

**The Observation Survey of Early
Literacy Achievement Tasks:
Developing Norms for Manitoba Students**

Irene Huggins

**A thesis submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy, University of Manitoba
2003**

THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION PAGE

The Observation Survey of Early Literacy Achievement Tasks:
Developing Norms for Manitoba Students

BY

Irene Huggins

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University
of Manitoba in partial fulfillment of the requirements of the degree
of

DOCTOR OF PHILOSOPHY

IRENE HUGGINS ©2003

Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilm Inc. to publish an abstract of this thesis/practicum.

The author reserves other publication rights, and neither this thesis/practicum nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

ACKNOWLEDGEMENTS

I wish to thank Dr Beverley Zakaluk, Dr. Stan Straw and Dr. Charles Mayenga for their support and guidance as supervisors of this research. I also want to thank Dr. Barbara Watson for providing the initial focus and Dr. Marie Clay for keeping me apprised of international research in this field.

I am also grateful to the superintendents, principals, and classroom teachers who volunteered to be part of this study. I would particularly like to thank the parents and caregivers who gave permission for their children to participate.

Two groups require special thanks. The children who gave their consent to the assessments required in this study and the Reading Recovery Trainer, Teacher Leaders and Teacher Leaders in training who completed the assessment. Without their support, this research could not have taken place.

My friends, colleagues and family have provided support, encouragement and help throughout.

Finally, I would like to thank my husband who has provided the ongoing support and encouragement over many years of academic work.

ABSTRACT

The aim of this study was to develop a set of standard scores for the six tasks in *An Observation Survey of Early Literacy Achievement* (Clay, 1993, 2002) based on the performance of students in publicly funded schools in Manitoba. Previous studies have focused on the development of norms for students in New Zealand and the United States and more current research has focused on the development of the tasks and norms for students in Spanish and French. Although many teachers now use the tasks from *An Observation Survey of Early Literacy Achievement* to monitor the early literacy progress of students in Kindergarten and Grade One, there was no well documented study of how these instruments perform in Manitoba. Therefore, it was important to provide a more valid comparison group for Manitoba Grade One students. A Manitoba Word Reading task was also developed, validated, and normed to determine if it should be used in place of the Clay Word Reading task.

Data from Grade One students collected eighteen months after school entry was analyzed to determine if separate stanine scores should be developed for students based on location (urban and rural), age in months, and gender. Although differences were found in student performance based on location and gender, it was concluded that one set of stanines would suffice. Future studies could examine gender differences by analyzing student achievement upon school entry and subsequent rates of progress.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
LIST OF TABLES	x
CHAPTER ONE	
Introduction	1
Definition of Emergent Literacy	3
Why Assess Emergent Literacy Skills	3
Statement of Problem	6
Significance of the Study	7
The Observation Survey Tasks	7
Standard Scores	9
Expected Performance Levels	11
Location	11
Age	11
Gender	11
Reading Recovery™	12
Word Reading Task	13
Research Questions	14
Scope of Study	14
Definition of Terms	16
Summary	18

CHAPTER TWO

Theory and Research	19
Children's Literacy Development	19
Emergent Literacy	19
Acquisition Phase	22
Assessment of Literacy Development	26
Assessment through Observation	26
Aspects of Literacy Development to Monitor	28
What to Assess	29
Text Reading	30
Information Sources Used	31
Word Reading	32
Word Learning	34
Phonological Coding	35
Phonemic Knowledge	36
Orthographic Knowledge	38
Lexical Retrieval	41
Letter Identification	43
Writing	45
Concepts About Print	46
Knowledge Categories and the Observation Survey Tasks	48
Running Records	49
Concepts About Print	51
Letter Identification	52

Word Reading	53
Writing Vocabulary	54
Hearing and Recording Sounds in Words (Dictation)	55
Manitoba Word Reading Task	56
Historical Development of Word Lists	56
Summary	61

CHAPTER THREE

Research Methodology	63
Purpose	63
Phase 1 Development of the Word Reading Task for Manitoba	63
Design Overview for the Word List	63
Identification of Text Materials	64
Word Frequency Counts	65
Comparison to Other Word Reading Tasks	66
Developing the Manitoba Word Reading Task	67
Validation of Manitoba Word Reading Task	69
Results of Validation	70
Order of Administration	70
Comparison of the Lists	72
Analysis of Words	72
Summary	73
Phase Two Norming of Six Tasks	74
Design of Norming Study	74

Sample Size	74
Selection of School Divisions	75
Selection of Schools Within the Divisions	75
Selection of Students	75
Time of Administration	76
Task Administrators	77
Number of Protocols	78
Summary	79

CHAPTER FOUR

Data Analysis	81
Overall Scores	82
Impact of Variables	82
Location	82
Age	84
Gender	86
Summary of Impact of Variables	87
Stanine Scores	87
Developing Stanine Scores	88
Comparing Stanine Scores	89
Location	88
Gender	90
Comparison of Stanine Scores	92
Comparison of Clay and Manitoba Word Reading Tasks	94

Conclusion	95
Standard Scores	95
The Manitoba Word List	95
CHAPTER FIVE	
Results and Discussion	97
Summary of Findings and Discussion	97
Location	98
Age	98
Gender	98
Study Limitations	100
Implications for Future Research	102
Implications for Classroom Instruction	103
Conclusions	104
REFERENCES	107
LIST OF APPENDICES	122
APPENDICES	123

LIST OF TABLES

Table 3.1	Frequency of Use of Text Series in Manitoba Schools Per Grade	65
Table 3.2	Location of Words on List Based on Frequency Count	69
Table 3.3	Pearson Correlations for List A Based on Order of Administration	71
Table 3.4	Pearson Correlations for List B Based on Order of Administration	71
Table 3.5	Pearson Correlations for List C Based on Order of Administration	71
Table 3.6	Pearson Correlations of Word Lists	72
Table 3.7	Percentage Correct and Point Biserial Correlations for Words on Manitoba Lists	73
Table 3.8	School Division Population and Percentage of Students	76
Table 4.1	Analyses of Variance for Each Task by Location	83
Table 4.2	Number of Students in Each Quarter of Year Proportion of Total Sample	84
Table 4.3	Proportion of Boys and Girls per Quartile	85
Table 4.4	Analyses of Variance for Tasks Based on Age	86
Table 4.5	Analyses of Variance for Each Task by Gender	87
Table 4.6	Stanine Scores for Urban and Rural Students: Running Records of Text Reading	89
Table 4.7	Stanine Scores for Urban and Rural Students: Clay Word Reading Task	90

Table 4.8	Stanine Scores for Urban and Rural Students:	
	Hearing and Recording Sounds in Words	90
Table 4.9	Stanine Scores for Urban and Rural Students:	
	Manitoba Word Reading	90
Table 4.10	Stanine Scores for Boys and Girls: Running Records	
	Of Text Reading	91
Table 4.11	Stanine Scores for Boys and Girls: Writing Vocabulary	91
Table 4.12	Stanine Scores for Boys and Girls: Hearing and Recording	
	Sounds in Words	91
Table 4.13	Stanine Scores for Boys and Girls: Manitoba Word Reading	91
Table 4.14	Comparison of Stanines for Concepts About Print	93
Table 4.15	Comparison of Stanines for Letter Identification	93
Table 4.16	Comparison of Stanines for Writing Vocabulary	94
Table 4.17	Comparison of Stanines for Hearing and Recording	
	Sounds in Words	94
Table 4.18	Comparison of Stanines for Clay Word Reading	95

CHAPTER ONE

Introduction

No person can reach full human stature in our society without competence in reading. Ten years of compulsory failure at school can be crippling enough for the poor reader without the continuing experience of deprivation which [s/he] faces in a society based on the expectation of literacy. The written word influences modern living more deeply every day. Print is persistently increasing its impact on the lives of ordinary people, and in much more complex ways than it influenced a literate minority in the past (Holdaway, 1980, p. 11).

This introduction to Holdaway's book *Independence in Reading* was written more than two decades ago, but still rings true today. Holdaway's message was echoed nearly twenty years later in a report by Snow, Burns, and Griffin (1998), *Preventing Reading Difficulties*. Reading is essential to success in our society. The ability to read is highly valued and important for economic advancement; it influences the social, personal, as well as civic aspects of our lives. Communication through intricate telecommunication systems, which requires the ability to read and write, has become indispensable in the workplace and plays a part in home activities. Learning how to read is not only the key to success in education and communication in the modern world but is also essential in an information economy. Interestingly, the ability to access that key to success may be determined as early as grade one (Snow, et al, 1998).

Research suggests that children who do not read at the end of the grade one fail to achieve in almost every other academic area (Slavin, Karweit, & Wasik, 1993). While

learning to read and write in grade one does not guarantee that students will succeed in future academic learning, lack of success in learning to read and write nearly always guarantees school failure (Adams, 1990). Students who are the lowest achievers in grade one are often the lowest achievers three years later and have established a cycle of failure and low self-esteem that lasts throughout their schooling and impacts on their adult lives (Clay, 1966). In a discussion of the results of a longitudinal study on children learning to read in New Zealand, Clay (1979) reinforced this point by stating:

There is an unbounded optimism among teachers that children who are late in starting will indeed catch up. Given time, something will happen! In particular, there is a belief that the intelligent child who fails to learn to read will catch up to [his/her] classmates once [s/he] has made a start. Do we have any evidence of accelerated progress in late starters? There may be isolated examples which support this hope, but correlations from a follow-up study of 100 children 2 and 3 years after school entry lead me to state rather dogmatically that where a child stood in relation to [his/her] age-mates at the end of [his/her] first year in school was roughly where one would expect to find [him/her] at 7:0 or 8:0 (p.13).

The *National Assessment of Educational Progress Report* (2000) indicates that by grade four, the gap between the highest and lowest performing fourth graders has widened and continues to grow in subsequent years. Therefore, it is important that education systems ensure that most students make an easy transition into formal literacy. Fortunately, teachers now have tools to monitor students' early progress in various areas of literacy development and consequently, are able to adapt instruction.

This brief overview of the research illustrates the importance of success in literacy learning in the early grades of school. It also highlights the need for teachers to

monitor the effectiveness of classroom programs to ensure success for all students. But, as indicated below, what aspects of literacy learning to assess and how to assess them are still controversial topics that are influenced by the teacher's philosophy and understanding regarding how young children become literate.

Definition of Emergent Literacy

Current theory suggests that literacy development begins early in childhood and that students' literacy development "emerges" (Clay, 1966) throughout the preschool years and before children begin formal literacy instruction. *Emergent literacy* is the term used to describe the behaviors of very young children when reading and writing even though they can not read or write in the conventional sense. The term "reading readiness" suggests that there is a point in time when the child is ready to begin formal literacy instruction, whereas the term emergent literacy suggests that the development of literacy knowledge and skill occurs within the child and is a gradual process that takes place over time, beginning at birth. Over the past decade in Manitoba, the concept of *emergent literacy* has gradually replaced the notion of reading readiness. It seems important, then, to monitor students' growing reading and writing achievement.

Why Assess Emergent Literacy Skills

Diverse groups of investigators and theorists have examined the process of literacy acquisition with a variety of research methods, both qualitative and quantitative (Adams, 1990; Clay, 1991; Goodman, 1986; Goswami & Bryant, 1990; Gough, Juel, & Griffith, 1992; Meek, 1992; Smith, 1988; Sulzby & Teale, 1991). Both conventional wisdom and research findings suggest that there are basic competencies that must be

mastered in order for students to become successful readers and writers. Systematic monitoring of students' achievement in these basic competencies throughout the early years is important if the classroom program is to be designed to meet the needs of all students.

Smith and Martin (1997) measured children's emergent literacy knowledge upon preschool entrance and found a strong positive relationship between this knowledge and reading ability five years later. Both this longitudinal study and a study conducted by Scarborough (1991), that examined initial experiences with print in school, offer insights into the difficult process of sorting out the complexities of the impact of early literacy experiences on later literacy achievement. It seems, however, the more literacy knowledge that children bring to school, the greater their success. Nevertheless, a low level of achievement in literacy upon entry does not necessarily lead to poor achievement in subsequent years. Snow et al. (1998) found that three years of high quality classroom instruction could eliminate achievement gaps evident upon school entry.

Recent international debates about the effectiveness of literacy instruction (Ministry of Education, 1999; Snow, et al., 1998) draw attention to the impact of classroom instructional programs on student progress. Researchers working on the sociocognitive and sociocultural dimensions of children's literacy learning have been interested in the role other people, particularly the adults, play in children's literacy acquisition (Cazden, 2001; Duncan, 1999; McNaughton, 1995; Wood, 1998). These researchers challenge the notion that there is a universal pattern of literacy development and the idea that one instructional program fits the needs of all students. Clay (1991) argues that classroom programs impact on the literacy development of young children,

particularly those experiencing difficulties. Effective instruction requires adults to match their interactions to the child's intentions and knowledge (Schrader, 1990) and to phase in and out of more and less directive roles in literacy instruction (Neuman & Rosko, 1993).

Rather than make changes to school entry age requirements or hold students back because they are not developmentally ready, proponents of emergent literacy theory argue that instruction needs to match individual needs (Clay, 2001; Duncan, 1999; McNaughton, 1995). Accordingly, responsibility for planning and monitoring the impact of instructional programs requires not only knowledge related to reading and writing theory, as well as tools to observe achievement, but also understanding of how to support tentative attempts to build effective reading and writing processing systems. Since any information collected through assessment is interpreted based on the theory of reading and writing held by the teacher, and subsequent literacy development activities are also based on this knowledge, monitoring the impact of instructional decisions is critical.

Assessing the effectiveness of beginning reading and writing instruction and student progress is also dependent on knowing what might be expected of students after a particular interval, for example, after six months of formal instruction. Teachers need a frame of reference to interpret and evaluate program effectiveness. Otherwise they may emphasize one area of literacy learning, neglect others, and fail to notice students who require more support. Indicators based on average or expected levels of performance would serve as helpful guidelines for interpreting student progress and identifying individual needs.

Successful literacy learning is more likely to occur if various components are interwoven into literacy instruction. Pearson (2000) suggests that balanced literacy

instruction is not defined by equal amounts of various activities but rather by varying the focus based on the needs of the students. This raises the question of how should teachers monitor the impact of instruction on early literacy development and know when to vary the focus or amount of instruction.

An Observation Survey of Early Literacy Achievement (Clay, 1993a, 2002) offers a multifaceted approach to assessing early literacy learning. But we need guidelines to interpret performance, to determine not only the skills and knowledge that the young child has acquired, but also the child's level of expertise in applying this knowledge. The challenge is not just to acquire knowledge but also to use it to problem-solve while reading and writing (Clay, 1991, 2001, 2002; Paris, Wasik & Turner, 1991). Historically, tools to monitor the development of these problem-solving abilities have been difficult to administer and interpret. Assessments have often been delayed until tests could be administered by clinicians or other specialists. However, given the importance of the task, it is critical that classroom teachers assume responsibility for monitoring children's literacy development and have the "know how" to capitalize on the assessment information to inform future instruction. In order to accomplish this, Manitoban, not New Zealand norms are required.

Statement of the Problem

Given the level of responsibility placed upon early years teachers, a number of Manitoba teachers are asking: Is my instructional program providing opportunities for students to acquire the skill and knowledge necessary to support successful literacy learning as they move through the first year of formal instruction? More specific questions associated with developing successful literacy learning are concerned with not

only what to assess and how, but also the interpretation of assessment results based on expected levels of achievement at various points in the school year. By providing a comprehensive array of early literacy competencies, Clay's *Observation Survey* (1993a, 2002) addresses the need of what and how to assess. This study addresses the question of expected levels by providing performance data from Manitoba students.

Significance of the Study

The purpose of this study is to develop norms for the tasks in *An Observation Survey of Early Literacy Achievement* (Clay, 1993a, 2002) in support of teachers' efforts to monitor the progress of individual students and the effectiveness of their classroom literacy program. These tasks, when used in addition to classroom observations, provide evidence of on-going student progress through the early years. The survey includes assessment techniques to measure progress on a range of competencies necessary for success in literacy learning.

The Observation Survey Tasks

An Observation Survey of Early Literacy Achievement (Clay, 1993a, 2002) is a portfolio of six standard observation procedures for systematically recording the reading and writing behaviors of children aged five to seven years. The *Observation Survey* is used by classroom teachers, Reading Recovery™ teachers, and researchers to determine the literacy progress of young children through the first two years of schooling. Clay (1966) developed the *Observation Survey* over thirty years ago when she was observing students' literacy development during their first year of instruction. As a result of her observations, she perceived a need for a variety of assessment tasks that could be used to

monitor student progress and to provide specific information that would be helpful for teachers in planning classroom instruction, both group and individual. The six tasks included in the *Observation Survey* are based on the notion that "children have to extend their knowledge along each of several different dimensions of learning as they approach formal literacy instruction" (Clay, 1993a, p. 6). The six tasks in the *Observation Survey* include (1) Taking Records of Reading Continuous Texts (Running Records), (2) Concepts About Print, (3) Letter Identification, (4) Word Reading, (5) Writing Vocabulary, and (6) Hearing and Recording Sounds in Words (Dictation).

Although the *Observation Survey* does not include a task for measuring oral language, the absence of an oral language measure does not minimize the role of language in literacy development (Clay, 2002). Language is integral to the development of reading and writing. Having a large vocabulary and flexibility with words, being able to communicate in acceptable grammatical forms and sustain a conversation, as well as being motivated to use language in problem-solving contexts are important language competencies that constitute the foundations of literacy (Snow, et al., 1998). The *Observation Survey* is made up of formal tasks to monitor other aspects of literacy development, including print knowledge, but also provides an opportunity for the teacher to observe the child's use of language informally.

The purpose of the survey is to provide a systematic, standard, multidimensional, and reliable tool for classroom teachers to use when monitoring the literacy learning of their students. Performance on these tasks provides information regarding the child's current literacy knowledge repertoire and also how the child is using this knowledge to facilitate reading and writing. Clay (2002) suggests that to assess literacy knowledge, teachers must select methods that capture: process, repertoire, strategic activity, and

problem-solving (p. 144). Information gathered by administering the six tasks thus provides classroom teachers with evidence of competency in each of these critical areas, information that informs subsequent instruction. To appraise the level of competency, Manitoba teachers must also have evidence of expected performance levels on each of the various tasks.

Standard Scores

The *Observation Survey* tasks are designed for standard administration and are accompanied by a standard scoring system that supports the reliability of comparisons made between tasks and individuals (Clay, 1993a, p. 7). The standard scores used in the *Observation Survey* are stanine scores so that comparisons between tasks and students are possible (Janda, 1998). To develop stanine scores for a particular test, the test must be administered to a large group of children of the same age or grade to estimate how other children in that age or grade will score. The sample on which the tasks are normed should represent, as well as possible, the population that will take the survey in the future.

Currently, there are two sets of stanine scores included in the *Observation Survey*, one set from a norming study conducted in New Zealand in 2000 with students ranging in age from 5.0 to 7.0, and a second set developed from a study of first graders in Columbus, Ohio in 1990. It seems inappropriate to use New Zealand and American students as a basis for interpreting the performance of Canadian students. These *Observation Survey* tasks are sensitive to classroom instruction and the stanine scores reflect a particular emphasis within the instructional program (Clay, 1991). New Zealand begins formal literacy instruction at age five in full day programs. It may not be helpful

to use the scores of New Zealand students when examining and reporting the progress of Manitoba students. Stanines developed in the United States in 1990-91 also reflect the classroom programs in place in Columbus, Ohio and similarly may not be reflective of the progress of students in Manitoba schools. The computation of Manitoba stanine scores developed in this study will thus provide a more valid basis for comparing the early reading and writing development of our grade one students in 2002-2003.

While the stanines generated in this study provide Manitoba teachers with an appropriate comparison group, these comparisons should be used to determine students' relative strengths and identify areas that need further instruction. The stanine scores do not indicate expected performance levels; they only indicate where students' scores are relative to that of other Manitoba students who have completed the same tasks.

Young children enter school with specific and divergent literacy experiences influenced by family and preschool opportunities. However, classroom factors also influence the rates of change in academic skills over time. Teacher beliefs about the impact of location, age, and gender on student achievement influence not only teacher expectations for achievement but also impact on the learning opportunities they provide. This study examines student performance based on location (urban and rural [farming and northern communities]), age in months, and gender in order to establish whether separate sets of norms should be developed along these dimensions.

Expected Performance Levels

Location. Children differ in the experiences they have had with print before entry to Kindergarten as a result of opportunities provided in homes and communities (Bronfenbrenner & Morris, 1998) or preschool centers (Hamre & Pianta, 2001).

Although separating out factors associated with family income and residence in poor neighbourhoods is difficult (Snow, et al., 1998), location has an impact on preschool experiences and, therefore, places some groups of children at risk. Their emergent literacy skills may not match that of peers from more affluent communities. Classroom teachers must be prepared for diversity in knowledge and skill levels and provide the requisite experiences that children may lack.

Age. Debates concerning the effects of age on achievement in early grades have been on-going. Reviews of the literature on age of entry, extra-year programs, and kindergarten retention (Meisels, 1992; Shepard & Smith, 1988) have consistently shown that academic differences between younger and older kindergarten students are eliminated by grade three. Bickel, Zigmond, and Strayhorn (1991) have noted "that any real or imagined advantage for older students may be specific to curricula that has not been sufficiently individualized or are otherwise unsuitable for young children" (p. 107). The current study determined if separate stanines should be developed for students in grade one based upon four age groups: January to March, April to June, July to September, and October to December.

Gender. The relationship between gender and achievement was also examined to determine if separate norms should be developed for boys and for girls. There is evidence that boys, on average, achieve at significantly lower levels than girls in all areas of literacy development throughout the primary years (Rowe, 2000). Findings remain inconsistent however. Some studies report significant differences in reading achievement based on gender, with girls outperforming boys on several measures across grade levels (Gates, 1961; Rowe, 2000; Stroud & Lindquist, 1942). Other studies indicate low correlations between reading achievement and gender, with no significant

differences beyond the first two years of school (Lummis & Stevenson, 1990; MacFarlane, 2001).

Since findings are not consistent across studies, the question as to whether or not there are significant gender (boys and girls) differences in grade one literacy achievement in Manitoba schools is relevant. This study therefore also examined differences in performance based on gender to determine whether or not separate stanines should be developed for boys and girls.

