WRAT Raw Score
Grade One**



Grade 2. For the Grade 2 children, where the sample sizes were consistently high across the 4 years, a more compelling picture emerged. Figure 14 shows a decrease in the number of children who scored extremely poorly on the spelling measure, from 29 in the Baseline Year to only 9 children in Follow-up Year 2 (see Table 35). The chi squares, respectively, for the Intervention Year, and Follow-up Years 1 and 2 were, $\chi^2(1, N = 114) = 13.53$, $\chi^2(1, N = 132) = 15.07$, $\chi^2(1, N = 114) = 14.75$, all $ps < .005$.

**Figure 14. Spelling WRAT Raw Score
*Grade Two***



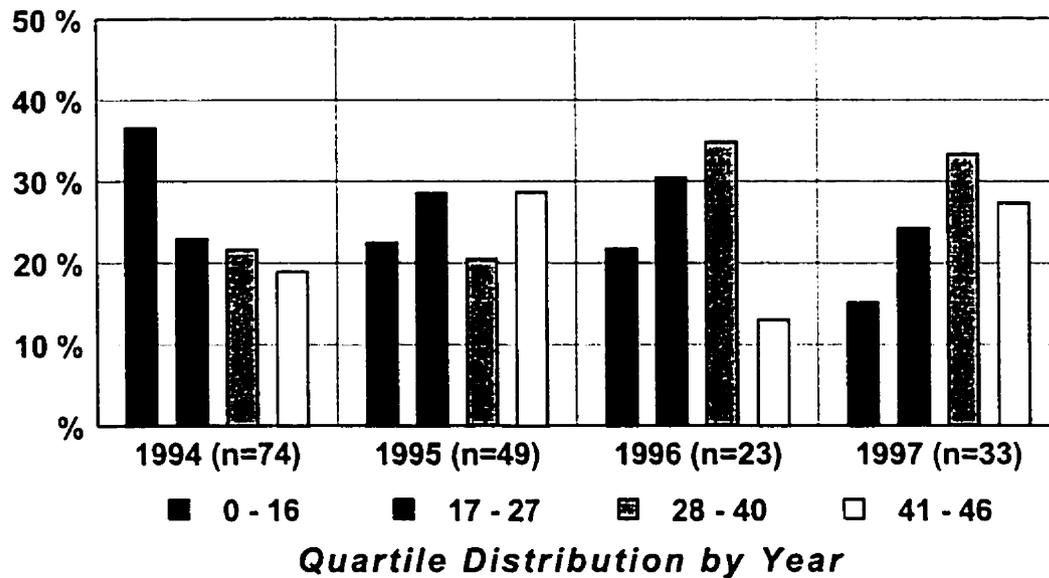
Together, these data seem to demonstrate a clear impact of professional development on the effectiveness of teachers' spelling instruction, as the number of children "at risk" for extremely poor spelling performance decreased over the four years of the study for both Grade 1 and Grade 2 children.

Reading Comprehension

Grade 1. Figure 15 shows more variability in the percentage of reading comprehension scores within each quartile across the years of this study. Again, the sample sizes of Grade 1 children in each year of this study was more variable, therefore, these data are not as reliable.

Nevertheless, there is a decrease in the absolute number of children in the lowest score range, from 27 in the Baseline Year, to 11 children scoring poorly during the Intervention Year, $\chi^2 (1, N = 123) = 4.03, p < .05$, to only 5 children scoring in the lowest range during Follow-up Year 1, $\chi^2 (1, N = 97) = 3.99, p < .05$, and Follow-up Year 2, $\chi^2 (1, N = 107) = 5.18, p < .05$. Caution is again warranted in drawing firm conclusions from the small samples in the two Follow-up Years.

**Figure 15. Reading Comprehension
(Gates-MacGinitie Raw Score)
*Grade One***



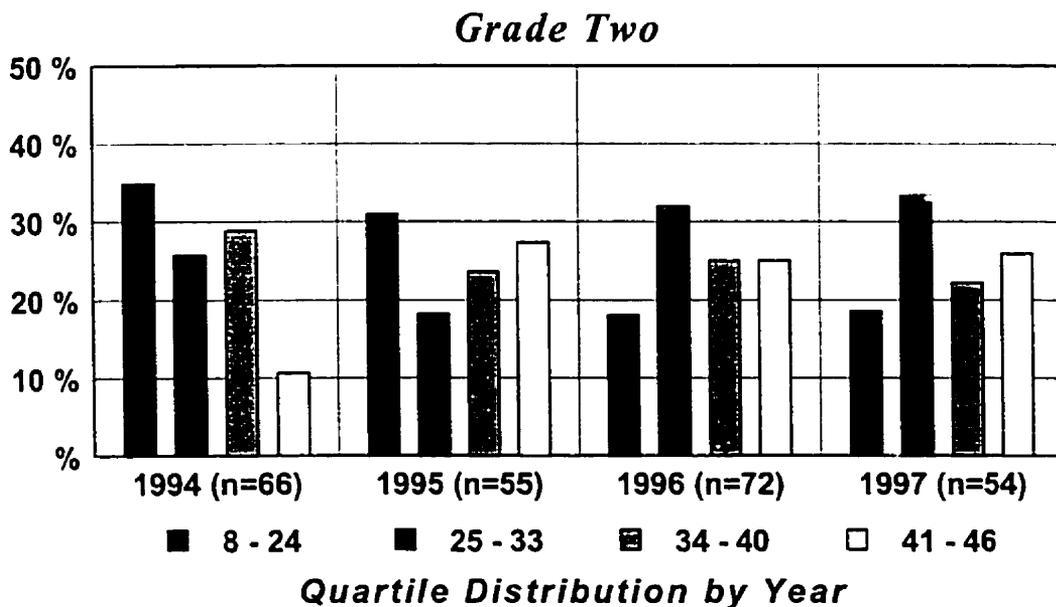
Grade 2. The Grade 2 findings were somewhat weaker. There was, however, a decrease in the proportion of Grade 2 children whose scores were in the lowest range (see Figure 16).

Although the effect was not as strong as for the word recognition and spelling measures, in absolute numbers there were 23 children from the three classes of Grade 2 children whose

scores were so low that they would probably warrant assistance during the Baseline Year. However, in the subsequent years, after the Grade 2 teachers had participated in the in-service PD, there were fewer children in Grade 2 who would likely require Special Education programming; the absolute scores for the Intervention Year, Follow-up Year 1 and Follow-up Year 2 being 17, 13 and 10, respectively. The Chi-square findings for the Intervention Year, and Follow-up Years 1 and 2 were $\chi^2(1, N = 121) = 3.33, .10 > p > .05$, $\chi^2(1, N = 138) = 3.95, p < .05$, $\chi^2(1, N = 120) = 3.46, .10 > p > .05$. Thus despite the overall downward trend in the numbers the significance levels were modest.

Although the impact of teachers' participation in the PD on the reading comprehension skills of their Grade 1 and Grade 2 students was not as great as the impact on their word recognition and spelling skills, the need for remedial assistance to improve reading comprehension skills was reduced somewhat.

**Figure 16. Reading Comprehension
(Gates-MacGinitie Raw Score)**



Summary of the Impact of Teachers' Participation in PD on Students' Rate of Literacy Failure

Based on the descriptive statistics presented in this section, it appears that participation of classroom teachers in professional development may be a contributing factor to the reduction in the numbers of their students who require Special Education programming to improve their word recognition, spelling and, to some extent, reading comprehension skills. Although there is some evidence to support this conclusion in the Grade 1 data, the drastically reduced sample sizes in 1996 and 1997, because of the loss of in-serviced teachers limit the reliability of results at that grade level. The evidence from the Grade 2 classrooms, however, where the sample size remained larger over the four years of the study, is very compelling. On all the literacy measures, the absolute numbers of children whose performance was in the lowest quartile diminished greatly over the years of the study. The overall need for low-ratio Special Education programming for "at risk" students was greatly reduced in this sample.

Even with the best classroom teaching, not all children acquire literacy skills with ease. Factors outside the control of the classroom teacher, like within-child factors such as learning disabilities and within-home factors such as poverty, contribute to children's difficulties in acquiring literacy skills (Jaeger, 1992). However, based on these findings, whether or not students are "at risk" of literacy failure clearly depends, in part, on the expertise of their classroom teachers. Better teaching reduces the number of children "at risk" for failing to learn literacy skills. Through professional development teachers can acquire knowledge and instructional strategies to improve their teaching practices.

CHAPTER 8

Summary and General Discussion

The purpose of this study was to evaluate a professional development system – *The Balanced and Flexible Literacy Diet* – which combines the key components for effective literacy instruction with the key strategies for promoting transfer of new skills to the classroom. This chapter will begin by providing a brief overview of the major findings of the research. It will then present a discussion of the contributions and limitations of the study, will make recommendations for future research, and will draw some final conclusions.

Overview of the Major Findings

This evaluation of the impact of in-service PD on student literacy outcomes was undertaken in three main parts: (1) the first involved tracking teachers who participated in the PD to determine whether they became better teachers of beginning reading and writing after their participation; (2) the second involved tracking students who were consistently in the classrooms of teachers who participated in the PD to determine whether they showed cumulative benefits in the form of better literacy outcomes after being in the classrooms of in-serviced teachers; and (3) the third involved an examination of the literacy outcomes of “at risk” students to determine whether classroom teachers’ participation in in-service PD can reduce the proportion of students in their classrooms requiring special education in reading and writing. The major findings relating to each of these three types of analyses were as follows:

First, based on analyses that tracked in-serviced teachers, the findings from this longitudinal study indicate that professional development was effective in improving the instructional skills of participating teachers, with the associated raise in the literacy achievement levels for students in their classrooms. This improvement in literacy achievement levels of Grade 1 and 2 children

in the classrooms of in-serviced teachers was progressive over the years, apparently reflecting a consolidation of knowledge and an improvement in instructional practices. However, not all teachers made great gains. A teacher-by-teacher examination of the literacy outcomes of students in the classrooms of three teachers who remained in the same grade over the 4 years of the study (Baseline Year, Intervention Year, Follow-up Year 1 and Follow-up Year 2), suggested that there was considerable variability in the extent to which teachers profited from the in-service PD; their students' gains ranging from exceptional to marginal.