The norms developed in this study can be used by classroom teachers to monitor the literacy development of young children. When used with the whole class, norms can identify strengths and limitations in the instructional program (Hill & Crevola, 1997; Pinnell & Fountas, 1996) and provide teachers with evidence of individual student progress that can be used for instructional planning and decision-making. The norms should identify students who require additional classroom support. The six tasks from the *Observation Survey* are used to determine how to select the lowest-achieving grade one students for one-to-one instruction in the Reading Recovery™ program.

Reading Recovery™

Reading Recovery™ is an early literacy intervention program that is designed to reduce the number of students who are experiencing difficulty with reading and writing. It is a second-chance, prevention program delivered to the lowest-achieving students in grade one (Clay, 1993b, 2001). These at-risk students have, in addition to classroom instruction, thirty-minutes of individual instruction daily directed at bringing students to achievement levels that are similar to that of their more average classmates within twelve to twenty weeks. These same tasks are also used at the end of the program to confirm the

Reading Recovery™ teacher's decision that the student has acquired the skills and knowledge necessary to discontinue the individualized program and is able to continue building literacy skills and strategies within the regular classroom.

A critical role of the *Observation Survey* is that it provides a system for monitoring the progress of students participating in the Reading Recovery™ program. Pearson (2000) contends that its use provides a great deal of indirect evidence for validity and reliability. Perhaps the best evidence for the construct validity of the tasks is that students who do well experience much greater success in transferring their skills to classroom reading and writing. According to Pearson, it is the convergence of multiple indices of reading and writing success, with all important dimensions of the literacy domain being assessed, that is part of the evidence supporting its construct validity. Manitoba teachers, however, question the validity of Clay's Word Reading task.

Word Reading Task

The Word Reading task in the *Observation Survey* was developed by Clay in 1966 using the 48 most frequently-occurring words found in the 12 books that were used to teach New Zealand children to read (Clay, 1993a, 2002). These books are no longer in print and Manitoba teachers question whether the words found on the Clay Word list are the most frequently-occurring words found in the books currently read in our classrooms. Significant differences in the word lists would suggest that a new word reading task should be developed for Manitoba schools.

This study also included the development of a word reading task based on the text materials currently used in Manitoba kindergarten and grade one classrooms. The word list was compared to Clay's three existing lists and subsequently validated and

normed to determine if the Manitoba list should be used in place of the Clay Word Reading task.

Research Questions

The purpose of the study was to develop a set of stanines scores based on the performance of Manitoba students on the *Observation Survey* tasks to be used as guidelines for monitoring and interpreting student performance. The study also focused on two other related questions:

1. Are there significant differences in the mean performance of students on the *Observation Survey* tasks based on location (i.e., urban and rural [farming and northern communities]), age groups (i.e., between January to March, April to June, July to September, and October to December), and gender (i.e., between boys and girls)?
2. Are there significant differences in the mean performance of students on the Manitoba and the Clay Word Reading tasks?

Scope of the Study

This study develops Manitoba norms for Clay's *Observation Survey* by analyzing the February test scores of a representative group of grade one students. The *Survey* tasks were administered by 15 Reading Recovery™ trained personnel to 400 children from 50 schools across Manitoba. Schools that volunteered to participate obtained parental consent for the students. From this list, students were randomly selected for assessment. The sample comprised an equal number of boys (200) and girls (200), as well as an equal number of urban and rural [farming and northern communities]) students.

After the data had been collected for each student on all of the seven measures, which included the six tasks in the *Observation Survey* and the Manitoba Word Reading task, the mean scores were analyzed to determine if it was necessary to develop separate sets of stanines for students based on location (urban and rural [farming and northern communities]), age (January to March, April to June, July to September, and October to December), and gender (boys and girls). The raw scores were then transformed into stanines.

A second purpose of this study was to develop a word list based on the reading selections most commonly used in Manitoba Kindergarten and grade one classrooms. A survey of Manitoba elementary schools identified the *PM Storybooks* as the most commonly-used textual material. Once a frequency count of all the words in the books in the series ranging from Magenta to Orange levels was completed, the 61 most frequently occurring words were placed in three lists of equal number according to order of frequency, with the most frequently occurring word used as the practice word. As a preliminary step, the lists were administered to 150 urban and rural (farming and northern communities) students and the results analyzed to determine whether the lists were of equal difficulty. The data for individual words were also analyzed to establish if words should be removed or moved to another list. With the equivalency of the Manitoba Word Reading task established, individual lists were administered to the norming sample and the mean scores of the Clay and Manitoba Word Reading tasks compared to determine if the Manitoba Word Reading list should be used in place of Clay's list.

Definition of Terms

Terms used in the study have been categorized into three broad groupings of:
literacy learning, decoding, and assessment.

Literacy Learning	
Emergent Literacy	The development of the association of print with meaning that begins early in a child's life and continues until the child reaches the stage of conventional reading and writing (Harris & Hodges, 1995, p. 70).
Conventional Literacy	Conventional readers and writers read and write in ways that most people in our literate society recognize as 'really' reading and writing. For example, they use a variety of reading strategies, know hundreds of sight words, read texts written in a variety of structures, are aware of audience, monitor their own performances as writers and readers, and spell conventionally (McGee & Richgels, 1996, p. 30).
Reading Readiness	Readiness refers to the point at which students are able to profit from beginning reading instruction (Harris & Hodges, 1995, p. 212).
Concepts About Print	Relates to a learner's understanding of the forms, functions, and uses of various aspects of print including book handling skills, directional behavior, language concepts such as letter, word, first and last, as well as uses of punctuation (Clay, 2002).

Decoding	
Phonological Awareness	Describes awareness of the constituent sounds of words in learning to read and spell (Harris & Hodges, 1995, p. 187).
Phonemic Awareness	Is the insight that every spoken word can be conceived as a sequence of phonemes (Snow, et. al., 1998, p. 52).
Orthographic Knowledge	Is the nature and use of symbols in a writing system (Harris & Hodges, 1995, p. 185).
Lexical Knowledge	Refers to understanding of word meanings as contrasted to understanding grammar or syntax (Harris & Hodges, 1995, p. 138).
Strategies	Includes the mental activities initiated by the child to problem-solve the puzzle of getting the messages from a text, or putting messages into texts (Clay, 2002, p. 34). The strategies include the use of semantic, syntactic, and visual cues and cross-checking to ensure that all three are used together.
Assessment Tools	
Running Records	A technique used to code oral reading behaviors of students systematically (Clay, 2002).
Concepts About Print	The Concepts About Print assessment in the Observation Survey contains 24 items that measure what children understand about the conventions of written language (Clay,

Assessment Tools	
<hr/>	
	2002).
Letter Identification	The identification of all the lower and upper case letters as well as the literary <i>g</i> and <i>a</i> on a page of randomly arranged letters (Clay, 2002).
Word Reading	The ability to identify words already learned in isolation (Harris & Hodges, 1995, p. 283).
Writing Vocabulary	Task in which the child is asked to write down all the words s/he knows how to write (Clay, 2002).
Hearing and Recording	In this task, the child is asked to record a dictated sentence.
Sounds in Words	The child is given credit for representing the sounds by the appropriate letters (Clay, 2002).

Summary

The purpose of this study was to develop a set of stanine scores that can be used by Manitoba teachers to monitor the literacy progress of grade one students. Chapter 1 delineated the nature of the problem and defined the research questions, while Chapter 2 reviews the related literature. Research procedures are outlined in depth in Chapter 3, including a description of the sampling population, data gathering, and statistical analysis. Chapter 4 presents the findings from the data analyses, while Chapter 5 summarizes findings and makes recommendations both for classroom practice and further research.

CHAPTER TWO

Theory and Research

Children's Literacy Development

This chapter examines the research related to the development and assessment of early reading and writing skills. In the first part of the chapter, the concept of emergent literacy is discussed and areas for assessment are identified. In the second part of the chapter, the tasks in *An Observation Survey of Early Literacy Achievement* (Clay, 1993a, 2002) are described and the research related to the development of word reading tasks is reviewed.

Emergent Literacy

To parents and many educators it may seem unnecessary to debate the issue of whether a young child's early encounters with literacy learning should be called prereading, reading readiness, or emergent literacy. However, the adoption of the term *emergent literacy* (Clay, 1966) signaled a break with the concept of readiness that suggested young children needed to be taught a series of prerequisite skills prior to learning to read and write and that writing should be delayed until after children were reading. The term *emergent literacy* implies that literacy learning is developmental and takes into account all the literacy behaviors exhibited by the child before s/he becomes literate in the conventional sense.

While discontinuities between fluent readers and young children's literacy behavior were recognized, historically, few theories previously addressed how young children actually become literate. In fact, most reading theories only attempted to explain

successful, conventional reading behaviors. In 1966, Clay explored the experiences that lead to the development of conventional reading ability as well as behaviors that signal growing control. Teale and Sulzby (1986) also examined the similarities between early literacy behaviors and conventional reading and writing, stating that "these behaviors are not pre-anything. It is not reasonable to point to a time in a child's life when literacy begins. Rather... we see children in the process of becoming literate as the term emergent indicates" (p xix).

The issue for teachers is that:

Children's movement into reading is not marked by a clear boundary between readers and nonreaders. Very young children may know where there is something to read but be unable to read it. Somewhat older children may be able to read isolated words in context, but not in isolation. Skilled older readers may be able to read isolated words by storing partial letter-sound associations in memory but they may not be able to read isolated words by decoding the letters into sounds. Which are we to consider readers and which are nonreaders? Reading acquisition is better conceptualized as a developmental continuum rather than the all-or-nothing phenomenon (Mason & Allen, 1986, p. 18).

Since the publication of Clay's work, research has expanded upon our understanding of emergent literacy. It is now accepted that children's literacy development begins long before formal instruction in grade one (Clay, 1991; Snow et al., 1998; Teale & Sulzby, 1986). This literacy development is supported by caring adults and opportunities to engage in reading and writing-like activities (Clay, 1991; Durkin, 1966; Teale & Sulzby, 1986). It proceeds along a continuum in a variety of ways and at different ages (Strickland & Morrow, 1989). It is influenced by the child's developing

understanding of literacy concepts and the efforts of caregivers to promote literacy development (McNaughton, 1995).

The type of activity introduced in literacy learning is important in defining emergent literacy development. McGee and Purcell-Gates (1997) would exclude studies of emergent literacy development during which formal instruction and/or interventions impact on student performance. These authors thus exclude any study that included students who received systematic training during the preschool years (in phonemic awareness, for example) as a study of emergent literacy behavior.

Yaden, Rowe, and McGillivray, (1999) suggested that the essential tenets of an emergent literacy perspective include:

1. an optimistic view of children's ability to learn on a forward trajectory from unconventional to conventional literacy;
2. a positive view of children as constructors of their own literacy knowledge; and
3. a belief that emergent literacy occurs informally in holistic, meaning-driven reading and writing events (p. 33).

Emergent literacy skills can be divided into at least six areas (Sayeski, Burgess, Pianta, & Lloyd, 2001). These include (a) alphabet knowledge, (b) phonological awareness, (c) expressive and receptive language, (d) verbal memory, (e) concepts about print, and (f) early writing. Sulzby's (1985) categories of Storybook Reading could be added to include information regarding emergent reading levels. What is not made clear by these authorities is the level of expertise required of the child in each of these areas in order to be successful in the learning of conventional reading and writing.

Clay (2002) suggests that a number of skills are necessary, but not sufficient, for literacy acquisition. Consequently, success in learning individual sets of emergent skills should not be viewed as a predictor of future success in becoming literate but rather as information regarding current levels of achievement in each area. Teachers should not be trying to determine literacy predictors but rather adapt their instructional programming to interfere with or thwart the predictions.

Most research has tended to highlight specific areas of literacy development and endeavored to establish a causal relationship (Clay, 1991). However, the importance of specific skill areas may not be critical to all children in all circumstances. According to Clay (1998) the developmental path of different children may vary in significant ways, even though these children eventually achieve the same outcome, success in reading and writing. Evidence for the theory of different paths for different children is demonstrated by the fact that most children learn to read and write in a variety of classroom programs. What is clear, however, is that teachers need an understanding of how children become literate and the skills and knowledge that support this development so that they can observe and monitor student progress and provide the required instructional assistance.

Acquisition Phase

From the first months of life, children's experiences with language development and literacy build the foundations for later success in literacy learning (Snow, et al., 1998). From three to four years, children demonstrate rapid growth in knowledge about literacy. They begin to engage in reading behaviors, such as telling a story from pictures in a favorite book (Holdaway, 1979), and writing, using not only scribbles, but also letter-like forms and random strings of letters (Clay, 1975).

Around age five, children enter school and begin to engage in more formal literacy activities guided by the teacher. They continue to expand their knowledge and skills in all aspects of literacy. A few students have already learned to read and write based on their preschool experiences (Durkin, 1966). At some point during kindergarten, depending on the philosophy of the school or the classroom teacher, many children begin to move into conventional literacy.

In grade one, when most children are six years of age, there is a deliberate effort by the classroom teacher to move all students into conventional literacy. Grade one signals the beginning of more formal instruction in literacy learning and an expectation on the part of both teachers and parents that children will become literate that year. The ease with which the students move successfully into conventional literacy is determined by the skills and knowledge that were learned in the emergent phase and the quality of the instructional program. It is important to note that students who have not acquired literacy skills and knowledge in the preschool years will still become literate with the support of skilled, knowledgeable teachers (Snow, et al., 1998).

Emergent and conventional literacy are not "discrete stages but a continuum of learning that varies with the complexity of each individual's development" (Fountas & Pinnell, 1996, p. 177). As the students are moving into conventional literacy, their developing literacy growth is referred to in terms of *early reader*, *transitional reader*, and *fluent reader* (Duncan, 1999) or *emergent*, *transitional*, and *self-extending* (Fountas & Pinnell, 1996). Most students at the grade one level are, or will become, early readers.

Learning to read is not just about reading the words in the text but also gaining understanding from the text. However, early readers and writers have additional challenges to face as they begin to understand the complexities of the printed message.

Characteristics of early or emergent readers include: (1) knowing how to use early reading strategies and read appropriately-selected, introduced texts independently (Fountas & Pinnell, 1996); (2) attending to print and applying knowledge of phoneme identity to letters (Snow, et al., 1998); (3) starting to use beginning and ending letters to decode unfamiliar words (Clay, 1991); (4) attending to more than one source of information while reading (Clay, 1991); (5) building a small bank of known words in both reading and writing (Clay, 1993b); and (6) using letter-sound associations and patterns to spell and read more words independently (Goswami, 1988, 1998; Moustafa, 1997).

Transitional readers, which would typically include most children at the grade two level, can do everything the early reader can do, but with greater independence and speed. They are able to read unfamiliar, but still carefully selected, texts independently (Fountas & Pinnell, 1996), both fiction and nonfiction. They use information from text flexibly and are building a bank of known words to use when generating new words in both reading or in writing.

Fluent or self-extending readers are able to use all sources of information to read a variety of texts and are more independent problem-solvers (Clay, 2002). They continue to expand their knowledge in reading and writing (Clay, 1991, 2001). They can identify specific words or language structures that are causing comprehension difficulties and can problem-solve or seek assistance (Clay, 2001). The system has become self-extending (Clay, 1991) or self-improving (Clay, 1979).

Because reading and writing are thinking processes (Allington & Cunningham, 1986), emergent or early reading must be viewed within the context of the child's developing cognitive skills or strategic activity (Clay, 2002). Paris, Wasik, and Turner,

(1991), explain why it is important for teachers to monitor the development of reading from a strategic perspective.

First, strategies allow readers to elaborate, organize, and evaluate information derived from text. Second, the acquisition of reading strategies coincides and overlaps with the development during childhood of multiple cognitive strategies to enhance attention, memory, communication and learning. Third, strategies are controllable by readers; they are personal cognitive tools that can be used selectively and flexibly. Fourth, strategic reading reflects metacognition and motivation because readers need to have both the knowledge and disposition to use strategies. Fifth, strategies that foster reading and thinking can be taught directly by teachers. And sixth, strategic reading can enhance learning throughout the curriculum (p. 609).

Key to understanding literacy development is the idea that literacy knowledge is partially discovered; young children construct their own ideas about literacy as they actively participate in literacy activities. However, literacy is also learned through behaviors that are modeled and scaffolded by adults (Wood, 1998) on an ongoing basis. Adults play a critical role in encouraging children to change and refine their knowledge and skills to match more conventional reading and writing. Critical to this support is the teacher's understanding of the need to monitor student progress throughout the entire acquisition phase.

Assessment of Literacy Development

Assessment is becoming increasingly important for teachers in primary grades because administrators and parents want more detailed information about children's early literacy progress. More importantly, assessment is a vital part of teaching because instruction must be designed according to children's knowledge, skills, and interests. Well-designed assessment tools provide teachers with information to monitor the impact of their instructional programs. Effective instruction challenges children because it falls into "the zone of proximal development" (McNaughton, 1995; Vygotsky, 1978; Wood, 1998) or the cutting edge of independent abilities. Effective instruction is planned from assessment data which is used to group children for instruction, to select appropriate materials, and, during teaching interactions, to inform teaching decisions that result in providing the most powerful learning opportunities for each child (Fountas & Pinnell, 1996). Fountas and Pinnell (1996) refer to the use of assessment data to achieve this kind of planning and decision-making as "data driven" instruction (p. 73).

Assessment Through Observation

Educational researchers have long depended on their skills as observers: they have watched individual students and groups of students engaged in a wide range of authentic tasks and have drawn inferences about students' behaviors (Clay, 1966; Duncan, 1999; Sulzby & Teale, 1991). Teachers who assess young children using observational techniques promote learning because the assessment can be based on real tasks, can be individualized, and can be used for educational planning.

Most educators of young children feel that skill in observing and interpreting children's behavior is an essential part of being a good teacher. While the use of

standardized, norm-referenced tests, even for young children, is on the increase, this kind of measurement cannot give the teacher the information needed to make decisions about individual students close to the onset of instruction (Clay, 2002). Standardized tests are useful after students have made considerable progress in reading and writing. However, students who are making slow progress require systematic observations by the teacher in a variety of skill areas which will identify precisely what needs to be taught.

Observation requires teachers to be respectful of the intellectual efforts children bring to reading and writing tasks; it also requires teachers to be confident about their own abilities to interpret the behaviors they see. Teachers can observe students to determine areas of strengths and weaknesses, patterns of behavior, and the cognitive strategies they apply when engaged in learning tasks. From these observations, teachers can determine which students need additional help, or specialized instruction, and how the classroom program might be changed to encourage more and better learning opportunities for all children.

The tasks in the *Observation Survey* are designed to allow the teacher to examine children's behaviors during literacy activities more carefully. However, analysis of the information obtained from the assessments is difficult and may lead to the inappropriate use of early literacy assessments without full appreciation of the developmental properties of the constructs underlying them, or the possible differences suggested by alternative versions of the same index (Pearson, 2000). Teachers may use the *Observation Survey* tasks to obtain scores, however, it is the qualitative analysis that will reveal the student's existing repertoire, how the responses were generated, and whether the information is being related from one competency to another (Clay, 2002; Fountas & Pinnell, 1996). This type of analysis can only be completed if the teacher has an

understanding of the theoretical constructs that are the basis for the *Observation Survey* tasks.

Aspects of Literacy Development to Monitor

The areas of literacy knowledge that the teacher selects to assess will be directed by the theory held either by the teacher or the school regarding early literacy acquisition. Constance Weaver (1994) suggested that there are three main views of what it means to learn to read, and these views highlight the areas that the theorists hold to be most important in beginning the process.

1. View #1: Learning to read means learning how to pronounce words.
2. View #2: Learning to read means learning to identify words and get their meaning.
3. View #3: Learning to read means learning to bring meaning to text in order to get meaning from it (p. 15).

The first view presents a simple view of reading and does not explain the complex cognitive processes that young children must develop if they are to be successful in learning to read (Clay, 2001). The second view suggests that children can construct an understanding of a story if they know the identity and meaning of the words they are reading. However, this does not explain the complexity of developing an understanding of the story that goes beyond the words in the story and requires an understanding of how authors construct texts and use a variety of techniques to tell a story. It is only the third view that begins to explore the actions of the constructive child in developing an understanding of how print (text) works as s/he begins to work with words and anticipate the author's message. To read a selection, the child must go beyond

the author's words to negotiate a possible meaning or message and therefore, it is important for teachers to observe students reading. However, it is also important that the teacher ensure that students have gained sufficient knowledge in a variety of print related tasks.

What to Assess

In the earliest stages of reading development, it is important to understand how students process words and use other text features in the context of reading and writing. The tasks from the *Observation Survey* provide this information. The Records of Reading Continuous Text (Running Records), Concepts About Print, Letter Identification, Word Reading, Writing Vocabulary, and Hearing and Recording Sounds in Words tasks all provide clues to the child's developing understandings of these text features. Clay (1998) suggests that students will take multiple and varied pathways to literacy development, thereby explaining the success of students in a diversity of instructional programs. More importantly, administering one or two assessment measures yields results that do not effectively represent the complexity and diversity of literacy development. Clay (1993a, 2002) recommends that all the tasks in the *Observation Survey* be administered when making critical decisions concerning students' educational programming, with only the Records of Reading Continuous Text (Running Records) serving as a stand alone task. Running Records of real reading provide information related to the child's integrated application of all literacy skills and strategies.

Text Reading

Arguably, performance on text reading is the most significant of the criteria that can be used to judge early reading progress. There are several reasons for this. Current models of reading are constructive (Clay, 1991, 2002; Rumelhart, 1994; Singer, 1994), requiring the consideration of multiple variables, as well as the integration of multiple sources of information. In order to assess this integration, observation of text reading is necessary. Clay (1972, 1991, 1993a, 2002) reported that Running Records would provide teachers with important insights into how children process text in terms of the information systems they use, and further, as important tools for documenting the reading behaviors of children. This record of text reading can be analyzed for three kinds of information: (1) text appropriateness for the particular child; (2) strategies employed (monitoring and searching); and (3) information (meaning, syntactic, visual) used by the reader when making an error, reading at too difficult a level, or self-correcting (Clay, 1993a, 1993b, 1991, 2002).