Second, based on the analyses tracking children, there were cumulative benefits to the literacy skills of Grades 1 and 2 students. Children with more years of exposure to the instruction of teachers who participated in the in-service PD were more successful in reading and spelling than were those with less exposure. Thus, students benefited from having on-going instruction by teachers who are knowledgeable and skilled in literacy education.

Third, the results from this research indicate that participation of teachers in PD can contribute to a reduction in the numbers of children who are "at risk" of failing to attain age-appropriate levels of achievement in reading and spelling. Better instructional practices apparently led to improved literacy outcomes for children with reading and spelling difficulties. Thus, teachers' participation in effective in-service PD appeared to reduce the proportion of their students needing special education.

Overall, the findings from this research show that providing teachers with an opportunity to extend their knowledge and hone their teaching skills in the area of literacy education can have significant long-term benefits. Through participation in professional development teachers can

significantly refine their teaching skills, and there is an associated rise in the literacy achievement of children in their classrooms.

Contributions of Current Research

As indicated in the literature review, there has been very little empirical research on the impact of teacher education – both pre-service and in-service – on teacher effectiveness. With respect to pre-service programs, there have been virtually no studies examining the impact of the pre-service training in literacy education on the teaching success of new teachers, as reflected in their students' literacy achievement. It is, however, increasingly argued that there is a need to reform pre-service programs in the area of literacy education (American Federation of Teachers, 1999; Lyon et al., 1989; MacGinitie, 1991; Moats, 1995; National Reading Panel, 2000). In view of the fact that the practicing teachers in the current research were able to make significant improvements in the literacy outcomes of their students by applying the knowledge and skills that they acquired through in-service PD, it would be interesting to explore if new teachers would benefit from a program that included the teaching of similar knowledge and skills.

The literature on in-service teacher education is also very sparse. There is some research and discussion concerning general principles of professional development to promote change in classrooms (e.g., Fullan, 2000; Guskey, 1986, 1995; Joyce & Showers, 1988; Sparks, 1988; Veenman et al., 1994). Within this literature on in-service PD, only a very small proportion deals with the specific subject area of literacy education in the general classroom. Six such studies, concerned with the role of PD in successful literacy education, were reviewed in the introduction (i.e., Baker, 1977; Conley, 1983; Cunningham et al., 1998; Miller & Ellsworth, 1985; Richardson, 1994; Stallings & Krasavage, 1986). In the review of these studies, enormous variability was found in the methodologies employed, the measures used to assess the

effectiveness of the PD, and in the focus of change in the research –the students, the teachers or both. In view of the dearth of research and the variability amongst existing studies, there is a clear need for large-scale, systematic, longitudinal research that assesses the impact of professional development on teachers' knowledge, attitudes and practices, and on students' literacy achievements. The present research was part of one such study.

This research was conducted as part of an attempt to provide converging evidence concerning the influence of teacher education on student literacy success (see Figure 1). The larger context of the research has involved studies examining the impact of in-service PD (1) on promoting change in teachers' knowledge, attitudes and practices (Barnes-Haig & Willows, 1999; Siegelman & Willows, 2000); (2) on how teachers and children spend their time during literacy instruction (Domi & Willows, 2000; Sumbler & Willows, 1998 and; (3) on student literacy outcomes resulting from the use of specific theory-based teaching strategies (Kwan & Willows, 1998; Morgan & Willows, 1997; Stornelli & Willows, 1998; Sumbler & Willows, 1998; Sumbler, 1999).

The results of these other studies showed that (1) knowledge, attitudes and practice of in-serviced teachers in literacy education were clearly different from those of control teachers who had not participated in the in-service; and that (2) classroom teaching practices during literacy time were clearly different after teachers had participated in the in-service PD. The time sampling in classrooms demonstrated changes in classroom instruction that were directly linked to in-service training, with control/comparison classes that had no exposure to the in-service.

From this present study alone, causal inferences cannot be drawn concerning the relation between teachers' professional development in literacy education and student literacy outcomes.

However, this study was conducted in the context of this larger research project that involved the impact of professional development on the attitudes of educators, the impact of professional development on classroom practice, and the impact of professional development on student outcomes. In the context of this larger picture it is reasonable to draw at least tentative causal inferences.

The focus of this thesis was only on “the bottom line,” student literacy outcomes. It has confirmed and extended what is known about the impact that in-service professional development can have on teachers’ competence and on students’ success. Some previous research has pointed to higher achievement of students exposed to the instruction of teachers who have participated in professional development incorporating the features for the successful transfer of skills into classroom practice (e.g., Guskey, 1986; Showers & Joyce, 1988). A few other studies, specifically dealing with literacy, have shown a positive impact of professional development on improving children’s literacy achievement (i.e., Baker, 1977; Conley, 1983; Cunningham et al., 1998; Miller & Ellsworth, 1985; Richardson, 1994; Stallings & Krasavage, 1986). As discussed earlier, however, these studies varied greatly on a number of dimensions. The professional development system evaluated here – *The Balanced and Flexible Literacy Diet* – attempted to combine what has been learned from the literature concerning key components associated with successful literacy instruction with key features associated with successful transfer of new skills to the classroom. Thus the present study extended previous research on the effect of in-service teacher education on students’ literacy success by employing a more comprehensive PD system, by including a wider range of literacy measures, by using a longitudinal research design and by focusing on whole classrooms of students as well as “at risk” students taught by regular classroom teachers. The outcomes of the research are very promising. They indicate that in-service PD can have a substantial effect on classroom success,

and that this effect benefits both the teachers – in terms of their effectiveness over the short and long terms – and the children – who demonstrate not only short-term benefits to their literacy achievement but also cumulative benefits if they continue to be taught by in-serviced teachers. This research also found that in-service PD for classroom teachers benefits children thought to be “at risk” of literacy failure. Many previous studies have focused on improving instruction only for at-risk children (e.g., Clay, 1986a, 1986b; Hiebert et al., 1992; Taylor et al., 1992).

Limitations and Future Directions

Despite the promising results of the present research, it should only be considered as one small step toward a better understanding of the impact of professional development on teacher effectiveness and student success because of a number of limitations.

One limitation is that this study did not involve the use of a true experimental design (i.e., experimental and control groups with random assignment of teachers and students to conditions). Rather it involved a quasi experimental design using a pre-treatment baseline as a control.

Another limitation of this study was the participation in the PD of elementary grade staff in only one school. In order to generalize the findings to a broad range of schools, it would be important to replicate the findings from the present study in a number of schools, involving more classrooms and more teachers.

Other limitations reflect the realities of collecting data in a high mobility school over a number of years. This resulted in diminishing sample sizes over the years of the study. One consequence of the longitudinal nature of the study and of the mobility of the teacher and

student populations was that sample sizes became substantially smaller over the years of the study.

In future research it would be beneficial to calculate the sample size required at the beginning of the study in order to have a sufficiently large sample at the end of the longitudinal study.

Furthermore, information from the schools concerning the mobility rates of students would be helpful in calculating the required sample size. In a lower SES community there is even more likelihood of mobility, especially among lower achieving children with fewer resources available. Children who move are different in non-random ways from children who do not move. Therefore, there could be a confound of SES affecting the results. In future research, SES data should be collected to avoid the possibility of a confound such that the children who remained in the sample over the full term of the study might have come from more affluent, better educated families. In the present research, such a potential confound could have affected the results of analyses involving tracking students; they would not, however, have affected analyses tracking teachers, since these analyses included students who moved as well as those who stayed in the school. A second potential problem associated with mobility involved the teachers. As teachers moved grades or out of the school, students were placed in classrooms of teachers who had not participated in the in-service, so these students were lost to the study. The researchers had no knowledge of or influence on how students were placed in the classes of in-serviced or non in-serviced teachers each year. Perhaps placements were done in such a way as to create a confound with instructional quality, either enhancing or detracting from treatment effects. Future studies should monitor placement policies to insure that these could not affect research outcomes.

Another limitation is that the analyses undertaken to track teachers and to track children were not independent. That is the samples of children overlapped in the two sets of analyses. This was done to gain as much insight as possible from the data set, but ideally the analyses tracking teachers should have been conducted with an entirely independent data set. Furthermore, in the analyses tracking teachers, one focus was on individual differences. Unfortunately, only three teachers could be tracked in the same grade over four years. Thus, the findings are only preliminary. However, the analysis of particular teachers suggests that there is considerable variability in the impact of PD on the skills of individual teachers. It is clearly important that future research examine individual teachers, as well as groups of teachers.

The statistics used in the analyses of “at risk” students should be interpreted cautiously since the proportion of students above or below an extreme cut-off score (near the end of the distribution) is a much less stable statistic than the mean score (Yen, 1997).

A further limitation was the lack of any pre measures of student cognitive abilities. Since sample sizes were relatively small in the follow-up years of the study, the achievement scores of a few very weak or very strong students might have had an influence on the outcome of the study. In future research complete cognitive and linguistic screening data should be collected before the initiation of the research in order to have some basis on which to describe the learning abilities of the children who remained in and who left the sample.

Another weakness of this study was the lack of an attempt to link student outcomes with particular teaching practices. Within the context of *The Balanced and Flexible Literacy Diet* system, the teachers were given autonomy to develop teaching strategies that provide the necessary literacy components. For example, practices such as building sight word vocabulary

using a Word Wall, developing reading skills through Guided Reading practices and developing phonemic awareness and letter sound knowledge through the use of *Jolly Phonics* (Lloyd, 1992), were modeled and discussed in professional development sessions, but no measures were taken to determine whether the use of any particular practices were consistently and effectively implemented. Thus, it is not possible, in this study, to determine the contributions of any specific teaching practices to better student literacy outcomes. Therefore, future research on the impact of PD in literacy education should be designed to relate specific classroom approaches and materials to children's literacy success.

Final Comment

Inadequate teacher preparation in literacy education can have pervasive consequences. The 1998 National Assessment of Educational Progress (NAEP) report noted that 38% of fourth graders were reading below the basic level, "yet only 5 to 10% of the population are estimated to be dyslexic with serious reading problems in need of specialized instruction" (North, 1999, p. 7). There is a fundamental problem with teacher preparation and curriculum when more than one third of fourth graders are not proficient in reading. Recent research also reports that schools are under prepared to educate children from diverse backgrounds, experiences and interests (Allington, McGill-Franzen & Schick, 1997). The present research suggests that there is something that can be done to improve this situation.

Reference List

- Abell, S. K., Dillon, D. R., Hopkins, C. J., McInerney, W. D., & O'Brien, D. G. (1995). "Somebody to count on": Mentor/Intern relationships in a beginning teacher Internship program. *Teaching & Teacher Education, 11*(2), 173 - 188.
- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print* Cambridge, MA: MIT Press.
- Adams, M. J., & Bruck, M. (1995). Resolving the 'Great Debate.' *American Educator, 19*, 7, 10-20.
- Afflerbach, P. (1998). Reading assessment and learning to read. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 239-263). New York, NY: The Guilford Press.
- Allington, R. L. (1983). Fluency: The neglected reading goal in reading instruction. *The Reading Teacher, 36*, 556-561.
- Allington, R. L. (1994). The schools we have. The schools we need. *The Reading Teacher, 48*, 14-29.
- Allington, R. L. (1995). Literacy lessons in the elementary schools: Yesterday, today and tomorrow. In R. L. Allington & S. A. Walmsley (Eds.), *No quick fix. Rethinking literacy programs in America's elementary schools* (pp.1-15). New York: HarperCollins Publishers, Inc.
- Allington, R. L., & Cunningham, P. M. (1996). *Schools that work: Where all children read and write*. New York: HarperCollins Publishers Inc.
- Allington, R. L., McGill-Frantzen, A., & Schick, R. (1997). How administrators understand learning difficulties. *Remedial and Special Education, 18*, 223-232.
- Allington, R. L., & Walmsley, S. A. (Eds.). (1995). *No quick fix. Rethinking literacy programs in America's elementary schools*. New York: Teachers College Press.
- American Federation of Teachers. (1999). *Teaching reading is rocket science*. Washington, D. C. Author.
- Archer, J. (1999 February 3). Teacher quality: A report on the preparation and qualifications of public school teachers. *Education Week*.

- Armbruster, B. B., & Nagy, W. E. (1992). Vocabulary in content area lessons. *The Reading Teacher*, 45 (7), 550-551.
- Armbruster, B. B., Anderson, T. H., & Osterag, J. (1989). Teaching text structure to improve reading and writing. *The Reading Teacher*, 43, 130-137.
- Baker, J. E. (1977). Application of the in-service training/consultation model to reading instruction. *The Ontario Psychologist*, 9(4), 57-62.
- Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly*, 34, 452-477.
- Barnes-Haig, M., & Willows, D. M. (1999). *Changing perceptions of literacy through professional development*. Unpublished manuscript, OISE/UT.
- Bell, B., & Gilbert, J. (1994). Teacher development as professional, personal, and social development. *Teaching & Teacher Education*, 10, 483 - 497.
- Bennett, B. (1987). The effectiveness of staff development training practices: A meta-analysis. Unpublished doctoral dissertation, University of Oregon, Oregon.
- Bereiter, C., & Scardamalia, M. (1982). From conversation to composition: The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology: Vol. 2* (pp. 1-64). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Beynon, C. (1999). Where it's A.T. Focus on the associate teacher. *Professionally Speaking*, June 1999, 39-41.
- Biemiller, A. (1999). Language and reading success. In J. S. Chall (Series Editor), *From reading research to practice: A series for teachers*. Cambridge, MA: Brookline Books.
- Bolton, W. F. (1994). Language: An introduction. In V. P. Clark, P.A. Eschholz, & A. F. Rosa (Eds.). *Language: Introductory readings (Fifth edition)*. New York: St. Martin's Press. Pp. 3-15.
- Booth, M. (1993). The effectiveness and role of the mentor in school: The students' view. *Cambridge Journal of Education*, 23(2), 185-197.
- Borko, H., & Putnam, R. T. (1995). Expanding a teacher knowledge base. A cognitive psychological perspective in professional development. In T. R. Guskey and M.

- Huberman (Eds.), *Professional development in education* (pp. 35-66). New York: Teachers College.
- Brady, S. (1999). Professional development for teachers: Raising the bar. *Perspectives: The International Dyslexia Association*, 25(4), 1,8.
- Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Education phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia*, 44, 26-59.
- Bromley, K. (1999). Key components of sound writing instruction. In L. B. Gambrell, L. M. Morrow, S. B. Neuman & M. Pressley (Eds.), *Best practices in literacy instruction* (pp.152-174). New York, NY: The Guilford Press.
- Bromley, K., Irwin De Vitis, L., & Modlo, M. (1995). *Graphic organizers: Visual strategies for pro active learning*. New York: Scholastic.
- Brown, G. A., & Ellis, N. (Eds.) (1994). *Handbook of spelling: Theory, process and intervention*. New York, NY: Wiley.
- Burns, P.C., & Roe, B. D. (1985). *Burns/Roe Informal Reading Inventory*. Boston, MA: Houghton Mifflin Company.
- Chall, J. S. (1967, 1983, 1996). *Learning to read: The great debate*. New York: McGraw-Hill (Harcourt Brace).
- Chall, J. S. (1983/1996). *Stages of Reading Development*. New York: McGraw-Hill.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis. Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chambers, R., Presseault, L., & Sheehan, L. (1999). Preparing Ontario's Teachers. *Professionally Speaking, September 1998*, 46-48.
- Chard, D. J. (1999). Reading certification for special education teachers: Making the difference. *Perspectives: The International Dyslexia Association*, 25(4), 10-13,
- Cisero, C. A., & Royer, J. M. (1995). The development and cross-language transfer of phonological awareness. *Contemporary Educational Psychology*, 20, 275-303.
- Clay, M. M. (1986a). The Reading Recovery research reports, Part I. *Reading-Canada-Lecture*, 4, 93-106.
- Clay, M. M. (1986b). The Reading Recovery research reports, Part II. *Reading-Canada-Lecture*, 4, 158-170.

- Clay, M. M. (1991). Attention to concepts about print. In *becoming literate: The construction of inner control* (pp. 141-154). Portsmouth, NH: Heinemann.
- Cohen, J. (1988). *Statistical power analysis for the behaviour sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Conley, M. M. W. (1983). Increasing students' reading achievement via teacher inservice education. *The Reading Teacher, April*, 804—808.
- Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*(2), 175-187.
- Cunningham, A. E., & Stanovich, K. E. (1993). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*, 264-274.
- Cunningham, P. M. (1995). *Phonics they use (Second Edition)*. New York: HarperCollins College Publishers.
- Cunningham, P. M. (1999). What should we do about phonics? In L. B Gambrell, L. M. Morrow, S. B. Neumann & M. Pressley (Eds.), *Best practices in literacy instruction* (pp. 68-89). New York, NY: The Guilford Press.
- Cunningham, P. M., & Allington, R. L. (1994,1999). *Classrooms that work: They all can read and write*. New York: HarperCollins College Publishers.
- Cunningham, P. M., Hall, D. P., & Defee, M. (1998). Nonability-grouped, multilevel instruction eight years later. *The Reading Teacher, 51*, 652-664.
- Daniels, H. A. (1994). Nine ideas about language. In V. P. Clark, P. A. Eschholz, A. F. Rosa (Eds.), *Language: Introductory readings (Fifth edition)*. New York: St. Martin's Press (pp. 201-223).
- Darling-Hammond, L. (1996). The quiet revolution, rethinking teacher development. *Educational Leadership, 53*, 4-10.
- Del Schalock, H., & Myton, D. V. (1988). A paradigm for teacher licensure: Oregon's demand for evidence of success in fostering learning. *Journal of Teacher Education, 39*, 8-16.
- Depre, H., & Iverson, S. (1994). *Early literacy in the classroom*. Richmond Hill, Ontario: Scholastic.

- Domi, H., & Willows, D. M. (2000). *The relation between classroom practice and literacy outcomes in grade one classrooms*. Unpublished manuscript, OISE/UT.
- Ebmeier, H., Twombly, S., & Teeter, D. J. (1991). The comparability and adequacy of financial support for schools of education. *Journal of Teacher Education, 42*, 226-235.
- Ehri, L. (1986). Sources of difficulties in learning to read and spell. In M. I. Woolraich & D. Routh (Eds.). *Advances in developmental and behavioral pediatrics (Vol. 7)*, pp. 121-195. Greenwich, CT: JAI Press.
- Ferguson, L. (1999, Winter). Literacy for all. *Tuftonia, 8-13*.
- Fielding, L. G., & Pearson, P. D. (1994). Reading comprehension: What works. *Educational Leadership, (Feb. 1994)*, 62-68.
- Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research, 65(2)*, 145-190.
- Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55.
- Freedman, S. W. (1998). Some things we know about learning to write. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues for teaching and learning* (pp. 189-204). New York, NY: The Guilford Press.
- Fresch, M. J., & Wheaton, A. (1997). Sort, search, and discover: Spelling in the child centered classroom. *The Reading Teacher, 51(1)*, 20-31.
- Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of Dyslexia, 36*, 69-102.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children, 60(6)*, 518-537.
- Fullan, M. (1995). The limits and potential of professional development. In T. R. Guskey & M. Huberman (Eds.), *Professional Development in Education* (pp. 253-267). New York: Teachers College Press.