As a tool for observing a young reader's oral reading behaviors, Clay (1966, 1993a, 2002) developed the procedures for obtaining a Running Record described in the *Observation Survey*. While the student reads aloud, the administrator simply takes a sheet of paper and records the students' reading behaviors in a controlled and systematic way. Advantages associated with taking a Running Record include its flexibility for use at any time on any book. Coding the child's reading does not appear to be a testing situation (Johnson, 1992) since it is an activity that the students engage in during the course of regular classroom instruction. Only the teacher's role changes.

It is important that students read stories that are at their instructional level; in other words, students should be able to recognize most of the words (Beck & Juel,

1995). When students read texts at their level, they have increased opportunities to read for meaning (Juel & Roper-Schneider, 1985). With successful word recognition, students do not have to concentrate on the words and can pay attention to meaning-making (Chall, 1996). Running Records provide the opportunity for the teacher to observe the student's reading and to determine that the text is at an appropriate instructional level and therefore assess that comprehension follows. Desirable progress shows that students are meeting the challenges of increasingly difficult texts.

An important accomplishment in becoming literate is the coordination of decoding and comprehension processes; therefore, Clay (1993a, 2002) suggests that it is important to examine how readers process words and use text features within the meaningful context of a book selection. The Running Records provide evidence of growth in reading achievement, but the analysis of errors provides evidence of the strategies that are being used. Leu (1985) cautioned against the use of oral reading to estimate the kind of linguistic processing going on inside the head of the reader, but Johnson (1992) argued that the analysis of the young child's oral reading can be informative. Applegate, Quinn, and Applegate (2002) suggested that after the completion of the Running Record the teacher should call upon the child to respond to the reading either by evaluating the story, defending an idea, or drawing conclusions. The child's response enhances the use of Running Records by providing the teacher with valuable information concerning the child's reading comprehension.

Information sources used. Braunger and Lewis (1998) suggest that word identification in text relies on four sources of information, representing the knowledge that the readers use as they unlock upcoming words, including: (a) pragmatic cues [social context]; (b) semantic cues [meaning]; (c) syntactic cues [structural or

grammatical]; and (d) graphophonic cues [the alphabetic, orthographic, or phonological]. All of these cueing systems must be operating together for the reader to comprehend the text. Effective readers are active in the reading process, making decisions, and solving problems as they proceed (Clay, 2001, Rumelhart, 1994; Singer, 1994).

When examining oral reading behaviors, the recorder can look for consistencies in how the student searches for information in the text. What appears on the surface to be the simple, word-by-word reading of a story actually involves connecting many and various sources of information. When students are required to problem-solve in this way, they access these various sources and make a series of decisions as they work through the text. To examine only letter and word substitutions, or responses to questions, is to ignore the problems faced by readers in solving the parts within the whole, in sequence, in order to make meaning. This behavior is captured in the Running Record.

Word Reading

Most educators argue that the most valid and important measure of reading achievement is reading comprehension, the product of reading a passage, story or selection. While it seems questionable to focus attention on the recognition of words, especially the reading of words in isolation, the ability to read words quickly, accurately, and effortlessly is critical to skillful reading comprehension (Stanovich, 1986). There are numerous debates about how to view instruction at the word level and how to assess and monitor the development of word identification. As suggested, however, children who recognize words more easily are able to focus more attention on the meaning of the selection (Chall, 1996; Ehri, 1995; Juel, Griffith & Gough, 1986; Samuels, 1994; Samuels, Schermer, & Reinking, 1992). Deficient word recognition skills are usually the

primary factor preventing effective reading comprehension (Perfetti, 1985; Stanovich, 1986). It also appears that the more words an individual knows and recognizes, the easier it is to build on that knowledge (Ehri, 1998; Moustafa, 1997).

The importance of success in early word recognition is supported by many studies (Ehri, 1994; Juel, 1988; Lesgold & Resnwick, 1982; Lundberg, Frost, & Peterson, 1988). Students' progress at the end of the first year of formal instruction in reading is closely linked to the ability to identify words. Nevertheless, it is important to emphasize that decoding does not necessarily lead to text comprehension (Oakhill, Cain, & Yuill, 1998). Comprehension, or the production of meaning, is the major goal of reading (Clay, 1991, 2001; Goodman, 1994; Rumelhart, 1994; Singer, 1994; Smith, 1988). As Stanovich (1985) suggests:

While it is possible for adequate word recognition to be accompanied by poor comprehension abilities, the converse virtually never occurs. It has never been empirically demonstrated, nor is it theoretically expected, that some instructional innovation could result in good reading comprehension without the presence of adequate word recognition (p. 418).

Early learning of the code leads to wider reading both in and out of school (Cunningham & Stanovich, 1993; Juel, 1994; Stanovich, 1986). What Stanovich (1986) terms the Matthew effect, the rich get richer, occurs through wide reading. The more one reads the more adept at word recognition one becomes, and the more the student develops world knowledge, vocabulary concepts, and an understanding about how texts are written. As children go beyond the early or emergent stages of learning to read, word learning accelerates because they have established ways of learning words. Rapid, automatic word recognition is related to competent, fluent reading with understanding

(Biemiller, 1970; Blanchard, 1980; Calfee & Pinontkowsky, 1981; Chall, 1989; Herman, 1985; Juel, 1988; Lesgold, Resnick, & Hammond, 1985; Samuels, 1994; Stanovich, 1985). For fluent reading with understanding, readers need instant recognition of about 95% of the words in the text (Adams, 1990). It is important therefore to understand factors that may interfere with word acquisition.

Word learning. If the ability to decode individual words accounts for most of the variance in first-graders' reading comprehension (Ehri, 1992; Stanovich, 1992), and the ability to name unfamiliar words in the first grade is a good predictor of reading comprehension in the fourth grade (Juel, 1994), how do students develop this knowledge? Morton (1969) proposed a "dual-route" model, that identified two independent pathways to the lexicon for word recognition. A direct phonological decoding route was hypothesized for unfamiliar printed words that were sounded out by applying the grapheme-phoneme correspondence rules. In contrast, a direct lexical route was hypothesized for highly practiced words that were read by the visual processing of the letters in a word, without phonological mediation. Coltheart (1978) and Patterson (1981) proposed similar dual-route models for word recognition.

The dual-route model seems consistent with observations of early readers. Novice readers may recognize a few common words visually as whole patterns (Ehri & Wilce, 1985), while some may sound out new words through applying grapheme-phoneme correspondence rules (Biemeller, 1970; Colheart, 1978). However, Clay (1991) argues that "the dual-process model leaves reading acquisition with all the unsolved problems of how such a dual-process comes into existence" (p. 315). Current researchers argue that the dual-route does not account for the use of analogies in word identification.

They argue that as students build knowledge of words through graphophonic and orthographic information, they also begin to use analogies based on the larger units they know from familiar words (Goswami, 1986; Marsh, Desberg, & Cooper, 1977; Moustafa, 1997). The ensuing discussion therefore, focuses further on the issue of word identification through the application of phonological and orthographic coding skills.

Phonological coding. Stanovich (1985) contends that the greatest contribution of cognitive science to the teaching of reading is the insight that phonological awareness is related to reading. Undeniably, evidence from correlational studies (Stanovich, Cunningham, & Cramer, 1984; Vellutino & Scanlon, 1987), experimental studies (Ball & Blachman, 1991; Bradley & Bryant, 1983) as well as observational studies (Winsor & Pearson, 1992) suggests that phonological awareness is related to success in learning to read.

Phonological awareness is defined:

as sensitivity to any size unit of sound. Thus the ability to generate and recognize rhyming words, to count syllables, to separate the beginning of a word from its ending and to identify each of the phonemes in a word may be each an indication of phonological awareness (Yopp & Yopp, 2000, p. 130).

The phonological system refers to the sounds of the language and as children develop phonological awareness, they become sensitive to these sounds (Ball & Blachman, 1991). They can hold up language and its sounds to conscious observation and analysis. From this analysis, they can complete a variety of tasks such as; phoneme-by-phoneme analysis, syllable analysis, or onset and rime analysis. Stanovich and West (1989) argued that a child with deficient phonological processes will have difficulty

learning to read. Performance on reading tests can be predicted by the young child's performance on phonological awareness measures and the ability to recite nursery rhymes (Bryant, Bradley, MacLean, & Crossland, 1989; Bryant, MacLean, Bradley, & Crossland, 1990).

The relationship between reading achievement and phonological awareness is not clear, however. How does phonological awareness help young children in the early stages of reading development identify new or unfamiliar words? How much awareness does the child have to have in order to be a successful beginning reader? What can be developed as the child learns to read? Stahl and Murray (1994) reported that one measure; initial sound segmentation was necessary, but not sufficient, for children to begin to read. Beach (1992) argued that only simple phonological skills were necessary for children to begin to read and that awareness, which is more complex, results from knowledge gained through actually reading.

Phonemic knowledge. A large body of research documents that one aspect of phonological awareness, namely, phonemic awareness is related to early reading development and to the ability to spell words (Adams, 1990; Blachman, 1984; Bradley & Bryant, 1983; Fox & Routh, 1984; Griffith & Olson, 1991; Lundberg, Frost, & Peterson, 1988; Perfetti, Beck, Ball, & Hughes, 1987; Treiman & Baron, 1981; Vellutino & Scanlon, 1987). "Phonemic awareness is the awareness that the speech stream consists of a sequence of sounds - specifically phonemes, the smallest unit of sound that makes a difference in communication" (Yopp & Yopp, 2000, p.130). This strong predictor of reading achievement (Ball & Blachman, 1991; Bryne & Fielding-Barnsley, 1993; Juel, 1988; Lomax & McGee, 1987; Stahl & Murray, 1994; Stanovich, 1985; Tunmer &

Nesdale, 1985) involves the ability to recognize that words can be broken into phonemes and syllables, and further, to be able to manipulate these elements.

According to Griffith and Olsen (1992), phonemic awareness underpins the ability to use letter-sound correspondences both for unlocking words in reading and spelling words in writing. Phonemic awareness may also be related to whole-word learning (Turner, Herriman, & Nesdale, 1988). However, Richgels and Burns (1989) found that only a portion of inventive spellers, as a group, could read words proficiently. These authors call into question the view (Share, 1999) that any amount of segmentation ability directly facilitates word reading ability. These differences in research findings may, however, be the result of the impact of an instructional intervention that interferes with normal word knowledge acquisition. For example, Richgels (1995) demonstrated that kindergarten students with no formal instruction in phoneme awareness, classified as good inventive spellers, were better able to learn phonetically simplified words than poor inventive spellers, thus strengthening the connection between higher levels of naturally-developing spelling ability and later word learning. In all of this research it is difficult to verify that certain early behaviors predict others, since the behaviors are inextricably linked developmentally in the first place.

Using phonemic knowledge to decode words requires looking at the letters and "recoding" them into their sounds. Finally, this recoded word is matched with the pronunciation of a word that is stored in memory or linked to lexical information (Daneman, 1991). As this sound/letter knowledge grows and some sounds and letters become automatic, fewer exposures to a word are needed for mastery (Velluntino & Denckla, 1991). Sight vocabularies expand rapidly, and fewer errors occur when reading text (Ehri, 1991; Mann, Shankweiler, & Smith, 1984). However, Clay (1993b) suggested

that "some children find it extraordinarily difficult to hear the sounds that make up words" (p. 32). Therefore, some students may require more instructional assistance.

Orthographic knowledge. The ability to relate sounds and letters builds as a child is able both to analyze spoken words further and tie them to orthographic elements. Reading in English requires knowledge in two areas for word recognition, the ability to read words where the letter sound correspondence is regular, and the ability to read irregular or exceptional words. The recognition of irregular words may be more dependent upon orthographic information from the whole word or from clusters of letters that co-occur. Orthographic knowledge, defined as understanding the relationship between spelling patterns and sounds, is an important component of word recognition (Adams, 1990). Ehri (1991) argues that deep knowledge of spelling patterns underlies the students' automatic recognition of words, both regularly and irregularly spelled.

However, in the earliest stage of learning to read, young children have not had the opportunity to build a phonological decoding system; so they often read words by sight (Ehri, 1991) or just begin to use orthographic information. These first words that students can recognize are very supportive to future word learning. As students learn more about word identification strategies, those who know how to recognize letters and can already read a small body of words are better able to transfer their knowledge of letter-sound relationships to decode new words (Clay, 1993b; Ehri & Wilce, 1985). In addition, Vellutino and Denckla (1991) suggest that some words in English require sight recognition because of inconsistencies in their letter-sound associations (i.e., *who, the, you, was*).

Researchers give different names and identify different transitions in classifying the sequential stages they believe children pass through as they learn English orthography. They do, however, give comparable sequential descriptions of each developmental stage (Chall, 1983; Ehri, 1991; Gough & Juel, 1991).

- The logographic (Frith, 1985), selective-association (Juel, 1988) or pre-alphabetic (Ehri, 1991) stage is the point at which the child can read a word only if it is associated with a cue or logo. Sometimes the cue is the shape of the word and letter, the first or last letter, or a chunk of the word (Gough & Juel, 1991). Students at this stage lack rudimentary phonological awareness (Stahl & Murray, 1994).
- Soon the children begin to associate letters and words. Chall (1996) refers to this stage as the "initial reading or decoding stage." In this stage, sometimes referred to as the phonetic cue reading stage (Bear & Barone, 1989), children learn the grapheme or letter and associate those letters with corresponding parts of spoken words. Students must acquire knowledge of the alphabet and the insight that words can be broken down into onsets and rimes to reach the second stage.

As they pass through these developmental stages, in contrast to drill in letter sounds and letter names, it is critical for students to have constant exposure to whole text and a range of reading materials. Students need opportunities to apply their growing knowledge about reading to problem-solve the words they encounter as they read (Chall, 1996; Clay, 1991; Cunningham, 1990; Stanovich, 1980; Stanovich, West, & Freeman, 1981).

- In the last stage, students are able to apply their knowledge of how words are constructed in English and to examine unusual patterns within words to aid in the

word identification. Therefore, it is the interrelationship between the phonological processing abilities and attention to letters and visual features of words (Moustafa, 1997), during real reading, that result in the child "having a flash of insight" (Chall, 1983, p.16), "mastering the cipher" (Gough & Juel, 1991, p. 52), or reaching the orthographic stage.

Thus, it appears that as student familiarity with specific printed words increases, the recognition process becomes very rapid and appears to occur without phonological mediation. If this is true, it might be more accurate to conceive of dual routes to word recognition: either one in which unfamiliar words are identified through phonological recoding and the assembly of word segments, or one in which unfamiliar words are read by linking whole word orthographic and phonological patterns (Patterson, 1981).

When students have achieved the orthographic stage, reading is fluent and automatic (Samuels, 1994) or almost automatic (Clay, 1991). When this system is deficient and word recognition is not achieved, an obligatory, automatic compensatory system is activated (Stanovich, 1986). Stanovich refers to this compensatory system as the context response system which is used similarly by both high and low progress readers when there is a decoding obstacle. Stanovich also contends that the low progress readers need to use this context response system more frequently due to weak phonological and orthographic processing. The end result is less cognitive capacity for comprehension (Stanovich, 1980). Studies by Stanovich (1980, 1986) show variance in the ways that different readers use context cues, with successful readers paying greater attention to the graphic information when decoding, particularly the internal structures of words. These successful readers used context less frequently for word recognition and more often as an aid to comprehending larger units of text.

Successful readers have to know how words work and how to use what they know about words and letters in order to identify novel words (Clay, 1993b) including the use of analogies. To unlock novel words, the student must have developed the "cipher"- the analogical system that has been internalized by the process called "cryptanalysis" (Ehri & Wilce, 1985; Gough & Juel, 1991). "Cryptanalytic intent" is the realization by the young child that there is a system to be mastered. When the cipher has been discovered, children begin to see reading and words in a new way, analyzing and deciphering words more efficiently and with greater precision (Chall, 1983; Gough & Juel, 1991). Both Clay (1991) and Chall (1996) agree that a major breakthrough in reading occurs when a child can let go of attachments to semantic and syntactic substitutions and view reading as a problem-solving process that links all sources of information in an integrated way.

Lexical retrieval. As young readers work with the graphophonic and orthographic information in text, they eventually must begin to link this information with lexical knowledge. Traditionally, reading researchers have looked separately at the acquisition of a reading vocabulary, the young reader's store of knowledge about word forms, and the acquisition of a meaning vocabulary or a lexicon. However, in interactive models of reading (Clay, 2001; Rumelhart, 1994; Singer, 1994), written word forms and meanings are viewed as being represented in larger, interconnected frameworks. These models provide a backdrop for examining the connections between the recognition of the written word and the attributes of word meaning. As young children have opportunities to engage in oral language activities, experience being read to, or begin to read themselves, they add to their lexical knowledge.

McFalls, Schwanenflugel, and Stahl (1994) identified two characteristics that may contribute to the acquisition of reading vocabulary or add to lexical knowledge; concreteness, and abstractness. The term concreteness refers to the "constellation of variables that distinguish abstract words with limited sensory referents from concrete words with more direct sensory referents" (McFalls, et al., 1994, p. 1). According to these authors, some words may be easier for young readers to learn because of their semantic connections. Research on early language acquisition (Gentner, 1982) has consistently indicated that nouns predominate in the child's early reading lexicon, so there may be something conceptually associated with nouns that make them easier to learn than words that fit into other grammatical categories. However, in the New Zealand classrooms in 1966 (Clay, 1993) more abstract words, such as *the*, *said*, and *is*, were identified as the most frequently-occurring words in texts used to teach students to read. In a more current review of the most frequently-occurring words in text material, Duncan (1999) found that the word *the* occurred 1,145 times, *a* occurred 666 times, and *I* occurred 474 times. These words would be categorized as abstract. In a 1994 study, McFalls, and her colleagues found that for both naming words and making lexical decisions, abstract basal reader words were read with less accuracy than concrete words.

It could also be argued from the perspective of Rumelhart's (1994) interactive model of reading that identification of the concrete words in text may be accomplished by using other text information. For example, the word *the* may be read accurately because the child expected the sentence to begin with the word *the*, and successfully applied knowledge of the syntactic information in the text to facilitate decoding. While Clay (1991) contends that children need to be taught to attend carefully to words, by segmenting the words into parts that can be decoded, for example, she also emphasizes

that the decoding that results should be cross-checked with other information (i.e., syntactic and semantic-contextual cues) to determine whether the word, as decoded, fits. The student may be using semantic or syntactic information to identify words with a quick check of the graphophonic information to confirm accuracy. However, Stanovich (1986) would argue that poor readers may be too reliant on this kind of top-down processing and over rely on context to provide word identification clues. High progress readers are less likely to be reliant on context since they are very competent at using graphophonic information. Clay's Word Reading task in the *Observation Survey* provides information regarding the ability to decode words without context clues and also determines skill at reading abstract words in isolation.

Young children learning to read must learn that print is unique; it is not like pictures. The recognition that print is unique is foundational to the development of word recognition (Johnston, Anderson, & Holligan, 1996). Written words are made up of letters that map to speech sounds (Adams, 1990). Further, research has shown that the relationships between print and speech facilitate learning to read. "These findings are buttressed by other information showing that knowledge of word meanings, an understanding that print conveys meaning, phonological awareness, and some understanding of how printed letters code the sounds of language contribute directly to successful reading" (Snow, et al., 1998, p. 320). Underpinning the student's ability to read words is knowledge of the alphabet.

Letter Identification

The alphabet is the foundation of all reading and writing since all words in English are based on this limited set of symbols. To identify these symbols, the young

child must be able to notice the differences that distinguish one letter from another.

Recognizing the letters of the alphabet is a necessary, although not sufficient, factor in mastering the alphabetic principle (Gibson, 1969). Eventually, young children also need to learn the names of the letters since this is the reference used by the teacher to draw attention to or exchange information about these symbols (Pressley, 1998).

It is now recognized that low letter knowledge is a roadblock to learning to read since letter cues are the primary means for identifying words (Adams, 1990). Decades of research have confirmed that knowledge of the alphabet is a strong predictor of a young child's success in reading instruction (Chall, 1996; Langenberg, 2000). Successful beginning readers: have letter knowledge prior to school entry (Durrell, 1958); learn to identify, as distinct entities, all letter forms used in English either alone or as sets through visual perception (Clay, 2002); gradually learn to distinguish new letter forms from known ones (Adams, 1990; Clay, 1991); and become fluid and automatic at letter name knowledge (Adams, 1990). Nonetheless, some research indicates that students who can identify letters may still make very slow progress in learning to read (Stanovich, 1992).

It seems prudent to provide instruction to help children learn letter forms and their associated speech forms so that they can learn how the system works. By increasing opportunities to learn these skills within the context of authentic reading and writing activities, letter and phoneme knowledge should increase. Concurrently, students will be building a repertoire of words that are conventionally spelled. Both reading and writing deserve curricular emphasis.

Writing

By considering the reciprocity between reading and writing, we can help children become better word solvers, not only to accelerate general reading ability (Stanovich, 1991), but also to assist students in becoming better spellers (Juel, Griffith, & Gough, 1985). In writing, students will have to say the word slowly, think about the sequence of sounds, and relate these sounds to sequences of letters. Letters can give abstract phonemes a concrete referent (Hohn & Ehri, 1983). The growth of this knowledge is evident in children's early spellings (Read, 1971).

Even when they are first learning to read, children may notice and use more than simple letter-sound relationships. As they work with print and write messages, they may notice larger letter chunks and easily learn links between groups of letters and groups of phonemes. Thus, reading and spelling instruction that begins with larger units may be more successful than instruction that begins at the phoneme level (Treiman 1992).

While some students may spontaneously learn these word-solving strategies as they encounter words in reading and writing, many students require direct instruction. Readers who are having difficulty going from letter to sound require instruction on letter-sound relationships in addition to explicit instruction on how to apply that knowledge to word solving and to do so while they are reading and writing continuous text (Adams, 1990; Anderson, Hiebert, Scott, & Wilkinson, 1984; Barr & Dreeben, 1983; Ehri, 1991; Gough & Hillinger, 1980; Johnson & Baumann, 1984; Juel, 1991; Mason, 1980).