- Fullan, M. (2000). The three stories of educational reform: Inside; inside/out; outside/in. *Phi Delta Kappan*, 81(8), 581-584.
- Gaffney, J. (1998). The prevention of reading failure: Teach reading and writing. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 100-110). New York: The Guilford Press.
- Gambrell, L. G. (1996). Creating classroom cultures that foster reading motivation. *The Reading Teacher*, 50 (1), 14-25.
- Gaskins, I. W., Ehri, L. C., Cress, C., O'Hara, C., & Donnelly, K. (1996/1997). Procedures for word learning: Making discoveries about words. *The Reading Teacher*, 50(4), 312-327.
- Gersten, R., & Brengelman, S. U. (1996). The quest to translate research into classroom practice: The emerging knowledge base. *Remedial and Special Education*, 17, 67-74.
- Gersten, R., Morvant, M., & Brengelman, S. (1995). Close to the classroom is close to the bone: Coaching as a means to translate research into classroom practice. *Exceptional Children*, 62, 52-66.
- Gersten, R., & Woodward, J. (1995). A longitudinal study of transitional and immersion bilingual education programs in one district. *The Elementary School Journal*, 95(3), 223-239.
- Geva, E. (2000). Issues in the assessment of reading disabilities in L2 children: Beliefs and research evidence. *Dyslexia*, 6, 13-28.
- Geva, E. & Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing: An Interdisciplinary Journal*, 12(1), 1-31.
- Geva, E., Wade-Woolley, L., & Shany, M. (1997). Development of reading efficiency in first and second language. *Scientific Studies of Reading*, 1(2), 119-144.
- Goldenberg, C. (1996). The education of language-minority students: Where are we, and where do we need to go? *The Elementary School Journal*, 96(3), 353-361.
- Gough, P. B. (1996). How children learn to read and why they fail. *Annals of Dyslexia*, 46, 3-20.

- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*(1), 6-10.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly, 25*(3), 375-406.
- Grant, P., & Mindell, N. (1989). A school based in-service education package using problem centered approach. *Educational Psychology in Practice, 5*(2), 97-102.
- Griffith, P. L., & Olson, M. W. (1992). Phonemic awareness helps beginning readers break the code. *The Reading Teacher, 45*, 516-523.
- Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*, 5-12.
- Guskey, T. R. (1995). Professional development in education: In search of the optimal mix. In T. R. Guskey & M. Huberman (Eds.), *Professional Development in Education* (pp. 114-131). New York: Teachers College Press.
- Guskey, T. R. (1998). The age of our accountability. *Journal of Staff Development, 19*(4), 36-44.
- Guthrie, J. T. (1996). Educational contexts for engagement in literacy. *The Reading Teacher, 49* (1), 432-445.
- Guthrie, J. T., & Wigfield, A. (1997). Reading engagement: A rationale for theory and teaching. In J.T. Guthrie & A. Wigfield (Eds.), *Reading engagement: Motivating readers through integrated instruction* (pp.1-12). Newark, DE: International Reading Association.
- Haas, E., & Willows, D. M. (1998). The development of spelling in a balanced literacy program: Assessing growth of phonological and orthographic knowledge. In *Changing School Literacy through In-service Teacher Education: A Four-Year Longitudinal Study*. A Symposium conducted at the National Reading Conference in Austin, TX.
- Hansen, J. (1981). The effects of inference education and practice on young children's reading comprehension. *Reading Research Quarterly, 16*, 391-417.
- Hansen, J. (1998). Young writers: The people and purposes that influence their literacy. In Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 205-236). New York, NY: The Guildford Press.

- Hargreaves, A. (1995). Development and desire: A postmodern perspective. In T. R. Guskey & M. Huberman (Eds.), *Professional development in education* (pp. 9 - 34). New York: Teachers College Press.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *SSLA*, 14, 25-38.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of the writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hiebert, E. H., Colt, J. M., Catto, S. L., & Gury, E. C. (1992). Reading and writing of first-grade students in a restructured Chapter 1 program. *American Educational Research Journal*, 29, 545-572.
- Hoover, N. L., O'Shea, L. J., & Carroll, R. G. (1988). The supervisor-intern relationship and effective interpersonal communication skills. *Journal of Teacher Education*, 39, 22-27.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Housego, B. E. J. (1994). How prepared were you to teach? Beginning teachers assess their preparedness. *The Alberta Journal of Educational Research*, 40(3), 355-373.
- Howey, K. R., & Zimpher, N. L. (1986). The current debate on teacher preparation. *Journal of Teacher Education*, 37(5), 41-49.
- Huberman, M. (1995). Professional careers and professional development. In T. R. Guskey & M. Huberman (Eds.), *Professional development in education* (pp. 193-224). New York: Teachers College Press.
- Jackett, E., & Willows, D. (1998). Development of story schemata in the written compositions of primary students: A longitudinal study of the *Balanced and Flexible Literacy Diet* system. In *Changing School Literacy through In-service Teacher Education: A Four Year Longitudinal Study*. Symposium conducted at the National Reading Conference in Austin, TX.
- Jaeger, R. M. (1992). World class standards, choice, and privatization: Weak measurement serving presumptive policy. *Phi Delta Kappan*, 74, 118-128.

- Jastak, S., & Wilkinson, G.S. (1984). *Wide Range Achievement Test - Revised*.
Wilmington, DE: Jastak Associates, Inc.
- Johns, J. L., & Vanleirsburg, P. (1994). Promoting the reading habit: Considerations and strategies. In E. H. Cramer & M. Castle (Eds.), *Fostering the love of reading: The affective domain in reading education* (pp. 91-103). Newark, DE: International Reading Association.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.
- Joyce, B., & Showers, B. (1988). *Student achievement through staff development*. New York: Longman, Inc.
- Juel, C. (1987). Beginning Reading. In D. Bloome (Ed.), *Literacy and schooling* (pp. 759-788). Norwo, NJ: Albex Publishing Corporation.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.
- Kagan, D. M., Freeman, L. E., Horton, C. E., & Rountree, B. S. (1993). Personal perspectives on a school-university partnership. *Teaching & Teacher Education*, 9, 499-509.
- Kameenui, E. (1998). The rhetoric of all, the reality of some, and the unmistakable smell of mortality. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp.319-338). New York: The Guilford Press.
- Kwan, A., & Willows, D. M. (1998, December). Impact of early phonics instruction on children learning English as a second language. In *More About Early Systematic Phonics in Balanced Literacy Classrooms*. Symposium conducted at the National Reading Conference in Austin, TX.
- Levine, M. (1987). *Developmental variation and learning disorders*. Cambridge, MA: Educators Publishing Service, Inc.
- Lilly, M. S. (1992). Research on teacher licensure and state approval teacher education programs. *Teacher Education and Special Education*, 15, 148-160.
- Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91, 501-536.

- Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis, 15*, 129-151.
- Lloyd, S. (1992). *The phonics handbook*. United Kingdom: Jolly Learning Ltd.
- Lundberg, I., Frost, J., & Peterson, O. P. (1988). Effects of an extensive program for stimulative phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263-284.
- Lyon, G. R. (Jan./Feb. 2000). *Why reading is not a natural process*. Web publication on idonline.org.
- Lyon, G. R., Vaassen, M., & Toomey, F. (1989). Teachers' perceptions of their undergraduate and graduate preparation. *Teacher Education and Special Education, 12*, 164-169.
- MacGinitie, W. H. (1991). Reading instruction: Plus ca change... *Educational Leadership, 48*, 55-58.
- MacGinitie, W. H., & MacGinitie, R.K. (1992). *Gates MacGinitie Reading Tests Second Canadian Edition*. Canada: Nelson Canada Ltd.
- Manzo, A. V. (1991). Education teachers to use content area reading strategies: Descriptions and appraisal of four options. *Reading Research and Instruction, 30*(4), 67-73.
- Maria, K. (1989). Developing disadvantaged children's background knowledge interactively. *The Reading Teacher, 42*, 296-300.
- Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1981). A cognitive developmental theory of reading acquisition. In T. G. Waller & G. E. MacKinnon (Eds.). *Reading research: Advances in theory and practice (Vol. 3) (pp. 199-221)*. New York: Academic Press.
- McKeown, M. G., & Beck, I. L. (1988). Learning vocabulary: Different ways for different goals. *Remedial and Special Education, 9*(1), 42-52.
- Metcalf-Turner, P., & Fischetti, J. (1996). Professional development schools: Persisting questions and lessons learned. *Journal of Teacher Education, 47*(4), 292-299.
- Mevarech, Z. R. (1995). Teachers' paths on the way to and from professional development forum. In T. R. Guskey & M. Huberman (Eds.), *Professional development in education (pp.151-170)*. New York: Teachers College Press.

- Miller, J. W., & Ellsworth, R. (1985). The evaluation of a two-year program to improve teacher effectiveness in reading instruction. *The Elementary School Journal*, 85, 485-495.
- Moats, L. C. (1995). The missing foundation in teacher education. *American Educator*, 19(9), 43-51.
- Moats, L. C., & Foorman, B. R. (1997). Introduction to special issue of SSR: Components of effective reading instruction. *Scientific Studies of Reading*, 1, 188-189.
- Moats, L. C., & Lyon, G. R. (1996). Wanted: Teachers with knowledge of language. *Topics in Language Disorders*, 16, 73-86.
- Monroe, E. E., & Smith, W. E. (1985). Guidelines for improving inservice reading education. *Reading Improvement*, 22, 123-125.
- Morgan, J., & Willows, D. M. (1997, March). Reducing the risks: An early literacy program for ESL students. Paper presented at the meeting of the AERA, Chicago, IL.
- Morris, J. M. (1985). The need for linguistics-informed teachers. *Early Child Development and Care*, 23, 41-52.
- National Commission On Excellence in Education. (1984). *A Nation at risk. The full account*. Cambridge, MA: USA Research.
- National Reading Panel (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC: National Institute of Child Health and Human Development (NICHD).
- Neuman, S. B. (1999). Creating continuity in early literacy: Linking home and school with a culturally responsive approach. In L. B. Gambrell, L. M. Morrow, S. B. Neuman & M. Pressley (Eds.), *Best practices in literacy education* (pp.258-270). New York: The Guilford Press.
- Nevin, A., & Thousand, A. N. J. (1986). What the research says about limiting or avoiding referrals to special education. *Teacher Education and Special Education*, 9(4), 149-161.
- Nolen, P. A., McCutchen, D., & Berninger, V. (1990). Ensuring tomorrow's literacy: A shared responsibility. *Journal of Teacher Education*, 41, 63-72.

- North, M.E. (1999). What constitutes adequate preparation for teachers of reading?
Perspectives: The International Dyslexia Association, 25(4), 4-7.
- Obiakor, F. E., & Utley, C. A. (1997). Rethinking preservice preparation for teachers in the learning disabilities field: Workable multicultural strategies. *Learning Disabilities Research & Practice, 12(2), 100-106.*
- Ontario College of Teachers. (1999). *Pre-service teacher education: Initial accreditation Handbook (3rd ed.)*. Toronto: Author. _
- Orton Dyslexia Society. (1997). *Informed instruction for reading success: Foundations for teacher preparation*. Baltimore: Orton Dyslexia Society.
- Palincsar, A. S., & Brown, A. (1987). Advances in improving the cognitive performance of handicapped students. In W. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), *Handbook of special education, research and practice, Volume 1* (pp. 93-111). New York: Pergamon Press.
- Pikulski, J. J. (1994). Preventing reading failure: A review of five effective programs, *The Reading Teacher, 48, 30-39.*
- Placier, P., & Hamilton, M. L. (1994). Schools as contexts: A complex relationship. In V. Richardson (Ed.), *Teacher change and the staff development process* (pp. 135-158). New York: Teachers College Press.
- Pressley, M. (1998a). Comprehension strategies instruction. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 113-133). New York: The Guilford Press.
- Pressley, M., & Wharton-MacDonald, R. (1998). The development of literacy, Part 4: The need for increased comprehension instruction in upper elementary grades. In *Reading instruction that works: The case for balanced teaching* (pp. 192-227). New York: The Guilford Press.
- Pressley, M. (1999). Self-regulated comprehension processing and its development through instruction. In L. B. Gambrell, L. M. Morrow, S. B. Neuman, & M. Pressley (Eds.), *Best practices in literacy instruction* (pp.90-97). New York: The Guilford Press.

- Pressley, M., Rankin, J., & Yokoi, L. (1996). A survey of instructional practices of primary teachers nominated as effective in promoting literacy. *The Elementary School Journal*, 96(4), 365-384.
- Reback, M. (1999). The role of teacher unions in improving the quality of reading instruction: Preservice, professional development and recertification. *Perspectives The International Reading Association*, 25(4), 16-19.
- Reiman, A. J., & Thies-Sprinthall, L. (1993). Promoting the development of mentor teachers: Theory and research programs for guided reflection. *Journal of Research and Development in Education*, 26(3), 179-185.
- Reynolds, R. J. (1995). The professional self-esteem of teacher educators. *Journal of Teacher Education*, 46, 216-226.
- Reynolds, M. C., Wang, M. C., & Walberg, H. J. (1992). The knowledge bases for special and general education. *Remedial and Special Education*, 13, 6-10, 33.
- Richardson, S. O. (1996). Coping with dyslexia in the regular classroom: Inclusion or exclusion. *Annals of Dyslexia*, 46, 37-48.
- Richardson, V. (1994). The consideration of teachers' beliefs. In V. Richardson (Ed.), *Teacher change and the staff development process* (pp. 90-103). New York: Teachers College Press.
- Richardson, V. (1998). Professional development in the instruction of reading. In J. Osborne & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 303-318). New York: The Guilford Press.
- Richardson, V., & Hamilton, M. L. (1994). The practical-argument staff development process. In V. Richardson (Ed.), *Teacher change and the staff development process* (pp.109-129). New York: Teachers College Press.
- Rogers, W. T. (1991). Education assessment in Canada: Evolution or extinction? *The Alberta Journal of Educational Research*, 37(2), 179-192.
- Rolheiser, C. (1996). Partnership and program renewal: Moving forward. In D. Booth & S. Stiegelbauer (Eds.). *Teaching teachers: The faculty of education, University of Toronto 1906-1996*. Hamilton, ON: Caliburn Enterprises Inc.
- Rolheiser, C. (1999). Redesigning teacher education: The delicate, demanding dance of "ready, fire, aim." In M. Nideen & P. Lemma (Eds.). *Ground level reform in*

- teacher education: Changing schools of education* (pp. 119-148). Calgary, AB: Detselig Enterprises
- Rust, F. O. (1988). How supervisors think about teaching. *Journal of Teacher Education*, 39, 56-64.
- Salinger, T. (1995). IRA, standards, and educational reform. *The Reading Teacher*, 49, 290-298.
- Samuels, S. J. (1997). The method of repeated readings. *The Reading Teacher*, 50(5), 376-381.
- Scanlon, D. M., & Vellutino, F. R. (1997). A comparison of the instructional backgrounds and cognitive profiles of poor, average, and good readers who were initially identified as at risk for reading failure. *Scientific Studies of Reading*, 1, 191-215.
- Scarborough, H. S., Ehri, L. C., Olson, R. K., & Fowler, A. E. (1998). The fate of phonemic awareness beyond the elementary school years. *Scientific Studies of Reading*, 2(2), 115-142.
- Scardamalia, M. (1981). How children cope with the cognitive demands of writing. In C. H. Fredericksen & J. E. Dominic (Eds.), *Writing: The nature, development and teaching of written communication: Vol. 2. Processes, development and communication* (pp. 81-103). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Shanahan, T. (1988). The reading-writing relationship: Seven instructional principles. *The Reading Teacher*, 41, 636-647.
- Shanahan, T., & Samuels, S. J. (2000). Building fluency. *Report of the National Reading Panel*.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *The New England Journal of Medicine*, 326, 145-196.
- Shen, J. (1995). Faculty fragmentation and teacher education in schools, colleges, and departments of education. *Journal of Teacher Education*, 46, 141-149.
- Sheppard, R. (1998, September 7). Why kids can't read. *Maclean's*, 111(36), 40-44.

- Showers, B. (1990). Aiming for superior classroom instruction for all children: A comprehensive staff development model. *Remedial and Special Education, 11*, 35-39.
- Showers, B., Joyce, J., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and state-of-the-art analysis. *Educational Leadership, 45*(3), 77-87.
- Siegel, L. S., & Ryan, E. B. (1984). Reading disability as a language disorder. *Remedial and Special Education, 5*(3), 28-33.
- Siegelman, K., & Willows, D. M. (2000). *Implementing change in early literacy education through professional development of teachers*. Unpublished manuscript.
- Simpson, R. L., Whelan, R. J., & Zabel, R. H. (1993). Special education personnel preparation in the 21st century: Issues and strategies. *Remedial and Special Education, 14*(2), 7-22.
- Sindelar, P. T. (1995). Full inclusion of students with learning disabilities and its implications for teacher education. *The Journal of Special Education, 29*(2), 234-244.
- Slavin, R.E., Madden, N., Karweit, N., Dolan, L., & Wasik, B. (1991). Research directions: Success for All: Ending reading failure from the beginning. *Language Arts, 68*, 404-409.
- Smylie, M. (1989). Teachers' views of the effectiveness of sources of learning to teach. *The Elementary School Journal, 89*(5), 543-558.
- Smylie, M. (1995). Teacher learning in the workplace: Implications for school reform. In T. R. Guskey & M. Huberman (Eds.), *Professional development in education* (pp. 92-113). New York: Teachers College Press.
- Snow, C. E., Barnes, W. S., Chandler, J., Goodman, I. F., & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy*. Cambridge, MA: Harvard University Press.
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Soltis, J. F. (1990). A reconceptualization of educational foundations. *Teachers College Record, 91*(3), 311-321.

- Sparks, G. M. (1988). Teachers' attitudes toward change and subsequent improvements in classroom teaching. *Journal of Educational Psychology, 80*, 111-117.
- Spiegel, D. L. (1998). Silver bullets, babies and bath water: Literature response groups in balanced literacy program. *The Reading Teacher, 52*(2), 114-124.
- Stahl, S. A. (1992). Saying the "p" word: Nine guidelines for exemplary phonics instruction. *The Reading Teacher, 45*(8), 618-625.
- Stallings, J., & Krasavage, E. M. (1986). Program implementation and student achievement in a four-year Madeline-Hunter follow-through project. *The Elementary School Journal, 87*(2), 117-138.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16* (1), 33-69.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.
- Stanovich, K. E. (1991). Cognitive science meets beginning readers. *Psychological Science, 2*, 70, 77-81.
- Stanovich, K. E. (1993/94). Romance and reality. *The Reading Teacher, 47*, 280-291.
- Stanovich, K. E. (2000). Progress in understanding reading: Scientific foundations and new frontiers. New York: The Guilford Press.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading difficulties: A regression-based test of Phonological-Core Variable Difference Model. *Journal of Educational Psychology, 86*, 24-53.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. (3rd edition). New Jersey: Lawrence Erlbaum Associates.
- Stornelli, D., & Willows, D. M. (1998, December). Effect of more and earlier phonics instruction on kindergarten outcomes. In *More About Early Systematic Phonics in Balanced Literacy Classrooms*. Symposium conducted at the National Reading Conference in Austin, TX.
- Sumbler, K. (1999). *Phonological awareness combined with explicit alphabetic coding instruction in kindergarten: Classroom observations and evaluation*. Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario, Canada.