Although writing contributes to knowledge about how to cipher print, it is also important to note that writing contributes to reading development in other areas such as:

- moving in a left to right direction and controlling serial order,

- drawing on language information stored in memory,
 - making and recognizing visual symbols,
 - using visual and sound information together,
 - holding "the message so far" [*italics added*] in mind,
 - drawing on the known words and structures of language,
 - searching, checking and correcting,
 - and managing to bring these activities together as a message is constructed
- (Clay, 2002, p. 20).

Thus, learning to write not only supports many of the cognitive strategies that are necessary in reading, but also supports the child's ability to work with graphophonic information. The number of phoneme-grapheme connections known by the child and the number of words the child can write independently are measures of this growing control and can supplement the teacher's observations of writing samples. This knowledge is assessed in the Hearing and Recording Sounds in Words and Writing Vocabulary tasks of Clay's *Observation Survey*.

Concepts About Print

Books are constructed according to a set of conventions that can be understood without being able to read (Clay, 1979); however, researchers agree that an understanding about how written language works (Adams, 1990; Clay, 1991, 1993a; Goodman, 1986; Holdaway, 1980; Meek, 1982; Tunmer, et al., 1988) is important for success in learning to read. Concepts about print refers to all the concepts related to the organization of print. A student's global awareness of the forms, functions, and uses of

print provides the basic conceptual framework in which reading and writing is best learned. Consequently, learning the mechanics of print is necessary for becoming a successful reader (Dickinson & Tabors, 1991). In addition, concepts about print may facilitate the acquisition of subsequent reading-related skills (Weir, 1989). Student performance on the Concepts about Print task has been found to predict future reading achievement and to be strongly correlated with other, more traditional measures of reading achievement (Adams, 1990).

Visual attention is hard to observe (Clay, 1991). The Concepts About Print task, however, gives teachers an opportunity to observe this aspect of learning at an important time when links are being encouraged between oral language and reading and writing (Clay, 2002). The ability to recognize word or line changes, as well as letter rearrangement is important for monitoring students' visual attention to print details as they are learning to read.

Directionality, which refers to the way that print is tracked through reading and produced through writing, is an important concept for young children to understand. Students in the early or emergent phase must come to understand that in English you begin at the top of the page of print and move toward the bottom, and start at the left side of the page and move to the right. Students must also learn that when you reach the end of one line of print you must move back to the left-hand side of the page (return sweep) before moving across the line to the right again. Children gradually develop an understanding of directionality during the emergent reading phase and during the transition into early reading.

Word-by-word matching is the ability to match printed words to spoken words. This concept applies in reading and writing as children gradually learn to relate spoken

to written words. However, this is not the same as word recognition, which is the ability to identify words at sight. This word-by-word matching is important because it supports the young child's attempts to match spoken words to the author's words and to notice when a mismatch has occurred, thus supporting the development of an important strategy, self-monitoring.

Concepts of word, letter, and first and last are important terms for early readers to understand if they are to follow the teacher's directions. During classroom instruction, the teacher may verbally draw the child's attention to various aspects of print including letters and words, as well as the first and last letter of a word. Clay (1991) suggests that without an understanding of these concepts, children may not be focused on the print detail that the teacher's instruction demands. What is important is that students understand the instructions given by the teacher. When terms such as *word*, *letter*, and related terminology are difficult, students may, as a consequence, lose out on instruction and fail to learn.

The understanding and use of punctuation and the concept of capital letter are also important for beginning readers. An understanding of punctuation is not only important for conveying the author's intent, but also supports the use of intonation and expression. Torgeson (1986) suggested that these aspects of oral reading may be related to how well students understand what they are reading.

Knowledge Categories and the Observation Tasks

The *Observation Survey* includes six measures for observing the reading and writing behaviors of young children that are closely aligned with important dimensions of literacy learning. These tasks include: (1) Taking Records of Reading Continuous

Text (Running Records), (2) Concepts about Print, (3) Letter Identification, (4) Word Reading, (5) Writing Vocabulary, and (6) Hearing and Recording Sounds in Words.

Each task on the *Observation Survey* is described in the next section.

Running Records

Clay (1966) developed the Running Record as a technique to track students' oral reading behaviors systematically. She developed a series of codes that allow teachers to record the behaviors of the student during oral reading. The use of the recommended coding system allows teachers to determine "how well children are learning to direct their knowledge of letters, sounds and words to understanding the messages in the text" (Clay, 2002, pp. 49-50). The codes help the teacher determine if the reader is using all knowledge systems (meaning, syntax, and visual cues) to unlock words and consequently how to scaffold the child's learning through instructional prompts during the reading of whole text.

Publishers have used a variety of methods for producing a series of leveled books with the lowest level varying on only one dimension. One line will occur on each page, for example, only the object of the sentence being changed. Students are able to read the sentence through the use of picture clues. The text reading measure of the *Observation Survey* assesses the highest gradient of text difficulty that the child can read with 90% accuracy or better. Fountas and Pinnell (1996) refer to identifying the book level as a "quantitative analysis" (p. 90). The implication is that in terms of readability, text levels used in future instruction should match the instructional level. This principle is based on earlier studies (Dunkeld, 1970; Scarborough, Bruns, & Frazier, 1957) that students make gains when instructed at a level which they can succeed, challenged but not frustrated.

In addition to identifying the book level the child can read with instructional support, the teacher gathers information from the observable reading behaviors coded in the Running Records (Clay, 1991). This is a "qualitative analysis" (Fountas & Pinnell, 1996, p. 92). The teacher begins to build a personal theory of what the surface reading behaviors imply about the student's underlying cognitive processes. Every teacher completing a 'qualitative analysis' of the Running Record begins to understand how each reader is using (a) pragmatic, (b) semantic, (c) syntactic, and (d) graphophonic cues. Analysis of Running Records also provides evidence of how the reader is using all of the information together as s/he problem-solves through the text. A student's ability both to notice error through comprehension monitoring and to initiate a search for more information to self-correct the error or miscue also provides valuable information about the student's growing control of complex reading strategies. In addition, the analysis of errors and self-corrections helps teachers determine the types of information the student is using to unlock new or challenging words.

An alternative term used for this qualitative analysis of Running Records is miscue analysis (Goodman, 1985). Goodman contends that readers use three information systems to comprehend: the graphophonic, syntactic, and semantic. In addition to learning about the student's use of the cueing systems for word identification, "the extent to which the miscues result in meaningful text or are self-corrected if they disrupt the meaning gives strong indications of the reader's concern for the ability to make sense of the text" (p. 831).

Concepts about Print

The Concepts about Print task (CAP) in the *Observation Survey* measures groups of behaviors that reveal students' understandings about the conventions of written language. The CAP task refers to all of the concepts related to how print is organized and used in reading and writing tasks. The CAP task consists of twenty-four items, which include:

Print orientation concepts: knowing how to open a book, knowing when the book, pictures, and print have the correct orientation (are upside down or right side up).

Print containing messages concepts: knowing where to read when the page contains both print and pictures.

Directional concepts: knowing where to start reading on a page, how to move left to right, return sweep and move through multiple lines of text.

Relationship between written and oral language concepts: matching speech to print in a one-to-one fashion by pointing with a finger while reading, or being read to, and by recognizing when the line of print, letters, or words is out of order.

Word, letters, capitals, space and punctuation concepts: understanding the basic symbols of English orthography and understanding and using a meta-language for talking about print (Clay, 2001).

The administration of this task includes the selection of one of four possible text samples designed to assess these concepts and the strict adherence to the administration protocol. The maximum score on this task is 24. Concurrent validity was reported as .79 when the criterion used was Word Reading using a sample of 100 urban children aged 6

(Clay, 1966). Predictive validity was reported at .79 after one year of school entry, .73 after two years, and .69 after three years (Clay, 2002). The internal reliability was reported as falling within the range of .73 to .89 using a sample of 56 Texas Kindergarten children (Day, 1979) and .94 using a sample of 40 urban New Zealand children aged 5 to 7 in 1968 (Clay, 1968). Clay (2002) reported a split-half reliability of .95, a Kruder-Richardson reliability of .85, and a test-retest reliability within a range of .73 to .89. While Gilmore (1998) reported Cronbach Alpha reliability of .87, Pinnell and her colleagues (1990) reported a Cronbach Alpha of .78

Clay (2002) also cited another study completed in New Zealand in which the scores of new entrants, on their fifth birthdays, were correlated to parts of the whole score. In this study (Gilmore, 1998), the correlations were high for knowing how reading is carried out ($r = .93$), and for understanding concepts about print ($r = .84$), but as indicated by the correlations, attention to punctuation marks was relatively low ($r = .68$) as was understanding the sequence of letters in words ($r = .33$). It is evident that the ability to analyze words visually, as measured by asking students to identify words that have been written incorrectly ($r = .33$), is difficult for most students (Clay, 2002).

Letter Identification

The letter identification task includes assessing the child on the ability to identify the entire set of upper and lower case letters as well as the literary *g* and *a* on a page of randomly arranged letters. Students are given credit for a correct response if they give either the letter name, the letter sound, or the name of a word that begins with that letter. The maximum score on this task is 54. The internal reliability of the letter identification task was reported as .97 (Clay, 1966, 2002) when the sample consisted of 100 urban

students six years of age. Pinnell, McCarrier, and Button (1990) reported a reliability of .95. Highlighting the link between letter knowledge and word learning, the concurrent validity of the letter identification task with word reading was reported as $r = .85$ using a sample of 100 urban students, aged six (Clay, 1966, 2002). Clay (1979, 1985, 1991, 2002) also reported the predictive validity of the letter identification task to literacy progress as .83 after one year of entry to school, .86 after two years, and .80 after three years.

Word Reading

The Clay Word Reading task is a measure of single word recognition developed to monitor growth in the ability to learn words in reading. The observer selects one of three possible lists that consist of 15 words, and the children are asked to read each word aloud. Students are helped to read the practice word if they cannot read it. The observer notes not only the child's correct and incorrect responses, but also notes attempts at saying the word. This helps the observer assess the student's ability to analyze unfamiliar words. The total possible score on the Clay Word Reading task is 15.

Clay (1966) recognized that it was not possible to assess student knowledge of all the words they were learning to read, so the Word Reading task was designed to assess knowledge of a sampling of words that occurred in a variety of texts. Teachers should not use these lists as a set of reading vocabulary for instructional focus, but rather as an index of the ability to remember words. This ability is critical since Juel and Minden-Cupp (1998) estimate that successful students in the primary grades must recognize about 80,000 words. To achieve this level of growth, learning must be generative. The

three lists provide test-retest opportunities, successive administrations indicating whether a change in the ability to identify words is occurring.

Clay (1966) first administered the word recognition task to 100 students during the course of their first year of instruction and established that the Word Reading task did differentiate among various achievement groups. Clay found, for students in their first year at school, that a list of 15 words systematically sampled from the 45 most frequently-occurring words in the 12 little books used to teach young children to read at that time was a good instrument for grouping students and for identifying low progress readers in their second year. A follow-up of the reading progress of these same students three years later showed that those who had experienced difficulty on the various literacy tasks, including the Word Reading task, were the low progress readers at age eight (Clay, 1966).

The Word Reading task takes about two minutes to administer (Clay, 2002). Clay (2002) reported the predictive validity to reading progress of the Clay Word Reading task at .90 after two years at school and .88 after three years. Clay (1966) reported a Kruder-Richardson reliability of .90, while Pinnell et al. (1990), using Ohio students as subjects, reported a Cronbach Alpha reliability of .92 based on a sample of 107 children.

Writing Vocabulary

The Writing Vocabulary task measures the number of words the students can write correctly in ten minutes. Students are prompted to write all the words they know, and performance is measured by the number of words written correctly. When the student can write more than forty words in the time period provided, then the value of

the score diminishes in terms of informing us about changes in literacy control (Clay, 1993a).

There are two important reasons for students to acquire a core of words that they know how to write in every detail. First, as the frequently used words of the language become automatic, they require less attention and free the writer to attend to other challenges associated with producing written language. The second reason is that known words can be used to analyze words through analogy (Moustafa, 1997). Children see similarities in words and use this information to construct new words when they write (Clay, 1991, 1993a).

Robinson reported the concurrent validity for Writing Vocabulary compared to Word Reading at .82 (Clay, 2002). Clay (1973) established the test-retest reliability as being .97 when 34 urban students aged 5.6 were assessed. Pinnell, et al., (1990) reported test-retest reliability coefficients of $r = .62$.

Hearing and Recording Sounds in Words (Dictation)

Young children's emerging knowledge about letter-sound relationships is revealed in their invented spelling and can be assessed by teachers who ask students to listen to a dictated sentence and then write it. The observer reads one of five alternative sentences samples to the student and then dictates the sentence word by word. The recorder encourages the child to say the word slowly and write down the sounds heard. The number of phonemes correctly represented, even if a word is not correctly spelled, measures performance. It demonstrates the child's growing knowledge of phonemes and letters of the alphabet and their confidence with the alphabetic principle. The maximum score on this task is 37.

Clay (2002) reported the test-retest reliability of the Hearing and Recording Sounds in Words tasks at .64 when 160 children were assessed. Pinnell et al. (1990) reported a reliability of .96 using Cronbach Alpha when 107 students were assessed, while Pinnell, Lyons, DeFord, Bryk, & Seltzer (1994) found a Split-half reliability coefficient of .84 to .88 for 403 students.

Manitoba Word Reading Task

Manitoba teachers have been questioning the use of the Clay Word Reading task which was developed using text materials used in New Zealand primary classrooms in 1966. For the purposes of this study an alternate Manitoba Word Reading task was developed. It is necessary then to review the research related to the development of word reading lists.

Historical Development of Word Lists

Word lists have been developed from the study of children's oral vocabularies (Dale, 1943; Horn, 1925), children's spelling vocabularies (Greene, 1954; Jones & Wepman, 1966; Rinsland, 1947), adult reading vocabularies (Thorndike, 1921), and children's reading vocabularies (Dolch, 1960; Gates, 1925). A few lists have been developed based on the frequently-occurring words in classroom instructional texts (Carroll, Davis, & Richman, 1971; Harris & Jacobson, 1972). Often these lists were designed to provide classroom teachers with lists of words for instructional emphasis, rather than lists of words to document growth in reading achievement.

Researchers (Clay, 1966; Farr, 1973; Farr & Summers, 1969) have argued that reading tests should be designed to provide a general estimate of reading ability, and

words included on the list should match the words that are contained in the text materials that the students are reading. Pany and Jenkins (1978) analyzed the words introduced in seven commercially-produced reading series with the vocabulary occurring in standardized reading tests and found a discrepancy. These authors suggest that more functional assessments of the ability to read words could be completed by developing word reading tests from classroom text materials.

At the beginning of this type of word list development, Bird (1972) and Heimberger (1972) carried out a survey of a stratified sample of school districts throughout the United States to determine what texts in reading, language, mathematics, social studies, and science were most frequently used in the elementary schools. They itemized all the words in the preprimers, primers, first and second grade readers in nine, commonly-used reading series that had been published in the late sixties. In studying the characteristics of the words, Bird (1972) and Heimberger (1972) found that the first ten words in the list in order of frequency were: *the, and, a, of, to, in, you, is, he, and it*. There were two articles, one conjunction, three prepositions, one verb and three pronouns, but no nouns.

An alternative method for categorizing words in a list was to examine the relative frequency of individual words and groups of words from the total list (Sartain, 1981). When lists are developed using this approach, it is estimated that when students read the ten most-frequently occurring words, they are able to read about twenty-two percent of all the running words encountered in the most-commonly used texts. Again, these lists were developed to provide teachers with a list of words to teach, rather than for formative assessment.

Summers (1984) also described a paradigm for developing words tests based on the frequency of the words. Although he considered secondary school text material, he used the same premise to generate the *Quick Word Reading Test*. Thirty-seven textbooks prescribed for use in the seven subject areas addressed in the British Columbia curriculum for Grades 8, 9, and 10 were used to generate a frequency count of 596 words. The words were then organized into three lists. Part I consisted of the 87 high frequency words (which constituted 47% of the total running words in the 37 texts), Part II consisted of 95 general subject vocabulary words, and Part III consisted of subject vocabulary content words. When administering the *Quick Word Reading Test*, the recorder asks the student to read all three lists, with 90% accuracy expected on Part I, 70% on Part II, and 60% on Part III. The author defended the expectation of 90% mastery on Part I of this test, stating that "lack of mastery indicates poor learning and/or lack of exposure to basic instruction at the beginning acquisition level in reading" (Summers, 1984, p. 25).

Standardized Word Reading tasks are based on the principle of sampling children's reading vocabulary from a variety of texts that are read over a period of time. However, Clay (1993a) suggested that for the early identification of struggling students, a different approach is required; that the word lists be compiled from the high frequency words in the text materials that are used to teach the students to read. This would limit the number of texts used, and consequently, would assure that the words included on the lists were indeed words that students would have had the opportunity to see most frequently. It would be unfair to assess students using words that they did not have the opportunity to learn after many encounters in text. Students should be able to identify them in isolation on a word reading task, thus providing evidence of their growing ability

to learn and remember words. This sampling of words from texts read would provide the same, equitable opportunities for success for high, average, and low progress readers.

Based on this premise, Clay (1966) developed the Word Reading task for the *Observation Survey*. The lists were based on the frequency count of words in the twelve-book New Zealand "*Ready to Read*" series introduced in 1960 (Simpson, 1962). The assessment task developed by Clay consisted of the 48 most frequently occurring words in these selections placed into three lists of 15 words, of equivalent difficulty, with the three most frequently-occurring words used as the practice samples. Subsequent norming of the test showed that it was a good instrument for ranking or grouping students in the first year of school and was useful for identifying students in the second year when they were having difficulty with literacy learning (Clay, 1993a, 2002).

By using this development process, limitations were placed on the Word Reading task. Clay (2002) highlighted these as follows:

- It does not give a reading age.
- It does not discriminate so well between better readers after one year of instruction. On the contrary, it groups them together.
- Differences of less than three score points are not sufficiently reliable to support any decisions about the child's progress, without other evidence.
- It does not sample a child's word reading if [s/he] is working beyond the level of early reading books. (p. 95)

More recently, Duncan (1999) developed an alternative word reading task based on the notion that after thirty years it was questionable whether Clay's Word Reading task still contained the high frequency words found in current instructional texts. In

developing her list, Duncan used the revised *Ready to Read* books introduced by the New Zealand government for use in primary classrooms in 1985 (Department of Education, 1985) to which titles have been added and deleted as needed. A study of the high frequency words in the current set of *Ready to Read* texts (Duncan, 1999) showed that approximately 60% of the words remained the same, but many had changed. A frequency count of the words in 176 books was completed. The total number of words in these books was 17,367. The frequency of individual words was computed and the 48 most frequently-occurring words were used to create two lists with one practice word and 23 test words in each list. Duncan decided to create two lists rather than three to enhance reliability.

A comparison of the words on the lists developed by Clay (1966) and by Duncan (1999) showed that of the fifteen most frequent words in the original list, all but four appeared on the new list. This means that these words have had relatively stable use in early reading texts. The Duncan Word Reading task and the Clay Word Reading task were normed in 2001. The results showed that there was no significant difference between the measurement qualities of the two, "...they could be considered optional alternatives, and correlation's between the two measures for four age groups were .90 and above" (Clay, 2002, p. 95).

An alternative word reading task for use by Australian students was developed using some of the principles defined by Clay (1993a). The Canberra Word Reading task, developed by Clough, McIntyre, and Cowey (1990), used the high frequency words in the reading materials used in Canberra schools in 1989. The 46 words included in these lists were taken from the "*Sunshine, Storybox, Eureka Treasure Chest, and Bookshelf*" series. Three lists were completed that included one practice word (*the*). There were

fifteen words in each list. These authors did not complete a comparison of the words on the original list with the Canberra list, but normed their word list using 300 urban students ranging in age from 6:0 to 7:0 in 1990.

Another word reading task, the Ohio Word Reading task (Pinnell, et al., 1987), was developed for use with students in the United States using the high frequency words from the *Dolch Word List*. The words from this list were assigned to three lists of twenty words of equivalent difficulty. Three practice words were included. The authors decided that the words on the Dolch Word List (Dolch, 1960) were reflective of the words that would be introduced to students in first grade classrooms in Columbus, Ohio.

It is debatable whether the original Clay Word Reading task would contain words Manitoba students would encounter when learning to read. There are variations in the language used in the texts from New Zealand and the texts used in Manitoba, especially differences in vocabulary. There was a need, therefore, to develop a Word Reading task based on the text materials used by Manitoba students.

Summary

There are several areas of literacy learning to be assessed by the teacher to provide information about the child's growing control over beginning reading and writing. An effective teacher monitors the child's knowledge in each area to determine if additional instruction is required. This chapter has reviewed the research that identified critical areas in literacy development and reported how the tasks in the *Observation Survey* provide useful measures in each knowledge area. It also reviewed the research related to the development of Word Reading tasks that can be applied to the development of the Manitoba Word Reading task. However, it is also important for

teachers to have a sense of average student progress in order either to establish that all competencies are developing successfully or to identify areas of concern that warrant further instruction. The norms developed in this study should provide interpretive information related to the expected performance of grade one students at a particular point in time.

CHAPTER THREE

Research Methodology

Purpose

The overall purpose of this study was to develop Manitoba norms for interpreting performance on the six tasks from the *Observation Survey of Early Literacy Achievement* (Clay, 1993a, 2002). The first phase of the study focused on the development of a Word Reading task for Manitoba based on words occurring in the reading material used for instruction in kindergarten and grade one classrooms. The second phase focused on the development of stanines for the six tasks of the *Observation Survey of Early Literacy Achievement*, including norming the new Manitoba Word Reading task created in phase one of the study.

The study focused on the following specific questions:

- 1) Are there significant differences in the mean performance of students on the *Observation Survey* tasks based on location (i.e., between urban and rural [farming and northern communities]), age groups (i.e., between January-March, April - June, July- September, and October - December), and gender (i.e., between boys and girls)?
- 2) Are there significant differences in the mean performance of students on the Manitoba Word Reading and the Clay Word Reading tasks?