- Sumbler, K., & Willows, D. M. (1998, December). Time management: Monitoring activities in *Jolly Phonics* and control classrooms. In *More About Early Systematic Phonics in Balanced Literacy Classrooms*. Symposium conducted at the National Reading Conference in Austin, TX.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. (3rd edition). New York: Harper and Row.
- Taylor, B. M., Short, R. A., Frye, B. J., & Shearer, B. A. (1992). Classroom teachers prevent reading failure among low-achieving first-grade students. *The Reading Teacher*, 45, 592-597.
- Tidwell, D. L., & Mitchell, J. N. (1994). Teaching reading and observing teachers' practices. In V. Richardson (Ed.), *Teacher change and the staff development process* (pp. 43-67). New York: Teachers College Press.
- Tillema, H. H. (1995). Changing the professional knowledge and beliefs of teachers: An education study. *Learning and Instruction*, 5, 291-318.
- Torgesen, J. K., & Goldman, T. (1977). Verbal rehearsal and short-term memory in reading disabled children. *Child Development*, 48, 56-60.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A., (1997). Prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading*, 1, 217-234.
- Turner, J., & Paris, S. G. (1995). How literacy tasks influence children's motivation for literacy. *The Reading Teacher*, 48(8), 662-673).
- Veal, M., & Rikard, L. (1998). Associate teachers' perspectives on the student teaching triad. *Journal of Teacher Education*, 49, 108-119.
- Veenman, S., Van Tulder, M., & Voeten, M. (1994). The impact of in-service education on teacher behaviour. *Teaching & Teacher Education*, 10, 303-317.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, 88, 601-638.

- Watson, N., & Hart, D., & Jacka, N. (1997, August). *Learning to teach: The OISE/UT two-year pilot program (A Program Evaluation Report)*. Toronto, ON: OISE/UT.
- Wilkinson, G. A. (1994). Self-direction empowers inservice teachers. *College Student Journal*, 28, 77-87.
- Willows, D. M. (1991). A "normal variation" view of written language difficulties and disabilities: Implications for whole language programs. *Exceptionality Education Canada*, 1 (3), 73-103.
- Willows, D. M. (1993). *Learning to read and write. Adapting programs to children's needs*. Unpublished manuscript, Ontario Institute for Studies in Education/University of Toronto.
- Willows, D. M. (1994). *The Balanced and Flexible Literacy Diet: Putting Theory into Practice*. Unpublished manuscript, OISE/University of Toronto.
- Willows, D. M. (1997, December). Symposium Discussant *Phonological and orthographic processes in literacy development: Longitudinal and cross-linguistic perspectives*. Presentation at the meeting of the National Reading Conference, Scottsdale, AZ.
- Willows, D. M. (2000). *The balanced and flexible literacy diet: Putting theory into practice*. Manuscript in preparation for publication to Pippin Publishing, Toronto.
- Willows, D. M., & Burgess, M. (1996, December). The balanced and flexible literacy diet: Putting theory into practice. In *Raising literacy levels through in-service teacher education*. A Symposium conducted at the National Reading Conference in Charleston, SC.
- Willows, D. M., & Dixon, M. (1998, December). Implementing and maintaining change in primary literacy program: A longitudinal case study of *The Balanced and Flexible Literacy Diet*. In *Changing school literacy through in-service teacher education: A four-year longitudinal study*. A Symposium conducted at the National Reading Conference in Austin, TX.
- Willows, D. M., & Ryan, E..B. (1986). The development of grammatical sensitivity and its relationship to early reading achievement. *Reading Research Quarterly*, 23, 253-266.

- Winograd, P., & Arrington, H. J. (1999). Best practices in literacy assessment. In L. B. Gambrell, L. M. Morrow, S. B. Neuman & M. Pressley (Eds.), *Best practices in literacy instruction* (pp.210-241). New York, NY: The Guilford Press.
- Yen, M. (1997). The technical quality of performance assessments: Standard errors of percents of pupils reaching standards. *Educational Measurement: Issues and Practice*, 16(3), 5-15.
- Yopp, H. K. (1992). Developing phonemic awareness in young children. *The Reading Teacher*, 45, 696-703).
- Zaborik, J. A. (1988). The observing-conferencing role of university supervisors. *Journal of Teacher Education*, 39, 9-16.
- Zimpher, N. L., deVoss, G. G., & Nott, D. L. (1980). A closer look at university student teacher supervision. *Journal of Teacher Education*, 31(4), 11-15.

Appendix A

Parental Consent Letter

Dear Parents,

In order to monitor the effectiveness of our primary literacy program, an assessment of children's reading, spelling and writing skills, as well as some aspects of oral language (audio taped), will be conducted over the next few weeks. This evaluation is a follow up of assessments undertaken in our school over the last three years. During May and June, each student will participate in two individual sessions of 20-30 minutes and a group session in the classroom of approximately 30 minutes. These sessions are designed to be fun and interesting for children and, based on past experience, they find them quite enjoyable.

The evaluation will be conducted under the supervision of Dr. Dale Willows, a professor at the Ontario Institute for Studies in Education / University of Toronto, and a well-respected reading and writing specialist who has been providing workshops to schools in the region. She and her assistants will be collecting the assessment information.

We would like permission to allow your child to be included in this data collection so that we can determine the success of our primary literacy program for all students, and so that we can design modifications that may be needed for future curriculum planning. In this research project Dr. Willows will be examining group results. However, for our teachers, the data can also provide useful information to aid in program planning for individual students.

It is hoped that you will sign the attached consent form and return it to the school as soon as possible, so that your child can participate in this valuable research.

If you have any questions about the research, please contact me.

Sincerely,

Principal

PERMISSION FORM

I give permission for _____, age _____, to
(Child's Name)
participate in the data collection to be conducted under the guidance of Dr Dale Willows.

My child's first language is _____

My child speaks which languages at home? English _____ and _____
(Yes or No) (Other language)

My child's date of birth is _____
Day / Month / Year

Date

Signature

Name (please print)

Appendix B

Effect of Home Language on Literacy Acquisition

Appendix B

Effect of Home Language on Literacy Acquisition

A potential risk factor for poor literacy skills is home language. Some educators argue that the increasing cultural and linguistic diversity in school populations is one of the causes for the large number of children experiencing problems with learning to read. National estimates of the growth in the United States in total school-age population project an increase of more than 20 percent, from 34 million to 42 million in 2010, and it is estimated that children of immigrants will account for more than half of this growth. The number of Anglo-American children is expected to decline by approximately 27 percent during the same period (Obiakor & Utley, 1997). The number of children acquiring English as a second language (L2) grew by 85 percent in the United States between 1985 and 1992, and over half of these children are concentrated in grades Kindergarten to four (Snow et al., 1998). In Canada, the 1996 Census showed that 33% of the total population reported a home language other than English.

Non-English speaking students also tend to come from lower socioeconomic backgrounds and attend schools with a high proportion of children in poverty. As mentioned previously, poverty is a known risk factor for poor achievement in literacy skills.

The terms ESL or L2 learners refer to students who are living in a place where they do not speak the majority's language (Fitzgerald, 1995). Many educators argue that ESL (L2) students are frequently misidentified, misassessed, and misinstructed by poorly prepared teachers (Obiakor & Utley, 1997). Since the passage, in 1990, of the Individuals with Learning Disabilities Act in the USA, distinguishing children with true learning disabilities from underachievers is a major challenge facing educators. Furthermore, a recurring issue is the overrepresentation of culturally diverse and low-income students in special education. Many

teachers have difficulty distinguishing between learning problems that reflect characteristics of second language learners or cultural differences and those that result from a learning disability.

As the population in our schools becomes more diverse, educators have become more concerned about the literacy levels of ESL (L2) students. However, not only are teachers confronted with students new to the English language but also with students who have limited exposure to English print materials in the home. Research is examining the reading processes of L2 learners and different educational programs for teaching L2 students. Reading is viewed as the most important skill for second language learners to attain (Cisero & Royer, 1995; Cummins, 1980; Fitzgerald, 1995; Gersten & Woodward, 1995; Geva, 2000; Geva & Siegel, 2000; Geva, Wade-Woolley, & Shany, 1997; Grabe, 1991; Harrington & Sawyer, 1992). Thus, it is the responsibility of teacher educators to help prospective teachers to examine the current knowledge on the similarities and differences between native English students (L1) and ESL (L2) students in the development of reading skills, and to examine their own beliefs and attitudes toward people who are different from themselves.

The theory underpinning most bilingual education programs in the United States holds that many non or limited English proficiency students require instruction in their native language in order to gain high academic achievements in English. Cummins (1980) reported that it takes students 5 to 7 years to become sufficiently competent in a second language in order to succeed in the academic program of the second language (English) in the regular classrooms. According to Cummins (1980), the removal of L2 students from ESL programs into regular mainstream classrooms where English was spoken was believed to be detrimental to their academic progress until these students were fluent in English, with 5 to 7 years of ESL instruction. This bilingual educational theory supports the following ideas: a student learns readily and easily in the

language he/she knows best – the home language. A student can learn a great deal of academic knowledge and skills in his/her home language while simultaneously learning to speak and understand a second language, and the knowledge and skills learned in the first language will transfer to the second language (i.e., English). From this perspective, L2 students will be best served in programs that build academic knowledge in the students' home language while building proficiency in English.

Due to lack of success of many bilingual instructional programs to foster high levels of literacy skills in L2 students, Goldenberg (1996) commented on the need for an evaluation of bilingual programs. In an opinion paper, Goldenberg (1996) reported the results of Grade 2 and Grade 3 students who were taught only in Spanish, and were tested on the Comprehensive Test of Basic Skills in Spanish. These students scored only at the 27th percentile even when tested in their native language, Spanish. Hispanic students in bilingual programs continue to perform poorly on academic tasks, although they were being taught in their home language.

To study the effectiveness of teaching L2 students in their home language. Gersten and Woodward (1995) examined the outcomes of two bilingual programs over a four-year period from grades 4 to 7. In one program, the Bilingual Immersion approach, the students are introduced to English in Grade 1, with a 90-minute component of Spanish-native language that is reduced to 30 minutes by Grade 4. Therefore, the early use of English language arts and English reading instruction to foster the rapid acquisition of English language is the basic theory in this program. Supporting this theory is the research examining the critical periods for language acquisition. This research shows a gradual decline in the ability to acquire a second language from the age of seven until adulthood. "The decline in learning ability begins earlier than initially thought by most researchers" (Johnson & Newport, 1989, p. 96).

The Transitional Bilingual approach is supported by Cummins research (1980). L2 students do not begin any instruction in English until grades 4 or 5. Subject matter and concepts in all academic areas are initially taught in the students' primary language.

All students in Gersten and Woodward's (1995) study had completed at least 4 years of either Bilingual Immersion or Transitional Bilingual education. The major measure in the study was achievement on the Iowa Tests of Basic Skills (ITBS) in grades 4, 5, 6 and 7. Sample sizes for the longitudinal analyses were 111 for the Bilingual Immersion program and 117 for the Transitional Bilingual program. The results demonstrated little differences between the achievement levels of these students in these two programs by grade 7. Thus, both programs, as measured by the ITBS, are equally viable options, and question the assertion that meaningful literacy instruction in English cannot occur until students have experienced many years of native-language instruction. However, the Grade 7 students in the Bilingual Immersion program scored at the 24th percentile in reading, and the students in the Transitional Bilingual program scored at the 21st percentile in reading, compared to national norms. Therefore, although there was little difference between the two programs, neither program produced students with good reading skills. The reading skills of L2 students in both programs were below average. Neither program provided effective literacy instruction.

Therefore, more research is needed to examine the components of effective literacy instruction for L2 students. This current research explores the differences in the acquisition of literacy skills between L1 and L2 children who were exposed to English language instruction by classroom teachers who had participated in professional development to improve their instructional practices in literacy.

Research Study

Given our current population of highly diverse, multi-lingual students, home language is a salient potential risk factor that may be affecting children's acquisition of written language skills. Children who arrive in the education system with English as their secondary language (L2) are generally expected to have more difficulties learning to read and spell. The exploratory analyses in this section focused on the effect of a child's primary language, whether L1 or L2, on the acquisition of literacy skills in reading and spelling. In particular this section was designed to determine whether in-service professional development for classroom teachers might attenuate any potential disadvantage that may be associated with having English as a second language.

The present examination of home language as a potential risk factor is considered only exploratory because of the relatively small numbers of children in some groups. However, given the unique opportunity to examine data over a period of time, an L1/L2 analysis was undertaken despite the limitations of sample sizes. The examination will begin with a description of the children, L1 and L2. The analyses, results and discussion will follow.

The Research Question for exploring the risk factor of home language can be stated as follows:

Did a child's first language, whether English is primary (L1) or secondary (L2), differentially affect the scores from the measures of word recognition, spelling and reading comprehension, between children who were taught by in-serviced teachers and children who were not taught by in-serviced teachers (Baseline Group)?

Participant Characteristics

The data of L1 and L2 children were included only when there was a definite designation as L1 or L2 based on parent and teacher ratings at the time of the analyses in 1997 and 1998. When the project began, L1 and L2, were not considered as variables in the study, so many children could not be included in the L1/L2 analyses because home language was not designated.

Children who spoke a language other than or as well as English in their homes were designated as L2. Given the much-reduced sample size, these results must be considered preliminary.

As shown in Table B1, the L1 and L2 children in the Treatment group were children who were taught by in-serviced teachers over 3 years (Grade 1) or 4 years (Grade 2). These Treatment groups of children were comprised of all the cohort groups in Grades 1 and 2 taught by in-serviced teachers. Therefore, the Treatment groups were the Grade 1 children who were in Cohort 1A + Cohort 1B + Cohort 1C (1995 + 1996 + 1997) and the Grade 2 children who were in Cohort 2A + Cohort 2B + Cohort 2C and Cohort 2D (1995 + 1996 + 1997 + 1998). The L1 and L2 Baseline groups were the Grade 1 and Grade 2 children taught by teachers before the initiation of the in-service education (June 1994).

Year-by-year analyses of each group of L1 and L2 children would result in sample sizes that were too small; all of the cohorts of Treatment group children (Intervention Year plus Follow-up Years) were combined into a larger group of children whose teachers had received in-service instruction.

Table B1 Summary of the Participant Characteristics

<u>Grade 1</u>						
	Baseline Group (1994) (N = 28)			Treatment Group (Cohorts 1A + 1B + 1C) (N = 67)		
Language	L1, n = 18			L1, n = 38		
	L2, n = 10			L2, n = 29		
	<u>M</u>	<u>F</u>	<u>p <</u>	<u>M</u>	<u>F</u>	<u>p <</u>
Age (months)	L1 83.1 (3.1)	.99	.33	L1 83.21 (3.97)	.174	.68
	L2 81.9 (3.1)			L2 82.79 (4.19)		

<u>Grade 2</u>						
	Baseline Group (1994) (N = 49)			Treatment Group (Cohorts 2A + 2B + 2C + 2D) (N = 87)		
Language	L1, n = 37			L1, n = 50		
	L2, n = 12			L2, n = 37		
	<u>M</u>	<u>F</u>	<u>p <</u>	<u>M</u>	<u>F</u>	<u>p <</u>
Age (months)	L1 95.89 (3.65)	.003	.96	L1 94.81 (3.49)	.214	.65
	L2 95.83 (2.52)			L2 95.23 (4.87)		

Note. English is the primary language for L1 children, and English is a secondary language for L2 children.

There was no difference between the mean ages of the L1 and L2 children in the Grade 1

Baseline group, $F(1, 26) = .986, p < .33$ or Treatment group, $F(1, 65) = .174, p < .68$.

Similarly, there were no differences between the mean ages of the L1 and L2 children in the two different Grade 2 groups, with the Baseline group L1/L2 comparison at $F(1, 47) = .003, p < .96$, and the Treatment group comparison at $F(1, 80) = .214, p < .65$.

Analyses

In order to evaluate the differences in the acquisition of reading and spelling skills between L1 and L2 children exposed to teachers who had participated in the in-service PD, two-way analyses of variance (ANOVA) were performed. One variable was language status, L1 or L2,

and the other was in-service status of the teachers, Baseline or Treatment. The dependent measures were again, word recognition, spelling, and reading comprehension.

Results and Discussion

The results of these analyses are presented for Grades 1 and 2 children, both L1 and L2, for each of the independent measures: word recognition, spelling and reading comprehension. The performance of L1 and L2 children on the literacy measures were compared between the Baseline group before the initiation of the in-service, and the composite Treatment group made up of all those children exposed to the instruction of in-serviced teachers in the Implementation and Follow-up Years.

Table B2 summarizes the Means and Standard Deviations for the 4 groups of Grade 1 and Grade 2 children (Baseline L1, Baseline L2, Treatment L1, Treatment L2). In all cases, analyses were performed on raw scores. Figures B1, B2 and B3 show the word recognition, spelling and reading comprehension scores of the Grade 1 and Grade 2 L1 and L2 children in graphic form.

Table B2
Means and Standard Deviations of the Scores on the Three Literacy Measures for the Baseline and Treatment Groups of L1 and L2 Children in Grade 1 and Grade 2

Grade 1					
<u>Teacher In-Service Status</u>					
Measures	<u>Baseline</u>			<u>Treatment</u>	
		<u>n</u>		<u>n</u>	
<u>Word Recognition</u>					
	L1	18	36.28 (37.54)	37	75.57 (36.72)
	L2	10	60.50 (41.26)	29	76.55 (35.19)
<u>Spelling</u>					
	L1	13	8.31 (4.42)	38	11.82 (4.57)
	L2	10	9.80 (4.54)	29	12.48 (4.70)
<u>Reading Comprehension</u>					
	L1	15	24.87 (13.02)	37	30.62 (11.37)
	L2	9	31.00 (13.07)	28	31.54 (9.7)

Grade 2					
<u>Word Recognition</u>					
	L1	35	81.91 (34.39)	48	100.08 (36.95)
	L2	11	86.91 (34.95)	34	117.50 (25.64)
<u>Spelling</u>					
	L1	33	12.58 (6.04)	46	16.26 (5.52)
	L2	11	14.37 (6.73)	34	19.41 (5.37)
<u>Reading Comprehension</u>					
	L1	36	28.64 (10.26)	49	34.12 (10.49)
	L2	12	29.83 (10.50)	36	35.42 (7.39)

Notes: - L1 is English as a primary language and L2 is English as a secondary language.