Phase I – The Development of the Word Reading Task for Manitoba

Design Overview for the Word Task

The purpose of this phase of the study was to develop a Manitoba Word Reading task which contained three lists. Based on a literature review on the construction of word

lists for assessment presented in Chapter 2 and a survey of 250 Manitoba elementary schools that identified the *PM Storybook* series (Nelson Price Milburn, 1994) as the most commonly used materials in kindergarten and grade one classrooms, province-wide, the Manitoba Word lists were developed. These lists were developed based on the following principles:

- a) Word frequency counts were produced based on the most commonly-used beginning reading text materials in Manitoba schools (*PM Storybooks* levels Magenta to Orange).
- b) Words were then systematically divided by choosing every third item to create three separate lists suitable for use as tests and retests.
- c) The most commonly used word was used as a practice word that was presented before students were asked to begin reading any of the wordlists.
- d) Instructions required students to read only one of the lists and assistance was given with the practice word if necessary. No credit was given for reading the practice word correctly (Clay, 1993a).

These procedures are described in more depth in the next section.

Identification of Text Materials

In order to develop the Manitoba Word Reading task, it was necessary to identify the text materials used in Manitoba schools to teach young children to read and to complete frequency counts of the words in these texts. The texts were identified by surveying 250 randomly-selected Manitoba elementary schools. This sample represents approximately 43% of the 583 publicly-funded elementary schools in the province. One hundred and forty-two surveys were returned, which represents 56.8% of the schools

surveyed. Table 3.1 shows the survey results indicating that one text series, the *PM Storybooks*, were used most frequently in Manitoba schools.

Table 3.1

Frequency of Use of Text Series in Manitoba Schools Per Grade

Series	Publishers Publication Date	Kindergarten	Grade One	Total
PM Storybooks	Nelson Price Milburn, 1994	75	51	126
Sunshine	The Wright Group, 1993	38	24	62
Collections	Pearson Education Canada, 1999	1	21	22
Bookshop	Scholastic Canada, 1980	5	10	15
Nelson Language Arts	Nelson Canada Ltd. 1999	1	12	13
Total		120	118	238

*This list does not include any text series with less than 12 books in total.

Word Frequency Counts

To develop the Manitoba Word Reading task, a word count was completed on all the books from the *PM Storybooks* series which are used to teach grade one children to read. All the books in the Magenta, Red, Yellow, Blue, Green, and Orange levels (see Appendix A) from the *PM Storybook* series were used for the word count. The color band codes used in the *PM Storybook* series correspond to the Reading Recovery™ number ratings used in some Manitoba schools. The number coding for Reading Recovery™ levels provides a gradient of text difficulty with the following correspondence between color and number levels: Magenta level (Reading Recovery levels 1 and 2), Red level (Reading Recovery levels 3, 4 and 5), Yellow (Reading

Recovery levels 6, 7 and 8), Blue (Reading Recovery levels 9, 10, and 11), Green (Reading Recovery levels 12, 13, and 14), and the Orange level (Reading Recovery levels 15 and 16). There are 154 titles available in the *PM Storybook* series from the Magenta to the Orange levels.

The *PM Storybooks* were designed to introduce new, basic reading words in each selection. The number of new words increases at each color level until 350 are mastered by the end of the Orange level (Randell, 1999). The first level, the Magenta color, is most supportive of emergent readers. Texts at this level contain one or two lines with a repetitive sentence pattern and illustrations that offer high support (Peterson, 1988). The highest level used for the frequency-count, the Orange level includes varied sentence patterns and illustrations that provide only low support. Appendix B presents the word frequency-counts from the various levels of the *PM Storybooks* used to develop the Manitoba Word Reading task.

Using the information from this frequency count, the first sixty-one words were compared to the existing Word Reading tasks cited earlier to determine differences between the newly-generated word lists and existing word lists (Canberra, 1990; Clay, 1966; Duncan, 2000; Pinnell, et al., 1991). If the words generated in this frequency count were identical to the words on the Clay, Duncan, Canberra, and Ohio word lists, then there would no need to create a new list for Manitoba.

Comparison to Other Word Reading Tasks

A comparison of the words included on each Word Reading task cited above was completed to determine if there were significant differences compared to previously-developed lists. It had been predicted that many of the words would be the same since

they occur commonly. However, if only minor changes were evident from the alternative lists, then there would be no need to construct, validate, or norm the new Manitoba Word Reading task. Appendix C shows that there were some differences between the Clay (59%), Canberra (66%), and Duncan (69%) word lists, but substantial differences compared to the Ohio Word list (only 46%). This comparison showed that the Manitoba lists, which consisted of 61 words, contained additional words not found on any of the lists (see Appendix C). It was therefore important to create a Manitoba Word Reading task for this validation study. The next section describes procedures followed during the development of the Manitoba Word Reading task.

Developing the Manitoba Word Reading Task

The first consideration was to use the most frequently-occurring word for practice on all three lists developed for the Manitoba Word Reading task. In the Clay (1966), Duncan (2001), and Ohio (1991) Word Reading tasks, the three most commonly-occurring words were used as the practice words. Consequently, these words could be read by the child but not counted in the score. Since the purpose of the task was to demonstrate mastery, it did not seem prudent to use three of the most commonly-occurring words for practice. In the Manitoba Word Reading task, one practice word was used for all three separate word lists. This approach is similar to the Canberra Word Reading task (1990).

The second consideration was whether or not to include words such as *I* or *a*. Teachers frequently question the validity of including the words *I* or *a* since it is impossible to determine if children are identifying them as letters or words. Three of the current word reading tasks, the Clay (1966), the Canberra (1990) and the Duncan (2001)

include these words but the Ohio (1991) does not. However, for this study, the words *I* and *a* were included on the list since these letters are also words that appear frequently in the selections Manitoba students encounter.

The third consideration was whether or not words should be moved so that similar words such as *come* and *came* did not appear on the same list. The decision was made to keep the words in their original position based on the frequency count because moving similar words to different lists would ignore the frequency rating. For example, the word *come* appeared 171 times in the texts, but *came* appeared 61 times. *Come* could have been replaced with *it* which also had a frequency count of 171, but that would have placed the word *come* in the same list as *comes*. Care was also taken to have the size and font type on the Word Reading tasks match that used in the *PM Storybooks* series.

The final three lists consisted of one sample or practice word (*the*) and twenty words arranged sequentially in order of occurrence. This procedure was based on Clay's recommendation (1993a) and the results of the literature review (Summers, 1984), which suggested that the words in the lists be arranged in an actual hierarchy or generally accepted developmental sequence so that students could demonstrate mastery. Table 3.2 illustrates how words were placed on the three lists. The Manitoba Word Reading task in included in Appendix D. The next task was to validate the word lists.

Table 3.2

Location of Words on List Based on Frequency Count

Most Frequently Occurring Word (Practice Word)		
The (1)*		
List #1	List #2	List #3
2	3	4
7	6	5
8	9	10
13	12	11
14	15	16
19	18	17
20	21	22

*Placement or rank of the word on the frequency word counts list.

Validation of Manitoba Word Reading Task

The word lists trials were conducted using 150 randomly-selected grade one students in twelve Manitoba schools, six urban and six rural (farming and northern communities), in December, 2001. The purpose of this segment of the study was to determine both the equivalencies of the lists and their validity. Two Reading Recovery™ teacher leaders and one trainer administered the lists. The administration of the lists was counterbalanced: 50 students given List A, List B and then List C; 50 students given List B, List C and then List A; and 50 students given List C, List A and then List B. The resulting mean scores were then analyzed to determine:

1. if the order of administration influenced students' scores,
2. if the lists were equivalent in terms of difficulty, or
3. if words needed to be moved or removed from the list.

Results of Validation

The purpose of the validation study was to ensure that the order of administration did not result in significant differences in students' scores, that the three lists developed for the Manitoba Word Reading task were equivalent, and to identify words that should be moved or removed from the lists. The data from the 150 students who completed the tasks were analyzed using the following statistical methods.

1. Scores from each of the three lists were analyzed using Pearson correlations to determine if the order of administration influenced student scores.
2. Scores from all students completing the three lists were analyzed using Pearson correlations to determine the association between the lists.
3. Scores for each of the words included on the three lists were analyzed using Point Biserial correlation coefficients to determine if words should be removed or moved to another list. The Kruder-Richardson reliability coefficient was also used as a measure of internal consistency of individual items.

Order of Administration. Pearson product-moment correlation coefficients (Janda, 1998) were completed for each of the word lists to determine if order of administration was a factor in the scores obtained. The Pearson correlation coefficient is a statistical procedure used with an interval scale that allows the determination of the strength and direction of the relationship between two variables. This procedure yields a single number that can have an absolute value of 0.0 to 1.0. The closer the absolute value is to 1.0 the stronger the relationship, and the closer the absolute value is to 0.0, the weaker the relationship. Pearson correlation coefficients of less than .19 suggest that there is little if any association between the scores, and correlations below .29 reflect a weak

association. Tables 3.3, 3.4, and 3.5 illustrate that the order of administration did not impact on student scores (List A, $r = .003$, $r = .084$, and $r = .249$); (List B, $r = .173$, $r = .041$, and $r = .086$); and (List C, $r = .194$, $r = .107$, and $r = .171$).

Table 3.3

Pearson Correlations for List A Based on Order of Administration

List A Order of Administration	Pearson Correlation Sig. (2-tailed)		
	First	Second	Third
ABC	-.003 .985 ns		
CAB		-.084 .563 ns	
BCA			.249 .081 ns

Table 3.4

Pearson Correlations for List B Based on Order of Administration

List B Order of Administration	Pearson Correlation Sig. (2 tailed)		
	First	Second	Third
ABC		-.041 .766 ns	
CAB			-.086 .552 ns
BCA	.173 .229 ns		

Table 3.5

Pearson Correlation for List C Based on Order of Administration

List C Order of Administration	Pearson Correlation Sig. (2 tailed)		
	First	Second	Third
ABC			-.171 .236 ns
CAB	.194 .177 ns		
BCA		-.107 .458ns	

Comparison of the Lists. A series of Pearson correlations were conducted to determine if there were significant differences between the mean scores of all the students completing the three lists. The Pearson correlations $r = .917$, $r = .931$, and $r = .881$, as shown in Table 3.6, indicate a strong association between the three lists.

Table 3.6

Pearson Correlation of Word Lists		
List	Pearson Correlation Sig. (2-tailed)	
	A	B
B	.917**	
C	.881**	.931**

**Correlation is significant at the 0.01 level (2-tailed)

Analysis of Words

To determine if any words should be removed or moved to another list, an analysis of the each word was completed, including the proportion of students answering the item correctly (p) and the Point Biserial correlation coefficient (r_{pbi}) (see Table 3.7). The Point Biserial correlation coefficient (r_{pbi}) indicates how well an item discriminates between high and low performing students (Janda, 1998). The normal range of biserial scores for each item is between 0, indicating that this item has no discriminative quality, and 1, indicating the item discriminates well between high and low performing students. The normal range expected for items is between .20 and .80. As indicated in Table 3.7, the range of point biserial correlation coefficients suggests that the words on the lists do discriminate between high and low performing students.

The Kruder-Richardson reliability coefficient (Janda, 1998) was also used as a measure of internal consistency. The relatively high K-R₂₀ (.912 to .923) reported in Table 3.7 reflects the homogeneity of the items and the test's ability to measure accurately.

Table 3.7

Percentage Correct and Point Biserial Correlations for Words on Manitoba Lists

List A	<i>p</i>	<i>r_{pbi}</i>	List B	<i>p</i>	<i>r_{pbi}</i>	List C	<i>p</i>	<i>r_{pbi}</i>
said	.72	.71	and	.86	.71	to	.94	.51
I	.93	.53	is	.93	.43	a	.93	.33
in	.88	.44	he	.74	.66	here	.55	.58
on	.81	.55	little	.69	.64	up	.88	.64
went	.51	.64	at	.79	.66	look	.88	.67
you	.87	.61	Dad	.88	.39	she	.71	.85
for	.83	.60	are	.74	.75	they	.46	.69
it	.79	.74	come	.68	.66	Mom	.95	.44
my	.83	.58	down	.59	.71	big	.74	.69
go	.88	.55	can	.76	.62	we	.87	.63
bear	.41	.61	was	.50	.69	with	.45	.72
am	.61	.52	ran	.35	.55	looked	.49	.65
me	.83	.62	of	.53	.62	baby	.44	.73
will	.65	.75	all	.66	.50	out	.55	.73
not	.68	.79	going	.60	.74	I'm	.47	.64
where	.36	.62	like	.69	.58	Mother	.51	.71
comes	.49	.69	get	.53	.72	but	.54	.75
back	.39	.63	see	.89	.63	too	.87	.59
no	.61	.43	this	.56	.71	her	.34	.61
day	.61	.71	shouted	.20	.47	came	.23	.55
Average	.69	0.616	Average	.66	0.622	Average	.64	0.636
SD	5.269		SD	5.528		SD	5.519	
K-R ₂₀ =	.912		K-R ₂₀ =	.919		K-R ₂₀ =	.923	

Summary

Based on these analyses, the final Manitoba Word Reading task was developed for the norming study. To prevent relatively similar words from appearing in the same list, three words from list A were exchanged with words from list C, *look* was exchanged

for *go*, and *Mother* for *went*, and *to* for *my*. The final word reading task was then used in the norming study with the other tasks in the *Observation Survey*.

Phase Two - Norming of Six Tasks

Design of the Norming Study

In Manitoba, the six tasks from Clay's (2002) *Observation Survey of Early Literacy Achievement* are used with grade one students to determine areas of relative strength. This norming study included grade one students who ranged in age from 6.0 to 7.0 years. The final sample included 200 boys, 200 girls, and 200 urban and 200 rural (farming and northern communities) students, sampling procedures being described as follows.

Sample Size

In Manitoba, there are approximately 19,000 students entering grade one each year, with about 14,000 of these students attending publicly-funded schools. Of the 14,000 students, approximately 2,300 students are in Immersion or Bilingual programs. For the purposes of this study, only students in full-day English programs participated since the *Observation Survey* tasks are used primarily to measure progress of students in these programs. Based on the grade one population size of about 11,700 students, the ideal sample size is 375 students (Clark-Carter, 1997). Thus, the sample ($n = 400$) used in this study meets this requirement.

Selection of School Divisions

All fifty-four school divisions in Manitoba were considered for the study. Ten of these fifty-four divisions were not included either because of access difficulties, very low grade one enrollments, or because a high percentage students received their language arts instruction in French. The remaining forty-four school divisions, nine urban and thirty-five rural (farming and northern communities), were considered for inclusion. Eight school divisions that had previously participated in the validation of the Manitoba Word Reading task were eliminated, leaving thirty-six divisions. All six urban and twenty randomly-selected rural (farming and northern communities) divisions from this pool were invited to participate in this norming study. Superintendents from four urban and seven rural (farming and northern communities) school divisions gave permission for their schools to participate.

Selection of Schools Within the Divisions

All of the schools providing English Language Arts programs in these divisions were invited to participate in the study. Ninety-nine schools received letters inviting their participation and 26 rural (farming and northern communities) and 24 urban schools agreed. The rural (farming and northern communities) schools represented all areas of the province.

Selection of Students

The schools were provided with parental approval permission letters to allow each child from the grade one classrooms to participate in the study. From this list, students were randomly selected by name. In most schools, a random selection of ten

students, five girls and five boys was identified for assessment. An additional two students, one boy and one girl, were randomly selected to replace previously selected students in the event that parents did not give final approval, a student was absent the day of the assessment, or the student decided to withdraw from the study. In smaller schools with smaller numbers of grade one students, four or six students were randomly selected for assessment. The final sample included a possible 455 grade one students in fifty schools from various parts of the province. Table 3.8 shows the breakdown of the number of grade one students in each school division and the percentage of students assessed.

Table 3.8

School Division Population and Percentage of Students

Urban			Rural		
Number of Grade One Students per School Division	Number of Students Assessed	Percentage of Total Eligible Students Assessed	Number of Grade One Students per School Division	Number of Students Assessed	Percentage of Total Eligible Students Assessed
187	36	19%	80	26	33%
174	42	24%	90	31	34%
386	80	21%	125	36	29%
277	60	22%	210	52	25%
			121	20	17%
			116	24	21%
			117	30	26%
1024	218	23%	859	219	25%

Overall, 455 students were selected for assessment with the tasks of the *Observation Survey*. However, 437 students were assessed, 18 students were not included in the study because either permission from the parents could not be obtained ($n = 5$), the students were not present at school the day the assessments were completed ($n = 10$), or the student did not agree to complete the assessments ($n = 3$).

Time of Administration

The optimum time for the assessment was mid-grade one or February. If the tasks were administered at the beginning of the school year, students would not have had the benefit of formal literacy instruction, leading to low scores across tasks. On the other hand, four of the tasks had maximum scores that most students would reach by the end of the year. Scores obtained from the *Observation Survey* are valid for only three to six weeks as a result of the rapid learning that occurs in beginning reading. To ensure that students had benefited from the classroom program, the optimum period of time for data gathering was determined to be within the three-week period from February 4 to February 28. By restricting the highest level of text reading to Level 20, which was deemed realistic for the time of year (see Appendix E), it was possible to complete the administration of all of the *Observation Survey* tasks, including Running Records, in approximately 35 minutes. A protocol for administering the Running Records (see Appendix F) and a list of common titles from the *PM Storybooks* series were also provided (see Appendix G). This reduced the time required to prepare for the assessments.

Task Administrators

The assessments were administered by the Manitoba Reading Recovery™ Teacher Leaders. Teacher Leaders are prepared for their role in a year-long course that provides academic study and practical work with students and teachers. Within the Reading Recovery™ professional development model, Teacher Leaders are essential people because they teach children, work with teachers, educate local educators, negotiate the implementation of the program, and act as advocates. Most importantly for

this study, the Teacher Leaders are responsible for providing the inservice course for Reading Recovery™ teachers which includes helping classroom teachers work with the *Observation Survey* tasks to ensure that they have administrative expertise.

For this study, eight Manitoba teacher leaders, five teacher leaders-in-training, and two trainers administered the tasks. These teacher leaders were located in various regions across the province and had access to schools in both urban and rural (farming and northern communities) areas. Consequently, 437 students were assessed by 15 Reading Recovery™ personnel within the three-week period in February of the school year.

Number of Protocols

After completion, task protocols were returned to the investigator and checked to ensure quality of administration and scoring accuracy. A total of 437 students were assessed from February 4 – 28 of the school year. When the protocols were checked, however, the task sheets for a number of students were removed for the following reasons.

1. In accordance with ethical provisions, students were entitled to withdraw from the study at any time. Seven students chose not to complete all of the tasks. If one or more of the seven tasks were not completed, none of the tasks for these students were included in the analysis.
2. Twenty-one sets of tasks had to be removed because Running Records were not complete. The primary difficulty was that the instructional level had only been estimated. For example, the student may have read text at level 14 at the instructional level with a word accuracy score that fell between 90 to 95%, then read level 16 with

word accuracy that fell below 90%, suggesting difficulties. The student's exact level could not be determined however, because s/he was not assessed at Level 15. These estimated scores were dropped.

3. Data for eight students were removed because one of the task sheets had been omitted from the package submitted for analysis.
4. Data from one urban female student was randomly selected for removal so that the sample size would include equal numbers of girls and boys and urban and rural (farming and northern communities) students.

Summary

The first task was to develop and validate the Manitoba Word Reading task. The 61 most frequently occurring words from the *PM Storybooks*, the most commonly-used text materials in Manitoba kindergarten and grade one classrooms, were identified. These 61 words were placed into three lists, the most frequently occurring word being used as the practice word. Using Pearson correlation coefficients indicated that the three lists were of similar difficulty, $r = .881$, $r = .917$ and $r = .931$. Pearson correlation coefficients showed that the order of administration, which was counterbalanced, did not change student scores.

Using Point Biserial (r_{pbi}) correlation coefficients demonstrated that the words did discriminate between high and low progress readers (r_{pbi} greater than .30). It was therefore not necessary to remove any items. Three words were moved from list A to list C so that similar words (e.g., *to* and *too*) would not appear on the same list.

A sample of 437 students were randomly selected from 50 schools across the province and were assessed in February, 2002 by 15 Reading Recovery™ personnel

using the six tasks in the *Observation Survey* and the Manitoba Word Reading task. The protocols were checked and 36 removed because of administration difficulties or errors. The scores for one student were removed at random so that the final sample included 400 students, with an equal number of students from urban (200) and rural (farming and northern communities) (200) areas and boys (200) and girls (200). The scores of these students were used to develop stanine scores for each of the tasks.

The next chapter provides the results. Mean performances were analyzed using a series of analyses of variance (ANOVA's) to determine if separate stanine scores should be developed based on location (urban and rural [farming and northern communities]), age (January - March, April - June, July - September, and October - December), and gender (boys and girls). Differences in mean performance on the Manitoba and Clay Word lists were also analyzed using Pearson correlations to establish whether the Manitoba Word Reading task should be substituted for the Clay Word Reading task.

CHAPTER FOUR

Data Analysis

The previous chapter examined the methodology employed to develop and validate the Manitoba Word Reading task and described the procedures used to gather the data for developing a set of standard scores for Clay's *Observation Survey of Early Literacy Achievement* (Clay, 1993a, 2002) based on the performance of Manitoba grade one students. As stated earlier, the previous norming studies in New Zealand (2000) included students between the ages of 5 and 7 years, while the Ohio (1991) study examined the achievement of first grade students at three points during the first grade, fall, mid-year, and late spring. The Canadian norming study of the *Observation Survey* in French (Bourque, 2002), completed in 2001, included students who had been in school for eighteen months. No norming study has been conducted using English-speaking students in Canada.

The purpose of this study was to provide stanine scores that could be used by teachers in Manitoba schools as a basis for comparison to determine levels of achievement as students move through the grade one program. This chapter includes the analyses of data from all Clay's *Observation Survey* tasks collected from the 400 students who were assessed in this Manitoba study. The data analysis was conducted in response to two major questions:

1. Are there significant differences in the mean performance of the students in the *Observation Survey* tasks based on location (i.e., between urban and rural [farming and northern communities]), age (i. e., between January - March, April - June, July - September, and October - December), and gender (i.e., between boys and girls)?

2. Are there significant differences in the mean performance of students on the Manitoba and Clay Word Reading Tasks?

Overall Scores

Demographic and test data were recorded for 400 Manitoba students in February of the school year. Test data consisted of students' raw scores on seven measures: six subtasks of the *Observation Survey* (i.e., Running Records of Text Reading, Concepts about Print, Letter Identification, Clay Word Reading, Writing Vocabulary, and Hearing and Recording Sounds in Words or Dictation), and the Manitoba Word Reading task. The means and standard deviations were first calculated for each task, and then the data analyzed using a series of analyses of variance (ANOVA's) to assess the effects of three variables, location (urban and rural [farming and northern communities]), age (January - March, April - June, July - September, and October - December) and gender (boys and girls). Subsequently, sets of stanine scores for specific *Observation Survey* tasks were developed using procedures recommended by Lyman (1991). Stanine scores represent broad bands of scores so it might be possible that although there were significant differences between the mean scores of boys or girls and urban and rural (farming and northern communities) students, these differences were not evident when they became collapsed within the range of stanine scores. Stanine scores were analyzed for differences and finally, an overall set of stanine scores developed.

Impact Of Variables

Location

The specific question for study regarding location centered on whether or not there were significant differences in the mean performance of students on the *Observation Survey* tasks between urban and rural (farming and northern communities). The sample for this study was equally divided, 50% urban ($n = 200$) and 50% rural (farming and northern communities) students ($n = 200$).

Analyses of variance (ANOVA's) are used to determine whether the mean of a variable differs among groups (Janda, 1998). A series of analyses of variance revealed that urban students scored significantly higher than rural (farming and northern communities) students on four of the seven measures: *Running Records of Text Reading*, $F(1, 398) = 10.832, p = .001$; *Clay Word Reading*, $F(1, 398) = 12.031, p = .001$; *Hearing and Recording Sounds in Words*, $F(1, 398) = 9.021, p = .003$ and the *Manitoba Word Reading Task*, $F(1, 398) = 11.271, p = .001$. These results are reported in Table 4.1.

Table 4.1

Analyses of Variance for Each Task by Location

Measure	df	F	p-value
Running Records	1, 398	10.832	.001 *
Concepts About Print	1, 398	3.750	.052 ns
Letter Identification	1, 398	3.303	.070 ns
Clay Word	1, 398	12.031	.001 *
Writing Vocabulary	1, 398	1.806	.180 ns
Dictation	1, 398	9.021	.003 *
Manitoba Word	1, 398	11.271	.001 *

* $p < .05$

Age

To examine the impact of age, students' birth dates were aggregated into four quarters: January to March, April to June, July to September, and October to December. As shown in Table 4.2, the proportion of students whose date of birth fell within each of the four categories was fairly balanced, with each category representing between 23.8% and 27.7% of the students.

Table 4.2

Number of Students in Each Quarter of Year Proportion of Total Sample

Dates of Birth	Number of Students	Proportion of Total Sample
Jan.- Mar.	103	25.3%
April - June	111	27.7%
July - Sept.	91	22.8%
Oct. - Dec.	95	23.8%
Total	400	100.0%

Similarly, shown in Table 4.3, when age was broken down by gender, the number of boys and girls was fairly equally-distributed across the four quartiles of the year.

Table 4.3

Proportion of Boys and Girls per Quartile.

Dates of Birth	Number of Students	Proportion of Boys and Girls			
		Boys	Percentage	Girls	Percentage
Jan. - Mar.	103	52	50.5%	51	49.5%
April - June	111	59	53.1%	52	46.9%
July - Sept.	91	44	48%	47	52%
Oct. - Dec.	95	45	47%	50	53%
Total	400	200		200	

A series of analyses of variance (ANOVA's) conducted on the mean scores for each of the measures, for each birth date category, depicted in Table 4.4, yielded the following non significant results: *Running Records*, $F(3, 396) = 1.048, p = .371$; *Concepts About Print*, $F(3, 396) = .372, p = .774$; *Letter Identification*, $F(3, 396) = .207, p = .892$; *Clay Word Reading*, $F(3, 396) = 1.055, p = .368$; *Writing Vocabulary*, $F(3, 396) = 1.861, p = .136$; *Hearing and Recording Sounds in Words (Dictation)*, $F(3, 396) = 1.697, p = .167$; and the *Manitoba Word Reading*, $F(3, 396) = 1.720, p = .162$. As indicated, there were no significant differences, thus students' scores do not differ based on date of birth. This finding was similar to the study by Bourque (2002) which showed no statistically significant difference between the scores of grade one students based on age.

Table 4.4

Analyses of Variance for Tasks Based on Age

Task	df	F	p-value
Running Records	3, 396	1.048	.371 ns
Concepts about Print	3, 396	.372	.774 ns
Letter Identification	3, 396	.207	.892 ns
Clay Word Reading	3, 396	1.055	.368 ns
Writing Vocabulary	3, 396	1.861	.136 ns
Dictation	3, 396	1.697	.167 ns
Manitoba Word Reading	3, 396	1.720	.162 ns

* $p < .05$ *Gender*

The specific question for study was related to whether or not there were significant differences in the mean performance of students on the *Observation Survey* tasks based on gender (boys and girls). With an equally distributed sample that was 50% boys ($n=200$) and 50% girls ($n=200$), an analyses of variance (ANOVA's) conducted on mean scores on each measure by gender determined that female students scored significantly higher than male students on four of the seven measures: *Running Records of Text Reading*, $F(1, 398) = 7.498$, $p = .006$; *Writing Vocabulary*, $F(1, 398) = 17.848$, $p = .000$; *Hearing and Recording Sounds in Word (Dictation)*; $F(1, 398) = 5.766$, $p = .017$; and the *Manitoba Word Reading Task*, $F(1, 398) = 4.060$, $p = .045$, as indicated in Table 4.5.

Table 4.5

Analyses of Variance For Each Task by Gender

Task	df	F	p-value
Running Records	1, 398	7.498	.006*
Concepts About Print	1, 398	2.573	.109
Letter Identification	1, 398	.402	.526
Clay Word Reading	1, 398	2.428	.120
Writing Vocabulary	1, 398	17.848	.000*
Dictation	1, 398	5.766	.017*
Manitoba Word	1, 398	4.060	.045*

* $p < .05$ *Summary of Impact of Variables*

The statistical analyses shows the following findings:

- Urban students scored significantly higher than rural (farming and northern communities) students on *Running Records of Text Reading*, *Clay Word Reading*, *Hearing and Recording Sounds in Words*, and *Manitoba Word Reading* tasks.
- There were no statistically significant differences based on age.
- Females scored significantly higher than males on *Running Records of Text Reading*, *Writing Vocabulary*, *Hearing and Recording Sounds in Word (Dictation)*, and *Manitoba Word Reading*.

Stanine Scores

Stanines scores were then developed for each of the tasks in which there were significant differences based on location (urban and rural [farming and northern communities]) and gender (boys and girls). Standard scores facilitate comparisons for use in identifying performance differences. Based on the preceding findings, stanine

scores for location (urban and rural [farming and northern communities]) were developed for: the *Running Records of Text Reading*, *Clay Word Reading*, *Hearing and Recording Sounds in Words (Dictation)*, and *Manitoba Word Reading*. Similarly, stanine scores for gender (boys and girls) were developed for *Running Records of Text Reading*, *Writing Vocabulary*, *Hearing and Recording Sounds in Words (Dictation)* and *Manitoba Word Reading Task*.

Developing Stanine Scores

Stanines are normalized scores made up of a nine-unit scale with a mean of five and a standard deviation of two. Stanines allow monitoring of student progress as it changes from below average to what might be considered average or above. Stanines of four, five, or six fall within the average band. Each unit, except one and nine, constitutes .5 standard deviations. Accordingly, the spread of scores among children in the research sample was transformed to fit the normal distribution curve (Lyman, 1991).

Since raw scores are grouped into stanines, it might be said that stanines are not sensitive indicators. Bauman (1988) suggests that believing in the exactitude of a raw score should not be encouraged. Teachers must not, therefore, rely on raw score differences as being real rather than due to chance. Bauman argues that stanines provide a gross measure because a stanine of five represents a percentile range of 40 to 60. Bauman argues that since each is a single digit, stanines have the advantage of being easily interpreted. Stanines are advantageous for interpreting performance because they place all the different scoring systems on one scale. This facilitates comparisons across *Observation Survey* tasks.

Lyman (1991) recommended a set of procedures for converting raw scores into stanines that follows the normal curve. Consequently, 4% of the sample would fall into Stanine 1, 8% into Stanine 2, 12% into Stanine 3, 16% into Stanine 4, and 20% into Stanine 5. Working backwards again, 16% of the sample falls into Stanine 6, 12% into Stanine 7, 8% into Stanine 8, and 4% into Stanine 9. These percentages (Lyman, 1991) were used to divide the scores obtained in this study into stanine equivalents.

Comparing Stanine Scores

Location. An examination of the data from the stanine scores (Tables 4.6, 4.7, 4.8 and 4.9) depicting scores from *Running Records of Text Reading*, *Clay Word Reading*, *Hearing and Recording Sounds in Words (Dictation)*, and the *Manitoba Word Reading* tasks respectively, shows that the variation occurred mainly at the lower end (Stanines 1-3). As reported earlier, stanine scores, particularly those in the first three-stanine levels, are used to identify areas in which students need more instructional support. Despite the fact that there were statistically significant differences between urban and rural (farming and northern communities) students on these two tasks, it seemed practical to designate one set of stanine scores for interpreting the performance of both urban and rural (farming and northern communities) students.

Table 4.6

Stanine Scores for Urban and Rural Students: Running Records of Text Reading

Stanine Group	1	2	3	4	5	6	7	8	9
Urban Raw Score	0-1	2-3	4-6	7-9	10-13	14-16	17-19	20	
Rural Raw Score	0	1	2-3	4-8	9-13	14-19	20		

Table 4.7

Stanine Scores for Urban and Rural Students: Clay Word Reading Task

Stanine Group	1	2	3	4	5	6	7	8	9
Urban Raw Score	0-3	4-7	8-9	10-11	12-13	14	15		
Rural Raw Score	0-1	2-3	4-7	8-10	11-13	14	15		

Table 4.8

Stanine Scores for Urban and Rural Students: Hearing and Recording Sounds in Words

Stanine Group	1	2	3	4	5	6	7	8	9
Urban Raw Score	0-21	22-27	28-31	32-33	34-35	36	37		
Rural Raw Score	0-11	12-24	25-30	31-32	33-35	36	37		

Table 4.9:

Stanine Scores for Urban and Rural Students: Manitoba Word Reading

Stanine Group	1	2	3	4	5	6	7	8	9
Urban Raw Score	0-13	14-15	16	17	18	19	20-21	22	23-24
Rural Raw Score	0-11	12-14	15	16	17-18	19	20	21-23	24

Gender. On both writing tasks (Writing Vocabulary and Hearing and Recording Sounds in Words) boys in the sample scored significantly lower than girls. Similarly, on Text Reading and the Manitoba Word Reading task, there was also a significant difference based on gender. However, when comparing the results by stanine scores, there is very little difference between the stanine ranks for boys and girls except on the Writing Vocabulary task. As shown in the tables (Tables 4.10, 4.11, 4.12, 4.13), the

progress of boys in both the mid (Stanines 4, 5 and 6) and high range (Stanines 7, 8 and 9) was very similar to that of girls on these three tasks.

Table 4.10

Stanine Scores for Boys and Girls: Running Records of Text Reading

Stanine Group	1	2	3	4	5	6	7	8	9
Boys Raw Score	0	1	2-3	4-7	8-12	13-17	18-19	20	
Girls Raw Score	0	1-2	3-7	8-10	11-15	16-19	20		

Table 4.11

Stanine Scores for Boys and Girls: Writing Vocabulary

Stanine Group	1	2	3	4	5	6	7	8	9
Boys Raw Score	0-6	7-14	15-20	21-27	28-35	36-43	44-48	49-60	61-79
Girls Raw Score	0-11	12-16	17-26	27-34	35-42	43-49	50-60	61-67	68-81

Table 4.12

Stanine Scores for Boys and Girls: Hearing and Recording Sounds in Words

Stanine Group	1	2	3	4	5	6	7	8	9
Boys Raw Score	0-12	13-24	25-30	31-32	33-35	36	37		
Girls Raw Score	0-18	19-28	29-32	33-34	35	36	37		

Table 4.13

Stanine Scores for Boys and Girls: Manitoba Word Reading

Stanine Group	1	2	3	4	5	6	7	8	9
Boys Raw Score	0-2	3-7	8-11	12-15	16-17	18-19	20		
Girls Raw Score	0-3	4-9	10-13	14-16	17-18	19	20		

Although there were some differences, particularly at stanine levels 1-3, by mid-range the differences began to disappear. While, stanine scores help distinguish students at lower levels who require additional opportunities to build knowledge in these areas, it seems, nonetheless, that one set of stanine scores will suffice for all students regardless of location (urban and rural [farming and northern communities]), age (January- March, April - June, July- September, and October - December), or gender (boys and girls). (See Appendix H for set of overall stanines.)

Comparison of Stanine Scores

The Manitoba stanine scores for the Concepts About Print, Letter Identification Writing Vocabulary and the Hearing and Recording Sounds in Words tasks were compared to the ones developed in New Zealand (2001) and Ohio (1990). Comparisons of the Running Records of Text Reading could not be completed because New Zealand does not report stanines for text reading and the U.S. stanines reflect the use of the Scott Foresman testing package. Clay Word Reading task stanines were also compared with those developed from this Manitoba study but no Ohio comparisons were carried out since different assessments tools were used.

Tables 4.14, 4.15, 4.16, and 4.17 present the stanines scores for two age groups from the New Zealand norms representative of the age groups assessed in this study, plus mid-year Ohio and the Manitoba stanines. Comparisons of the stanines show relatively similar scores for Manitoba and New Zealand students on *Concepts About Print, Letter Identification* and *Hearing and Recording Sounds in Words*, and similar scores for Manitoba and Ohio students on *Concepts about Print*, and *Letter Identification*.

However, differences between stanine ranges occurred on the Writing Vocabulary task.

Both groups of New Zealand students, 6.00 - 6.50 and 6.51 - 7.0, scored higher than Manitoba and Ohio students, particularly at higher ranges. It is important to note that the students in the New Zealand sample of 6.51 - 7.0 had been in school full time for 18 to 24 months, but Manitoba and Ohio students had been in school half-time for 12 months and full-time for only 6 months.

Table 4.14

Comparison of Stanines for Concepts About Print

Stanine Group	1	2	3	4	5	6	7	8	9
New Zealand (6.0 - 6.5)	0-11	12-14	15-16	17	18-19	20	21-22	23	24
New Zealand (6.51 - 7.0)	0-13	14-15	16-17	18	19-20	21	22-23	24	
Ohio (Feb.)	0-10	11-12	13	14-15	16	17-18	19	20	21
Manitoba (Feb.)	0-12	13-14	15	16-17	18	19	20	21-22	23-24

Table 4.15

Comparison of Stanines for Letter Identification

Stanine Group	1	2	3	4	5	6	7	8	9
New Zealand (6.0 - 6.5)	0-34	35-48	49-50	51	52	53	54		
New Zealand (6.51- 7.0)	0-40	41-50	51-52	53		54			
Ohio (Feb.)	0-46	47-49	50	51	52	53		54	
Manitoba (Feb.)	0-47	48-50	51	52	53	54			

Table 4.16

Comparison of Stanines for Writing Vocabulary

Stanine Group	1	2	3	4	5	6	7	8	9
New Zealand (6.0 - 6.5)	0-4	5-13	14-25	26-36	37-49	50-59	60-69	70-83	84+
New Zealand (6.51- 7.0)	0-8	9-25	26-35	36-45	46-56	57-66	67-80	81-99	100+
Ohio (Feb.)	0-7	8-12	13-16	17-19	20-24	25-30	31-36	37-40	41-57
Manitoba (Feb.)	0-7	8-15	16-22	23-29	30-39	40-46	47-55	56-65	66-81

Table 4.17

Comparisons of Stanines for Hearing and Recording Sounds in Words

Stanine Group	1	2	3	4	5	6	7	8	9
New Zealand (6.0 - 6.5)	0-8	9-19	20-27	28-32	33-35	36	37		
New Zealand (6.51- 7.0)	0-14	15-28	29-32	33-35	36	37			
Ohio (Feb.)	0-13	14-19	20-23	24-26	27-30	31-32	33-34	35	36-37
Manitoba (Feb.)	0-15	16-26	27-31	32-33	34-35	36	37		

A comparison of Clay Word Reading task stanines based on the performance of Manitoba and New Zealand students for the appropriate age groups was also completed. Table 4.18 shows the similarities between the stanine groupings. Especially when the stanines for the range 6.0 to 6.5 were used as a basis of comparison, stanine ranges for word identification were relatively similar, although mastery for Manitoba students in February (score 15) fell into stanine 7.

Table 4.18

Comparisons of Stanines for Clay Word Reading

Stanine Group	1	2	3	4	5	6	7	8	9
New Zealand (6.0 - 6.5)	0-1	2-4	5-9	10-12	13-14	15			
New Zealand (6.51- 7.0)	0-4	5-10	11-12	13	14	15			
Manitoba (Feb.)	0-2	3-4	5-8	9-11	12-13	14	15		

Comparison of Clay and Manitoba Word Reading Tasks

A second purpose of the study was to determine if there were significant differences in the mean performance of students on the Manitoba and Clay Word Reading Tasks. Pearson correlations conducted on the mean scores for the Clay and Manitoba Word Reading tasks and the resulting coefficient, ($r = .914$) showed nevertheless, that the measurement qualities of the two word reading tasks were the same. Teachers could consider the two word reading tasks as optional alternatives, however, the Clay Word Reading task may be preferred because it is shorter (15 compared to 20 words).

Conclusion

Standard scores. This chapter provided the analysis of the scores of students from the six tasks of the *Observation Survey* and the Manitoba Word Reading tasks. Mean scores used to determine whether there were significant differences based on location (urban and rural [farming and northern communities]), age (January- March, April - June, July- September, and October - December), and gender (boys and girls), showed that there were no significant differences for students based on age, however,

there were differences on four of the tasks based on location (urban and rural [farming and northern communities]) and four based on gender (boys and girls). Sets of stanine scores were developed for these tasks and differences examined. Since stanines collapse raw scores into units based on the normal curve, it may be concluded that it is more practical to use one set of stanine scores with all students regardless of location (urban and rural [farming and northern communities]), or gender (boys and girls). That is both location and gender differences were washed out in the stanine transformations.

The Manitoba Word List. The mean scores for students on the two Word Reading tasks, Clay (1966) and Manitoba (2002) showed that the lists could be considered optional alternatives for determining students' ability to read words in isolation. The next chapter includes a discussion of the results of this study, the application for classroom teachers and limitations associated with the research.

CHAPTER FIVE

Results And Discussion

The purpose of this study was to develop norms for Manitoba students using *An Observation Survey of Early Literacy Achievement* (Clay, 1993a, 2002) tasks and to develop a Manitoba Word Reading task based on the reading materials used by Manitoba students in kindergarten and grade one. This study also focused on two related questions:

1. Are there significant differences in the mean performance of students on the *Observation Survey* tasks based on location (i.e. between urban and rural (farming and northern communities), age (i.e. between January to March, April to June, July to September, and October to December), and gender (i.e. between boys and girls)?
2. Are there significant differences in the mean performance of students on the Manitoba and the Clay Word Reading tasks?

This chapter consists of four sections: 1) summary of the findings and discussion, 2) study limitations, 3) implications for future research, and 4) implications for classroom instruction.

Summary of Findings and Discussion

Findings from the first question regarding significant differences in the mean performance of students on the *Observation Survey* tasks based on location (urban and rural ([farming and northern communities])), age (January to March, April to June, July to September, and October to December), and gender (boys and girls) showed the following results.

Location. There were significant differences for urban and rural (farming and northern communities) students in the province on four of the tasks, the *Running Records of Text Reading*, *Clay Word Reading* task, *Hearing and Recording Sounds in Words*, and *Manitoba Word Reading* task. However, differences between the scores occurred in the first three stanines only and placed students in similar rankings according to the range - below average, average, and high progress.

Age. Conventional wisdom suggests that children born in the latter part of the school year would not have had the same amount of time to acquire the literacy skills and knowledge as that acquired by six year-old students born in the first term of the year. However, the analysis of scores divided into quartiles showed that there were no statistically significant differences among task scores based on age in months.

Gender. There were significant statistical differences between the scores for boys and girls on four tasks, *Running Records of Text Reading*, *Writing Vocabulary*, *Hearing and Recording Sounds in Words (Dictation)* and the *Manitoba Word Reading* task. An examination of separate sets of stanine scores developed for boys and girls based on task performance showed that for three of the tasks, *Running Records of Text Reading*, *Hearing and Recording Sounds in Words (Dictation)*, and the *Manitoba Word Reading*, the scores in the average (Stanines 4, 5 and 6) and the high average range (Stanines 7, 8, and 9) were very similar. On the three tasks, the boys appeared to be one stanine grouping behind the girls until stanine 4 or 5 when the performance was comparable. The stanines developed for the *Writing Vocabulary* task showed more variation between

boys and girls across all stanine groupings, with the boys consistently remaining one stanine grouping below the girls.

Seltzer, Choi and Yeow (2002) suggested that evidence of critical importance in monitoring the effectiveness of classroom programs for both individual and groups of students requires the examination of mean rates of change as well as differences in rates of change for various demographic groups. To obtain this evidence, students must be assessed upon school entry. Since this study examined progress at only one point in time, eighteen months after school entry, it was not possible to explore rates of progress.

Nevertheless, Seltzer et al., (2002) demonstrated that although some groups may have entered school with lower achievement levels, the relationship between their initial status and their current status may demonstrate rates of progress that highlight the closing of the achievement gap. It is possible that the boys assessed in this current study may have had lower achievement levels than the girls upon entry and were, in fact, closing this achievement gap.

Seltzer and his colleagues (2002) found that initial achievement status and rates of progress were related to gender. They reported that among students with relatively low initial status, rates of change were appreciably faster for girls than for boys. On the other hand, among students with relatively high initial status, rates of progress tended to be more rapid for boys. Future studies that examined the rates of progress for boys and girls could determine whether alternative stanines should be used. Nevertheless, based on findings from this study, it seems appropriate to suggest that classroom teachers monitor the progress of boys carefully, especially when their achievement scores place them in stanines 1, 2 or 3.