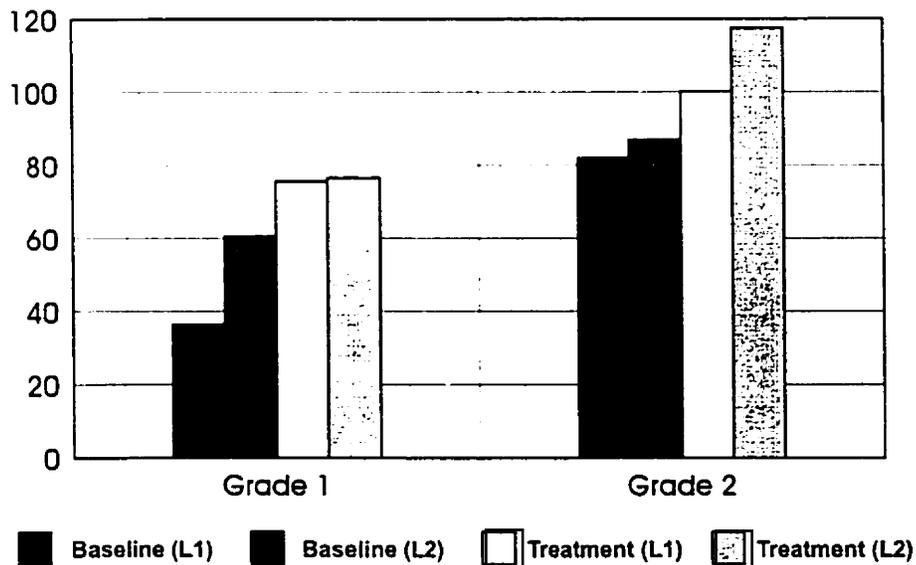
- All scores are raw scores.

- Sample sizes vary slightly across measure because not all children were present for all testing sessions.

Word Recognition

Grade 1. On the word recognition task (see Figure B1), the scores of the Grade 1 children showed a significant main effect, such that the Treatment group (children whose teachers had participated in the in-service PD) had higher word recognition scores than the Baseline group, $F(1, 90) = 10.37, p < .002$. There was, however, no significant main effect of language, $F(1, 90) = 2.15, p < .146$ and there was no significant interaction of Treatment x Language on the word recognition scores. L2 children's word recognition was at least as good as that of L1 children.

Figure B1. Mean Word Recognition for L1 and L2 Groups
(Burns & Roe Informal Reading Inventory)



In view of the fact that the mean word recognition scores of the L2 children in 1994 appeared to be higher than the word recognition scores of the L1 children, a *post hoc* t- test was performed to assess if there was a significant difference between these scores. Although there was an apparently large difference between the means of the L1 and L2 Baseline groups, this difference was not statistically significant, $t(26) = 1.58, p < .126$.

Post hoc t-tests were also performed to determine if there were significant differences between the word recognition scores of the Grade 1 Baseline and Treatment groups of L1 children and between the word recognition scores of the Grade 1 Baseline and Treatment groups of L2 children. For the L1 children there were significant effects of Treatment, $t(53) = 3.70, p < .001$. However, there was no significant effect of Treatment, instruction from teachers who had participated in the PD, on the word recognition scores of the L2 children, $t(37) = 1.19, p < .24$. Due to the superior scores of this particular sample of L2 children in the Baseline group, who were already achieving at the expected level of Grade 1 students, there may have been a ceiling effect. Therefore, it was difficult for the Treatment group of L2 children to show significant improvement. Secondly, this specific small sample of ten L2 children in the Baseline year may have more easily learned word recognition skills than the sample of L2 children assessed in the Treatment years.

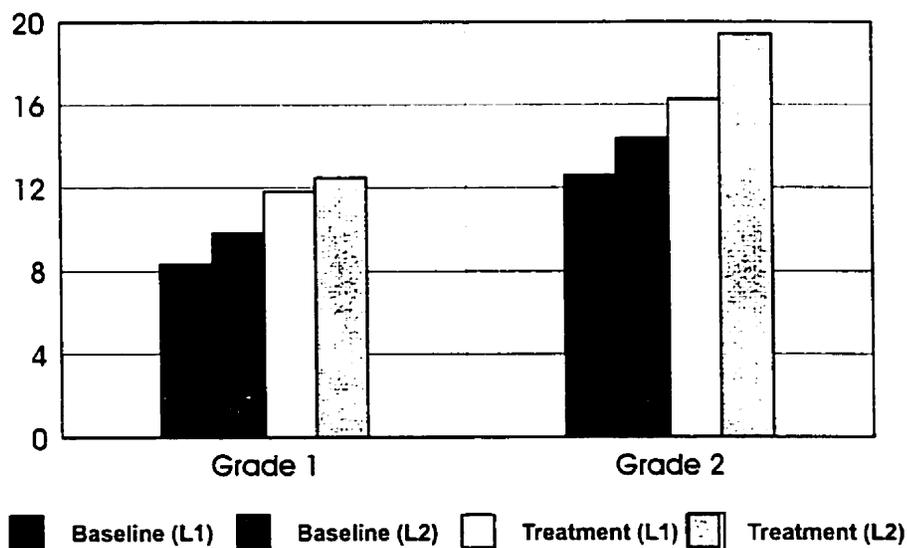
Grade 2. The word recognition scores of the Grade 2 children showed a significant main effect, such that the Treatment group had higher word recognition scores than the Baseline group, $F(1, 124) = 12.56, p < .001$. There was, however, no significant main effect of home language, $F(1, 124) = 2.65, p < .106$, and there was no significant interaction effect of Treatment x Language on the word recognition scores. These L2 children had similar word recognition skills when compared to the L1 children (see Figure B1).

Spelling

Grade 1. On spelling as well, Grade 1 children benefited from the instruction of in-serviced teachers. The main effect of Treatment was significant, $F(1, 86) = 7.47, p < .007$. The Treatment group had higher spelling scores than the Baseline group. Of particular importance is that there was no significant main effect of home language, nor was there a Treatment x Language interaction, $F(1, 86) = .93, p < .34$ and $F(1, 86) = .14, p < .713$, respectively.

Grade 2. Similarly, the spelling scores of the Grade 2 children showed a significant main effect of Treatment, $F(1, 120) = 13.48, p < .001$. Interestingly, there was a significant main effect of home language on the spelling scores, $F(1, 120) = 4.31, p < .04$. As seen in Figure B2, the mean spelling scores of the L2 children were, in fact, higher than the mean scores of the L1 children. However, there was no interaction effect of Language x Treatment on the spelling scores of the Grade 2 children, $F(1, 120) = .328, p < .57$. Thus both L1 and L2 children benefited from being in the classrooms of in-serviced teachers.

Figure B2. Mean Spelling Scores for L1 and L2 Groups
(WRAT Spelling Raw Scores)



Reading Comprehension

Grade 1. There were no significant main effects or interactions on the reading comprehension measure. Because of the appearance of a mean difference in the two Baseline groups (i.e., L2 group higher), a *post hoc* t-test was performed between the scores of the L1 and L2 Baseline groups. There was, however, no significant difference between these two groups, $t(22) = 1.12, p < .277$. T-tests were also performed between the Baseline and Treatment groups of L1 and L2 children separately. There was no effect of treatment for either the L1 or L2 groups, $t(50) = 1.59, p < .12$, and $t(35) = .132, p < .90$, respectively.

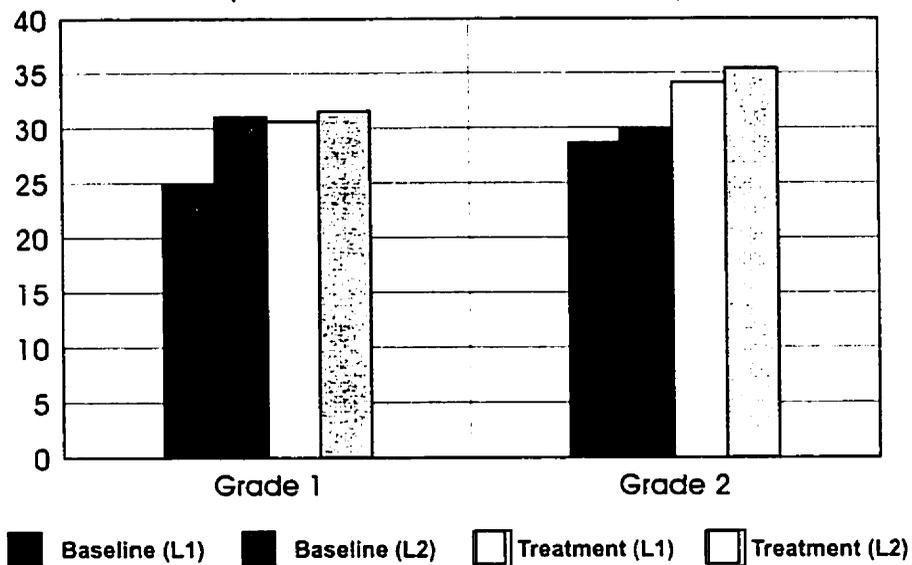
The mean reading comprehension score of the L1 children in the Baseline group was at the 48th percentile and the mean reading comprehension score of the L2 children in the Baseline group was at the 63rd percentile. Therefore, these children who were not exposed to the instruction of in-serviced teachers were already performing within the average range on the reading comprehension measure.

Grade 2. For the Grade 2 children, however, the reading comprehension scores showed a significant main effect of treatment, $F(1, 129) = 8.20, p < .005$. As with the word recognition and spelling measures, however, there was neither a main effect of Language nor a Treatment x Language interaction on the reading comprehension scores of the Grade 2 children, $F(1, 129) = .42, p < .52$, and $F(1, 129) = .001, p < .98$, respectively.

As seen in Figure B3, there is no difference between the mean reading comprehension scores of the Grade 2 children who were L1 and L2.

Figure B3. Mean Reading Comprehension
for L1 and L2 Groups

(Gates MacGinitie Raw Scores)



Summary of L1/L2 Comparisons

The main question of interest in this section was whether children who speak some language other than English (L2) in the home are more “at risk” of literacy failure than those who speak English as their first language (L1). The results across all three literacy measures in both Grades 1 and 2 indicate that, in every case, L2 children performed at least as well as L1 children. In fact, on the spelling measure, L2 children’s performance exceeded that of L1 children in Grade 2. Both Grade 1 and Grade 2 children benefited from instruction by in-serviced teachers, and the benefits were virtually the same for the L1 and L2 samples.