The second major question focused on whether there were significant differences in the mean performance of students on the Manitoba and the Clay Word Reading task. The practical question was whether or not a word list predicated on the reading materials used in the province would be preferable to one developed in New Zealand. The Manitoba Word Reading Test consisted of three word lists of twenty words each. The correlation between the three lists ranged from .881, to .931. The correlation between the Clay and Manitoba word reading measures was .91 indicating a strong correlation between the tasks. This analysis indicates that either list could be considered to measure the ability to read words in isolation. However, the Clay Word Reading task contains three lists of fifteen words and the Manitoba Word Reading task consists of three lists of twenty words. Using the Clay list, therefore, might be more efficient in terms of classroom time.

Study Limitations

The sample of 400 Manitoba students selected randomly from grade one classes including equal numbers of students based on gender and location provided useful data for analysis in this study. The administration of the tasks by teacher leaders and trainers provided consistency in task administration and assured the validity of the data collected. The stanines provided information on a variety of literacy tasks important to establishing student literacy achievement eighteen months after school entry.

The research had the following limitations, however. First, the study examined students' progress at one point in time, eighteen months after school entry. One of the critical areas to be considered is the relationship between the level at which students begin and how rapidly they progress through critical areas of literacy development.

Attending to mean rates of change (Seltzer, et al., 2002) and to differences in rates of change for various groups is critical in monitoring the effectiveness of instruction. It is important to understand the relationship between individual student achievement upon school entry and the rate of progress, as well as how rates of progress vary across groups of students. Clearly much can be learned by moving beyond snapshots of student achievement at single points in time to more comprehensive analyses and summaries of student growth.

Second, although there was an attempt to include students in the sample from schools that were not implementing the Reading Recovery™ program, in the final sample that consisted of 50 schools, only seven schools (six rural [farming and northern communities] and one urban) were not participating in the Reading Recovery™ program. There were 43 students in six rural (farming and northern communities) schools and ten students in the urban school. Two research studies, one by Hill and Crevola (1997) and another by Sylva and Hurry (1995) identified the impact of the Reading Recovery™ program on the literacy instruction of all children in the schools. The authors termed this the "leakage effect." These researchers reported improved achievement results for all students in the early years programs because of the impact of the Reading Recovery™ inservice course for teachers. It is possible that the differences between the performance of urban and rural (farming and northern communities) students can be accounted for by the large number of students in the urban samples who were from Reading Recovery™ schools.

That being said, while the data for 347 students in the study were collected from schools that were already implementing the Reading Recovery™ program, these schools began the implementation of Reading Recovery™ at different times. Eighteen of the

urban schools began the implementation of Reading Recovery™ in 1993, while only 6 of the rural (farming and northern communities) schools started at this time. Six of the urban and 13 of the rural (farming and northern communities) schools started the Reading Recovery™ program in 1998. It is possible that the students in the urban sample scored higher than the students in the rural (farming and northern communities) schools on four of the seven tasks because more of the urban schools had been implementing the Reading Recovery™ program for a longer period and teachers benefited from the professional development associated with Reading Recovery™, that in turn, impacted on student achievement. Future studies should explore the differences in the students' performance based on the time that the school began the implementation of the Reading Recovery™ program.

A third limitation of the study was the restriction to stop administering Running Records at level 20. This decision was made for two reasons: (1) a reading level of 20 demonstrates progress that is considered into the grade two level (see Appendix E), and (2) having the students read higher than level 20 would have increased the time required for administration, approximately thirty-five minutes having been set aside for individual assessment. This decision limited the upper ranges of the stanine scores and thus probably restricted the range of reading scores obtained.

Implications for Future Research

Future studies could include information concerning rates of progress through the early years in order to address the question of how differences in initial status relate to subsequent differences in rates of progress. It is therefore important to gather student achievement data upon school entry to determine the impact of gender (boys or girls) on

progress. Student achievement should also be studied six months after school entry or at reporting times in November and February to ensure that all students receive the instructional support that is necessary. However, if examining rates of progress for different groups of students provides more effective instruction, it may be important to gather initial data at the beginning of Kindergarten. It is also important to note that these initial scores should not be viewed as predictors of future success, but indicators of relative strength that inform instruction. The teacher must thwart negative predictions.

Future studies could also remove the artificial ceiling placed on text levels for conducting the Records of Reading Continuous Text (Running Records). The arbitrary level of 20, set in this study, that was considered appropriate for the month of February might have been too low. Many of the students assessed in this study were able to read this level of text with ease. As a result, the spread of scores obtained for this task in this research was limited.

Implications for Classroom Instruction

The stanines developed in this study provide useful information that teachers may use to gauge the achievement of grade one students, mid-point in the year. If the average performance of the class lags behind the achievement levels identified in the average stanine groupings (Stanines 4, 5, and 6), teachers, in consultation with administrators and other support services personnel, may want to examine whether instructional changes are required.

The stanines also provide information addressing the progress of individual students. It is critical that students develop a range of knowledge and skill in relation to literacy. Using these stanine scores, classroom teachers can determine areas of relative

strength and areas of concern, which can be used to plan further, appropriate instruction. It is especially important to address the needs of students whose scores are in the lowest stanine groupings, stanines 1, 2, and 3. Over time, without appropriate instruction, these students could fall farther and farther behind their classmates. This study has also highlighted the need for an examination of rates of progress for students, particularly, boys in the lowest stanine groupings. It appears that these students may be more at-risk and, consequently, need to be monitored more carefully so that timely adjustments can be made in instruction.

The scores on the two writing tasks included in the *Observation Survey, Hearing and Recording Sounds in Word*, and *Writing Vocabulary* illustrate student knowledge in two ways, the ability to hear and record sounds in words or to write known words quickly. The scores on the *Hearing and Recording Sounds in Words* task showed that students in stanine grouping of 4 to 9 had good knowledge of the phoneme-grapheme relationships with many students reaching the maximum score of 37. In the *Writing Vocabulary* task, however, the average-band students wrote between 23 and 46 words. Research suggests that if a young student can write about 45 different words accurately and quickly, s/he can represent most of the letter-sound associations in the English language and can use the most frequent and regular spelling patterns (Clay, 2002). Classroom teachers should be aware that the average students in this study are gaining this knowledge but must continue to grow in this area.

Conclusions

The analyses conducted in this study showed that although there were significant differences between student achievement based on location (urban and rural [farming

and northern communities]) and gender (boys and girls) on four of the seven Clay's *Observation Survey* (1993a, 2002) tasks, a separate set of stanine scores were not necessary. There were no significant differences in performance on any of the *Observation Survey* tasks based on age. The analysis of the mean scores for the Manitoba and Clay Word Reading tasks demonstrated that there was no difference between the measurement qualities of the two tasks. Manitoba teachers may use the new Word Reading task developed in this study which was based on reading materials used in the classrooms, or continue to use the Clay Word Reading task, which has the advantage of fewer words.

This norming study confirmed the findings of others (Bourque, 2002; Clay, 2002; Pinnell, Lyons, Young & DeFord, 1987) that the progress of Manitoba students eighteen months after school entry was similar to performance elsewhere. Students, after eighteen months of schooling, had made remarkable progress on all the *Observation Survey* tasks and were competent grade one literacy learners. It was also found that some students were falling behind their more average peers and required extra instruction to enhance their literacy learning.

This study also confirmed the decision to use maximum scores on all of the *Observation Survey* tasks to confirm the decision to discontinue individual tutoring sessions in the Reading Recovery™ program. In the Writing Vocabulary task, for example, the recommended minimum of 40 to 45 different words would place students receiving Reading Recovery™ tuition at the top of the average band and suggests that such students are capable enough to sustain continued learning in the classroom. The decision to recommend level 16 as a minimum for discontinuing individualized instruction will also place Reading Recovery™ students at the top of the average range,

once again supporting their continued success as they work more independently in the classroom. Stanine scores in the study also suggest that by the end of the year, Reading Recovery™ students must achieve high standards in both text reading and writing vocabulary if their continued success within the average band is to be expected.

This study also reinforced the importance of monitoring student performance across a variety of tasks and the need to provide additional instructional support for students who are not as successful in learning to be literate. The stanines are helpful for teachers in identifying areas of strength and areas needing further attention both for individuals and groups of students.

Lastly, this study has shown that most grade one students in Manitoba schools are demonstrating high levels of achievement in a variety of critical literacy learning areas. However, at mid-point in the school year, many students were early readers. These students continue to need teachers who will monitor their achievement on a variety of tasks and based on this information, institute high quality literacy instruction.

References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Allington, R., & Cunningham, P. (1986). *Schools that work*. New York: Harper Collins.
- Anderson, R.C., Hiebert, E.H., Scott, J. A., & Wilkinson, I. A. (1984). *Becoming a nation of readers*. Washington, D.C.: National Institute of Education.
- Applegate, M., Quinn, K., & Applegate, A. (2002). Levels of thinking required by comprehension questions in informal reading inventories. *The Reading Teacher*, 56, 174-180.
- Ball, E.W., & Blachman, B. A. (1991). Does phoneme awareness training in Kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 26, 49-66.
- Barr, R.C., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.
- Bauman, J. (1988). *Reading assessment: An instructional decision-making perspective*. New York: Macmillan Publishing Company.
- Beach, J. (1992). Literature's place in Language Arts Education: New Trends in a historical perspective. *Language Arts*, 69, 550-556.
- Bear, D. R., & Barone, D. (1989). Using children's spellings to group for word study and directed reading in the primary classroom. *Reading Psychology*, 10, 275-292.
- Beck, I., & Juel, C. (1995). The role of decoding in learning to read. *American Educator*, 19, 8-25.
- Bickel, D., Zigmon, N., & Strayhorn, B. (1991). Chronological age at entrance to first grade: Effects on elementary school success. *Early Childhood Research Quarterly*, 6, 105-117.
- Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, 6, 75-96.
- Bird, D. G. (1972). *An analysis of the grapheme-phoneme patterns in the vocabulary of library books and textbooks frequently read by early primary children*. Unpublished doctoral dissertation, University of Pittsburgh.

- Blachman, B. (1984). The relationships between rapid naming ability and language analysis skills to kindergarten and first grade reading achievement. *Journal of Educational Psychology*, 76, 614-622.
- Blanchard, J. S. (1980). Preliminary investigation of transfer between single-word decoding ability and contextual reading comprehension of poor readers in grade six. *Perceptual and Motor Skills*, 51, 1271-1281.
- Bourque, G. (2002). *Le Sondage d'observation en lecture-écriture*. Toronto: Cheneliere/McGraw Hill.
- Bradley, L., & Bryant, P. E. (1983). *Rhyme and reason in reading and spelling*. Ann Arbor: University of Michigan Press.
- Braunger, J., & Lewis, J. (1998). *Building a knowledge base in reading*. Newark, DE: International Reading Association.
- Bronfenbrenner, U., & Morris, P. (1998). The ecology of developmental processes. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), *Handbook of child psychology* (5th ed., pp. 983 - 1028). New York: Wiley.
- Bryant, P. E., Bradley, L. L., MacLean, M., & Crossland, J. (1989). Nursery rhymes, and reading. *Journal of Educational Psychology*, 87, 488-503.
- Bryne, B., & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1 year follow-up. *Journal of Educational Psychology*, 85, 104-111.
- Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection and learning to read. *Developmental Psychology*, 26, 429-438.
- Calfee, R.C., & Pintontkowsky, D.C. (1981). The reading diary: Acquisition of decoding. *Reading Research Quarterly*, 16, 346-373.
- Cazden, C. (2001). *Classroom discourse: The language of teaching*. Portsmouth, NH: Heinemann Educational Books.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. New York: American Heritage.
- Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.
- Chall, J. S. (1989). Learning to read: The great debate twenty years later. A response to "Debunking the great phonics myth." *Phi Delta Kappan*, 71, 521-538.

- Chall, J. S. (1996). *Learning to read: The great debate (revised)*. New York: McGraw-Hill.
- Clark-Carter, D. (1997). *Doing quantitative psychological research from design to report*. East Essex, UK: Psychology Press.
- Clay, M. M. (1966). *Emergent reading behaviour*. Unpublished doctoral dissertation, University of Auckland.
- Clay, M. M. (1968). A syntactic analysis of reading errors. *Journal of Verbal Learning and Verbal Behavior*, 7, 434-438.
- Clay, M. M. (1972). *The early detection of reading difficulties: A diagnostic survey with recovery procedures*. Exeter, NH: Heinemann.
- Clay, M. M. (1975) *What did I write*. Auckland: Heinemann.
- Clay, M. M. (1985). *The early detection of reading difficulties*. Auckland: Heinemann.
- Clay, M. M. (1979). *Reading: The patterning of complex behaviour*. Auckland: Heinemann.
- Clay, M. M. (1989). Concepts about print in English and other languages. *The Reading Teacher*, 42, 268-276.
- Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1993a). *The observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1993b). *Reading Recovery: A guidebook for teachers in training*. Auckland: Heinemann.
- Clay, M. M. (1998). *By different paths to common outcomes*. York, MA: Stenhouse.
- Clay, M. M. (2001). *Change over time in children's literacy development*. Portsmouth, NH: Heinemann.
- Clay, M. M. (2002). *The observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Clough, M., McIntyre, J., & Cowey, W. (1990). *The Canberra word reading task*. Canberra, AU: University of Canberra.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing*. (pp. 151-216). London: Academic Press.

- Cunningham, A. E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Psychology*, 50, 429-444.
- Cunningham, A. E., & Stanovich, K. E. (1993). Children's literacy environments and early word recognition skills. *Reading and Writing: An Interdisciplinary Journal*, 5, 193-204.
- Dale, E. (1943). *Easy word list*. Columbus, OH: The Ohio State University.
- Daneman, M. (1991). Individual differences in reading skills. In R. Barr, M.R. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 512-538). New York: Longman.
- Day, K., & Day, H. D. (1984). Kindergarten knowledge of print conventions and later school achievement: A five year follow-up. *Psychology in the Schools*, 21, 393-396.
- Dickinson, D.K., & Tabors, P.O. (1991). Early literacy: Linkages between home, school, literacy achievement at age five. *Journal of Research in Childhood Education*, 6, 30-46.
- Dolch, E. (1960). Needed vocabulary. *Elementary English*, 28, 530-534.
- Duncan, S. (1999). *Reading and writing relationship: Knowledge and processes*. Unpublished doctoral thesis, University of Auckland.
- Dunkeld, C. (1970). *The validity of the informal reading inventory for the designation of instructional reading levels*. Unpublished doctoral dissertation, University of Illinois.
- Durkin, D. (1966). *Children who read early*. New York: Teacher's College Press.
- Durrell, D. (1958). First grade reading success study. *Journal of Education*, 140, 23-41.
- Ehri, L. C. (1991). Learning to read and spell words. In L. Rieben & C.A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 57-73). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition*, (pp. 107-143). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ehri, L. C. (1994) Development of the ability to read words: Update. In R. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 323-358). Newark, DE: International Reading Association.

- Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of research in reading*, 18, 116-125.
- Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme & R. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 87-111). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly*, 20, 163-179.
- Farr, R. (1973). *Iowa Silent Reading Tests - Guide for interpretation and use*. New York: Harcourt, Brace, Jovanovich.
- Farr, R., & Summers, E. (1969). Guide to tests and measuring instruments in reading. In *Reading: What can be measured?* Newark: DE: International Reading Association.
- Fountas, I., & Pinnell, G. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann Educational Books.
- Fox, B., & Routh, D. K. (1984). Phonemic analysis and synthesis as word attack skills: Revisited. *Journal of Psycholinguistic Research*, 4, 331-342.
- Frith, U. (1985). Beneath the surface of phonetic dyslexia. In K. Patterson, J. Marshall, and M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading*. (pp. 301-330). London: Lawrence Erlbaum.
- Gates, A. I. (1925). *A reading vocabulary for the primary grades*. New York: Teachers College Press.
- Gates, A. (1961). Sex differences in reading ability. *Elementary School Journal*, 61, 431-434.
- Gentner, D. (1982). *Why nouns are learned before verbs: Linguistic relativity versus natural partitioning*. Center for the Study of Reading. Technical Report No. 257. Urbana, IL: University of Illinois.
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton Century Crofts.
- Goodman, K. S. (1985). Unity in reading. In H. Singer & R.B. Ruddell (Eds.) *Theoretical models and processes of reading* (3rd ed., pp. 813-840). Newark: DE: International Reading Association.
- Goodman, K. (1986). *What's whole in whole language?* Portsmouth, NH: Heinemann.

- Goodman, K. S. (1994). Reading, writing and written texts: A transactional-sociopsycholinguistics view. In R.B. Ruddell, M.R. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading*. (4th ed., pp. 1093-1130). Newark, DE: International Reading Association.
- Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, 42, 73-83.
- Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology*, 40, 323-340.
- Goswami, U. (1998). The role of analogies in the development of word recognition. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy*. Mahwah, NJ: Erlbaum.
- Goswami, U., & Byrant, P. (1990). *Phonological skills and learning to read*. East Sussex, UK: Lawrence Erlbaum.
- Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society*, 20, 179-196.
- Gough, P. B., & Juel, C. (1991). The first stages of word recognition. In L. Rieben, & C.A. Perfetti (Eds.), *Learning to read: Basic research and its implications*, (pp. 47-56). Hillsdale, NJ: Lawrence Erlbaum.
- Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling and the orthographic cipher. In P. B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition*, (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum.
- Greene, H. A. (1954). *The new Iowa spelling scale*. Iowa City, IA: The University of Iowa.
- Griffith, P. L., & Olson, M. W. (1992). Phonemic awareness helps beginning readers break the code. *The Reading Teacher*, 45, 516-523.
- Hamre, B., & Pianta, R. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72, 625-38.
- Harris, T., & Hodges, R. (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.
- Harris, A., & Jacobson, M. (1972). *Basic elementary reading vocabulary*. New York: MacMillan.

- Heimberger, M. J. (1972). *Analysis and comparison of the vocabularies in textbooks and library books frequently read by early primary children*. Unpublished doctoral dissertation, University of Pittsburgh.
- Herman, P. A. (1985). The effects of repeated readings on reading rate, speech pauses and word recognition accuracy. *Reading Research Quarterly*, 20, 553-565.
- Hill, P., & Crevola, C., (1997). *The literacy challenge in Australian schools*. Sydney: Mercer House.
- Hohn, W. E., & Ehri, L.C. (1983). Do alphabetic letters help prereaders acquire phonemic segmentation skill? *Journal of Educational Psychology*, 75, 752-762.
- Holdaway, D. (1980). *Independence in reading* (2nd ed.) Portsmouth, NH: Heinemann.
- Holdaway, D. (1979). *The foundations of literacy*. Sydney, AU: Ashton Scholastic.
- Horn, E. (1925). The commonest words in the spoken vocabulary of children up to and including six years of age. *National Society for the Study of Educational Yearbook, Part 1*, 24, 185-198.
- Janda, L. (1998). *Psychological testing: Theory and application*. Needham Heights, MA: Allyn & Bacon.
- Johns, J. (1980). First graders' concepts about print. *Reading Research Quarterly*, 15, 529-548.
- Johnson, P. (1992). *Constructive evaluation of literate activity*. New York: Longman.
- Johnson, D. D., & Baumann, J. F. (1984). Word identification. In P. D. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research*, (pp. 583-608). New York: Longman.
- Johnston, R. S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in prereaders: The nature of the relationship. *Reading and Writing: An Interdisciplinary Journal*, 8, 217-234.
- Jones, L. V., & Wepman, J. M. (1966). *A spoken word count: Adults*. Chicago: Language Research Associates.
- Juel, C. (1991). Beginning reading. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research, Vol. II* (pp. 759-788). New York, NY: Longman.
- Juel, C. (1994). *Learning to read and write in one elementary school*. New York: Springer-Verlag.

- Juel, C., Griffith, P. L., & Gough, P. S. (1985). Reading and spelling strategies of first grade children. In J. A. Niles & R. Lalik (Ed.), *Issues in literacy: A research perspective* (pp. 306-309). Rochester, NY: National Reading Conference.
- Juel, C., & Minden-Cupp, C. (1998). *Learning to read words: Linguistic units and strategies*. CIERA Report #1-008. Ann Arbor: University of Michigan.
- Juel, C., & Roper-Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly*, 20, 134-152.
- Langenberg, D. (Ed.) (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC: U.S. Department of Health and Human Services.
- Lesgold, A. M., & Resnick, L. B., (1982). How reading disabilities develop: Perspectives from a longitudinal study. In J.P. Das, R. Mulcahy, and A. E. Wall (Eds.), *Theory and research in learning disability*. New York: Plenum.
- Lesgold, A. M., Resnick, L. B., & Hammond, K. (1985). Learning to read: A longitudinal study of word skill development in two curricula. In G. E. MacKinnon and T. G. Walker (Eds.), *Reading research: Advances in theory and practice* (Vol. 4, pp.107 - 138). New York: Academic Press.
- Leu, D. J. (1985). Oral reading error analysis: A critical review of research and application. *Reading Research Quarterly*, 17, 420-437.
- Literacy Task Force (1999). *Report of the literacy task force*. Wellington, NZ: Ministry of Education.
- Lomax, R. G., & McGee, L. M. (1987). Young children's concepts about print and reading: Toward a model of word recognition. *Reading Research Quarterly*, 22, 237-256.
- Lundberg, I. (1987). Are letters necessary for the development of phonemic awareness? *Cahiers de Psychologie Cognitive*, 7, 472-475.
- Lundberg, I., Frost, J., & Petersen, O. P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 264-284.
- Lummis, M., & Stevenson, H. (1990). Gender differences in beliefs and achievement: A cross-cultural study. *Developmental Psychology*, 26, 254-263.
- Lyman, H. B. (1991). *Test scores and what they mean*. (5th ed.) Boston: Allyn & Bacon.
- MacFarlane, L. (2001). *Gender differences in reading achievement and early literacy experiences*. (ERIC Document Reproduction Service No. ED451476)

- Mann, V. A., Shankweiler, D., & Smith S. T. (1984). The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language*, 11, 627-643.
- Mason, J. (1980). When do children begin to read: An exploration of four year old children's letter and word reading competencies. *Reading Research Quarterly*, 15, 203-227.
- Mason, J., & Allen, J. (1986) A review of emergent literacy with implications for research and practice in reading. *Review of Research in Education*, 13, 3-47.
- Marsh, G., Desberg, D., & Cooper, J. (1977). Developmental changes in reading strategies. *Journal of Reading Behavior*, 9, 391-394.
- McFalls, E., Schwanenflugel, P., & Stahl, S. (1994). *Influence of word meaning on the acquisition of a reading vocabulary in second-grade children*. Reading Research Report No. 25, National Reading Research Center. Athens, GA.
- McGee, L., & Purcell-Gates, V. (1997). Conversations: So what's going on in research in emergent literacy? *Reading Research Quarterly*, 32, 310-318.
- McNaughton, S. (1995). *Patterns of emergent literacy processes of development and transitions*. Auckland: Oxford University Press.
- Meek, M. (1982). What counts as evidence in theories of children's literature? *Theory into practice*, 21, 284-292.
- Meek, M. (July, 1992). *Literature and literacy: The real book approach for children learning to read*. Paper presented at the Annual Conference of the International Association of School Librarians. Belfast, Northern Ireland.
- Meisels, S. (1992). Doing harm by doing good: Latrogenic effects of early childhood enrollment and promotion policies. *Early Childhood Research Quarterly*, 7, 155-174.
- Ministry of Education, (1998). *School Entry Assessment: The first national picture*. Wellington, NZ: Ministry of Education.
- Morton, E. (1969). A structural comparison of English and Tongan. *TESL Reporter* 2, 30-42.
- Moustafa, M. (1997). *Beyond traditional phonics: Research discoveries and reading instruction*. Portsmouth, NH: Heinemann.
- Neuman, S., & Rosko, K. (1993). *Language and literacy in the early years: An integrated approach*. New York: Harcourt Brace & Jovanovich.

- Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. In C. Hulme and R. Joshi (Eds.), *Reading and spelling: Development and disorders*. (pp. 343-367). Mahwah, NJ: Lawrence Erlbaum.
- Paris, S. G., Wasik, B., & Turner, J. (1991). The development of strategic readers. In R. Barr, M. I. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. II, pp. 609-640). New York: Longman.
- Pany, D., & Jenkins, J. R. (1978). Learning words meanings: A comparison of instructional procedures and effects on measures of reading comprehension with learning-disabled students. *Learning Disabilities Quarterly*, 1, 21-32.
- Patterson, K. (1981). Neuropsychological approaches to the study of reading. *British Journal of Psychology*, 72, 151-174.
- Pearson, D. (2000). *Standard tools for assessment: Tools for reform of early reading instruction*. Paper presentation at 2000 Michigan Reading Conference. Retrieved June 15, 2002, from <http://ed-web3.educ.msu.edu/pearson/pdpaper>.
- Perfetti, C.A. (1985). *Reading Ability*. New York: Oxford University Press.
- Perfetti, C.A., Beck, I., Bell, L., & Hughes, C. (1987). Children's reading and development of phonological awareness. *Merrill Palmer Quarterly*, 33, 39-75.
- Pinnell, G. S., Lyons, C., Young, P., & DeFord, D. (1987). *The Reading Recovery program in Ohio*, Volume VI (Technical Report). Columbus, OH: Ohio State University.
- Pinnell, G., McCarrier, A., & Button, K. (1990). *Constructing literacy in urban kindergartens: Progress report on kindergarten early literacy project*. Report No. 10. Columbus, OH: Ohio State University.
- Pinnell, G., Lyons, C., DeFord, D., Bryk, A., & Seltzer, M. (1994). Comparing instructional models for literacy education of high-risk first graders. *Reading Research Quarterly*, 29, 18-39.
- Pressley, M. (1998). *Reading instruction that works: The case for balanced teaching*. New York: The Guilford Press.
- Randell, B. (1999). Shaping the PM story books. *The Running Record*, 11, 1-12.
- Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard Educational Review*, 41, 1-34.
- Richgels, D. (1995). Invented spelling ability and the printed word learning in kindergarten. *Reading Research Quarterly*, 27, 125-138.

- Richgels, D., & Burns, J. (1989). An investigation of task requirements associated with the invented spellings of four-year-olds with above average intelligence. *Journal of Reading Behavior*, 21, 1-14.
- Rinsland, H. (1947). *A basic vocabulary of elementary school children*. New York: MacMillan.
- Rumelhart, D. E. (1994). Towards an interactive model of reading. In R. B. Ruddell, M.R. Ruddell, & H. Singer, (Eds.). *Theoretical models and processes of reading* (4th ed., pp. 864-894). Newark: DE: International Reading Association.
- Rowe, K. (August 2000). *Problems in education of boys and exploring real effects from evidence based research: Useful findings in teaching and learning for boys and girls*. Paper presented at Teaching Boys Developing Fine Men Conference, Brisbane, Australia.
- Samuels, J. (1994). Towards a theory of automatic information processing in reading, revisited. In R. Ruddell, M. Ruddell, & H. Singer (Eds.). *Theoretical models and processing of reading* (4th ed., pp. 816-837). Newark, DE: International Reading Association.
- Samuels, J., Schermer, N., & Reinking, D. (1992). Reading fluency: Techniques for making decoding automatic. In S.J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction?* Newark, DE: International Reading Association.
- Sartain, H.W. (August, 1981). *Frequency count of English words appearing in books read in elementary schools*. Paper presented at the Annual meeting of the European Conference on Reading. (ERIC Document Reproduction Service No. ED216326)
- Sayeski, K., Burgess, K., Pianta, R., & Lloyd, J., (2001). *Literacy behaviors of preschool children participating in an early intervention program*. CIERA Report #2-014. Ann Arbor: University of Michigan.
- Scarborough, H. (1991). Early syntactic development of dyslexic children. *Annals of Dyslexia*, 41, 207-220.
- Scarborough, R., Bruns, R., & Frazier, A. (1957). Improvement of reading through ability-level assignments. *School Review*, 165, 474-486.
- Schrader, B. (1990). Disabled readers as able teachers: A cross-age tutoring project. *Academic Therapy*. 25, 589-97.
- Schreiber, P.A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice*, 30, 158-164.

- Seltzer, M., Choi, K., & Yeow, T. (2002). *Examining relationships between where students start and how rapidly they progress: Implications for constructing indicators that help illuminate the distribution of achievement within schools*. CSE Technical Report, Los Angeles: University of California.
- Share, D. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72, 95-129.
- Shepard, L., & Smith, P. (1988). Escalating academic demand in kindergarten: Counterproductive policies. *The Elementary School Journal*, 89, 135-145.
- Simpson, M. M. (1962). *Suggestions for teaching in reading in infant classes*. Wellington: Department of Education.
- Singer, H. (1994). The substrata-factor theory of reading. In R. B. Ruddell, M.R. Ruddell, & H. Singer, (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 895-927). Newark: DE: International Reading Association.
- Slavin, R., Karweit, N., & Wasik, B. (1993). Preventing early school failure: What works? *Educational Leadership*, 50, 10-18.
- Smith, F. (1988). *Understanding reading* (4th ed). New York: Holt, Rinehart & Winston.
- Smith, J., & Martin, D. (1997). Investigating literacy and pre-literacy skills in Panjabi/English schoolchildren. *Educational Review*, 49, 181-197.
- Snow, C. E., Burns, M., & Griffith, S. (1998). *Preventing reading difficulties in young children*. Washington, D.C.: Committee on the Prevention of Reading Difficulties in Young Children, Commission on Behavioral and Social Sciences and Education, National Research Council.
- Stahl, S.A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology*, 86, 221-234.
- Stahl, S.A., & Murray, B. A. (1998). Issues involved in defining phonological awareness and its relation to early reading. In J. Metsala and L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 65-87). Mahwah, NJ: Erlbaum.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of fluency. *Reading Research Quarterly*, 16, 32-71.
- Stanovich, K. E. (1985). Explaining the variance in reading ability in terms of psychological processes: What have we learned? *Annals of Dyslexia*, 35, 67-96.

- Stanovich, K. E. (1986). The Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.
- Stanovich, K. (1991). Word recognition: Changing perspectives. In Barr, R., Kamil, M. L., Mosenthal, P. & P. D. Pearson, (Eds). *Handbook of Reading Research* (Vol. II, pp. 418-452). New York: Longman.
- Stanovich, K. E. (1992). Speculation on the causes and consequences of individual differences in early reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp 307-342). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stanovich, K., Cunningham, A., & Cramer, B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Psychology*, 38, 175-190.
- Stanovich, K. E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24, 402-433.
- Stanovich, K., West, R., & Freeman, D. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive compensatory model. *Journal of Experimental Child Psychology*, 32, 185-199.
- Stroud, J., & Lindquist, E. (1942). Sex differences in achievement in elementary and secondary schools. *Journal of Educational Psychology*, 33, 657-667.
- Sulzby, E., & Teale, W. H., (1991). Emergent literacy. In R. Barr, M. L. Kamil, P. Mosenthal, and P. D. Pearson (Eds.), *Handbook of Reading Research* (pp. 727-757). New York: Longman.
- Summers, E. (1984). A paradigm for developing Quick Word Reading Tasks for reading assessment in secondary grades. *B.C. Journal of Special Education*, 8, 11-32.
- Strickland, D., & Morrow, L. (Eds.) (1989). *Emerging literacy: Young children learn to read and write*. Newark, DE: International Reading Association.
- Stroud, J., & Lindquist, E. (1942). Sex differences in achievement in elementary and secondary schools. *Journal of Educational Psychology*, 33, 657-667.
- Sylva, K., & Hurry, J., (1995). *Early intervention in children with reading difficulties: An evaluation of Reading Recovery and a phonological training*. Oxford: Blackwell Publishers.

- Teale, W., & Sulzby, E. (1986). Emergent literacy as a perspective for examining how young children become readers and writer. In W. H. Teale and Sulzby (Eds.), *Emergent literacy: Reading and writing* (pp.i-xxv). Norwood, NJ: Albex
- Thorndike, E.L. (1921). *Teachers word book*. New York: Teachers College, Columbia University.
- Turner, W.E., & Nesdale, A. R. (1985). Phonemic segmentation skill and beginning reading. *Journal of Educational Society*, 77, 417-527.
- Turner, W.E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly*, 23, 134-158.
- Torgeson, K. (1986). Computers and cognition in reading: A focus on decoding fluency. *Exceptional Children*, 53, 157-162.
- Treiman, R. (1992). The role of intersyllabic units in learning to read and spell. In P. Gough, L.C. Ehri, and R. Treiman (Eds.), *Reading Acquisition* (pp. 65-106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Treiman, R. & Baron, J. (1981). Phonemic-analysis training helps children benefit from spelling-sound rules. *Memory and Cognition*, 11, 382-389.
- Vellutino, F. R., & Denckla, M. B. (1991). Cognitive and neuropsychological foundations of word identification in poor and normally developing readers. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of Reading Research* (Vol. II, pp. 571-608). New York: Longman.
- Vellutino, F.R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly*, 33, 321-363.
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, (Eds).) Cambridge, MA: Harvard University Press.
- Weaver, C. (1994). *Reading process and practice: From socio-psycholinguistics to whole language*. Portsmouth, NH: Heinemann.
- Weir, B. (1989). A research base for prekindergarten literacy programs. *The Reading Teacher*, 42, 456-460.
- West, P. (1999). Boys' underachievement in school: Some persistent problems and some current research. *Issues in Educational Research*, 9, 33-54.

- Winsor, P., & Pearson, D. (1992). *Children at risk: their phonemic awareness development in holistic instruction*. (ERIC Document Reproduction Service. ED3435209)
- Wood, D. (1998). *How children think and learn*. Oxford, UK: Blackwell Publishers.
- Yopp, H., & Yopp, R. (2000). Supporting phonemic awareness development in the classroom. *The Reading Teacher*, 54, 130-143.
- Yaden, D., Rowe, D., & McGillivray, L. (1999). *Emergent literacy: A polyphony of perspectives*. CIERA Report R-1005. Ann Arbor: University of Michigan.

List of Appendices

A	List of Titles Used for Word Frequency Count	125
B	Total Frequency Count	128
C	Comparison to Other Word Lists	131
D	Word Reading Task - Score Sheet	132
E	Expected Levels of Text Reading	134
F	Protocol for Administering Running Records	135
G	List of Titles for Assessment	136
H	Overall Stanine Scores	139

Appendix A:
List of Titles Used for Word Frequency Count

PM Starters Level 1 and 2 - Magenta

Me	The go-karts
Mom	In the shopping cart
Dad	Climbing
A house	The shopping mall
Big things	Playing
Little things	Look at me
Dressing up	Pets
Time for Dinner	The way I go to school
At the zoo	We go out
Moms and Dads	The skier
Out in the Weather	The farm in the spring
Cat and Mouse	We can run
Where are the babies?	Four ice-creams
Packing my bag	My little dog
The rock pools	At the library
Stop!	Looking down
Sally's new shoes	Fishing
Ben's red car	The pencil
Ball games	Can you see the eggs?

PM Library Red Level Set A and B

The Photo Book	The little Snowman
Hedgehog is hungry	A birthday cake for Ben
Wake up, Dad	The baby owls
Tiger, Tiger	The bumper cars
The lazy pig	The flower girl
The merry-go-round	Hide and seek
Baby lamb's first drink	A home for little Teddy
Sally and the daisies	Where is Hannah?
The big kick	Hot dogs
Kitty and the birds	Ben's teddy bear
Ben's treasure hunt	Lizard loses his tail
Father Bear goes fishing	Tom is brave

PM Library Yellow Level Pack A & B

Where are the sunhats?	Blackberries
Brave Father Mouse	Mumps
The hungry kitten	Sally's beans
Lucky goes to dog school	Ben's Dad

The new baby
 Hermit Crab
 Little Bulldozer
 Friend for little white rabbit
 Baby Hippo
 Sally and the sparrows
 Soccer at the park
 Tiny and the big wave

Baby Bear goes fishing
 Seagull is clever
 Sally's red bucket
 Fire! Fire!
 Jolly Roger, the pirate
 Choosing a puppy
 A lucky day for little dinosaur
 Snowy gets a wash

PM Library Blue level Pack A & B

Magpie's baking day
 Sally's friends
 Baby Bear's present
 Jane's car
 The house in the tree
 Mushrooms for dinner
 Cows in the garden
 Locked out
 Little Bulldozer helps again
 Lost at the fun park
 Tim's favorite toy
 Teasing Dad

The lion and the rabbit
 Honey for Baby Bear
 The best cake
 Tabby in the tree
 The Christmas tree
 Come on, Tim
 Late for soccer
 The lion and the mouse
 The duck with the broken wing
 Birthday balloons
 Chug the tractor
 Tiger runs away

PM Library Green Level Pack A & B

The Naughty Ann
 The clever penguins
 House-hunting
 Mrs Spider's beautiful web
 Ten little garden snails
 The little red bus
 The flood
 The waving sheep
 The flying fish
 Snow on the hill
 Father Bear's surprise
 Joey

Brave Triceratops
 Pete Little
 Candle-light
 Ben's tooth
 The fox who foxed
 The island picnic
 Pepper's adventure
 The cross-country race
 The rescue
 The babysitter
 After the flood
 Try again, Hanan

PM Library Orange Fiction Pack
 Traditional Tales and Plays

The dinosaur chase
 Jack and Chug
 Toby and the Big Tree

The Biggest Fish
 Toby and BJ
 The toy farm

Sarah and the Barking Dog
Just One Guinea Pig
Toby and the Big Red Van
Rebecca and the Concert
The Careful Crocodile
The Busy Beavers
Chicken-Licken
The Three Little Pigs
The Tale of the Turnip

Pterosaur's Long Flight
Jessica in the Dark
Mitch to the Rescue
Lost in the Forest
Skates for Luke
Two Little Goldfish
The Gingerbread Man
The Little Red Hen
The Three Billy Goats Gruff

Appendix B:

Total Frequency Count

Frequency Count of PM Storybook Levels							Total
Word	Magenta	Red	Yellow	Blue	Green	Orange	
the	195	113	167	260	279	652	1666
said	41	74	119	204	150	155	743
and	34	40	70	111	166	264	685
to	26	19	57	83	130	210	525
a	83	47	36	61	115	119	461
is	146	103	90	34	33	33	439
I	83	42	21	99	47	56	348
in	40	36	41	28	48	91	284
he	8	15	38	60	65	86	272
here	98	74	24	30	20	13	259
up	29	63	33	35	34	59	253
little	14	19	51	32	41	86	243
on	24	21	23	39	39	82	228
went	15	22	48	43	50	41	219
at	36	32	29	46	32	44	219
look	62	41	39	26	23	19	210
she	20	3	25	36	49	73	206
Dad	17	47	42	27	25	38	196
you	0	0	0	70	51	72	193
for	5	32	27	31	39	49	183
they	8	0	18	42	42	64	174
are	28	27	33	38	16	32	174
Mom	28	31	33	33	31	17	173
it	20	0	29	43	9	70	171
come	19	29	27	27	28	41	171
my	36	32	25	22	27	23	165
down	14	9	33	33	27	46	162
big	8	2	32	31	18	66	157
we	57	3	11	9	20	41	141
can	31	21	12	18	12	38	132
go	30	9	17	33	19	19	127
bear	0	17	30	47	32	0	126
was	0	0	0	0	30	95	125
with	21	0	27	19	17	38	122
looked	0	20	21	23	28	26	118
ran	15	0	20	23	24	31	113
of	0	0	4	15	19	71	109
me	21	8	16	34	10	20	109
am	62	23	4	10	8	2	109
baby	8	19	25	27	19	5	103
out	0	0	7	12	27	56	102
all	0	0	20	10	27	42	99
will	0	0	16	29	29	21	95
not	5	0	17	27	18	26	93
going	8	3	18	17	21	26	93
I'm	0	2	22	31	24	13	92
Mother	12	23	16	9	13	18	91
where	15	12	23	17	8	12	87
like	16	0	24	28	10	9	87
comes	21	13	40	2	2	8	86
get	0	0	6	33	21	22	82
but	0	0	0	17	28	36	81
too	8	7	14	15	12	22	78
see	7	11	11	10	18	16	73
back	8	0	5	10	25	23	71
no	0	7	14	9	12	26	68
this	21	4	10	5	19	8	67
her	15	0	3	14	7	27	66

Frequency Count of PM Storybook Levels							
Word	Magenta	Red	Yellow	Blue	Green	Orange	Total
came	0	0	0	17	24	20	61
shouted	0	13	16	13	6	11	59
day	7	0	6	11	18	17	59
help	0	0	11	15	15	16	57
dog	7	22	7	8	0	13	57
away	0	3	14	11	17	11	56
hungry	0	18	18	5	10	4	55
home	14	0	20	9	2	10	55
looking	0	9	6	15	9	13	52
some	6	0	3	24	17	1	51
Father	0	0	20	5	12	12	49
way	0	1	3	27	10	6	47
after	0	0	16	12	9	5	42
one	0	0	0	12	24		36
did	0	0	0	14	21		35
oh	0	10	4	11	7		32
this	0	0	0	12	19		31
his	0	0	13	12	6		31
asleep	0	22	7	1	1		31
car	0	7	6	16	1		30
Mrs.	0	0	0	5	23		28
saw	0	0	0	14	13		27
got	0	0	6	11	10		27
back	0	0	0	2	25		27
thank	0	5	6	13	2		26
my	0	0	0	22	4		26
sleep	0	5	17	1	2		25
now	0	0	0	10	13		23
your	0	0	7	5	11		23
play	7	0	13	2	1		23
have	0	0	0	8	14		22
good	0	1	3	6	12		22
us	13	0	4	1	2		20
wake	0	0	17	2	1		20
so	0	0	1	1	17		19
coming	0	0	0	7	12		19
has	7	0	0	5	7		19
good	0	0	3	4	11		18
into	0	1	0	15	2		18
yes	0	0	0	11	7		18
goes	8	2	2	0	6		18
then	0	0	0	10	6		16
children	0	0	0	11	5		16
make	0	0	0	13	3		16
way	0	0	3	3	10		16
who	0	0	4	10	2		16
liked	0	0	6	8	1		15
made	0	0	0	11	4		15
can't	0	0	0	11	3		14
fire	0	0	14	0	0		14
there	0	0	0	0	14		14
had	0	0	0	5	9		14
them	8	0	0	0	6		14
liked	0	0	0	7	7		14
jumped	0	0	0	6	7		13
helped	0	0	0	8	5		13
again	0	0	0	3	10		13
thank	0	0	0	11	2		13
new	0	0	0	7	5		12
stay	0	0	4	4	4		12
his	0	0	0	1	11		12
wait	6	0	6	0	0		12

Frequency Count of PM Storybook Levels							
Word	Magenta	Red	Yellow	Blue	Green	Orange	Total
cried	0	5	0	6	0		11
do	0	0	0	3	8		11
that	0	0	0	6	5		11
stayed	0	0	0	5	5		10
here's	0	0	0	6	3		9
he's	0	0	7	0	2		9
an	2	0	0	0	7		9
run	3	0	1	0	4		8
let	0	0	0	8	0		8
house	0	0	0	6	1		7
book	0	7	0	0	0		7
wakes	0	0	7	0	0		7
let's	0	0	0	7	0		7
fast	0	0	0	4	1		5
woke	0	0	1	1	3		5
playing	0	0	0	4	0		4
by	0	0	2	0	2		4
I'll	0	0	0	3	1		4
looks	0	0	0	3	1		4
please	0	0	3	0	0		3
helps	0	0	0	3	0		3
eat	0	0	0	0	3		3
helping	0	0	0	2	0		2
played	0	0	0	1	1		2
staying	0	0	1	0	1		2
that's	0	0	0	1	1		2
sleep	0	0	0	1	1		2
didn't	0	0	0	0	2		2
jump	0	0	0	1	0		1
she's	0	0	0	1	0		1
Dad's	0	0	0	1	0		1

Appendix C: Comparison to Other Word Lists

Manitoba 2002	New Zealand 1963		Ohio 1991		Canberra		New Zealand 1995	
Manitoba Words	Common Words	Different Words	Common Words	Different Words	Common Words	Different Words	Common Words	Different Words
the	√	Father	√	pretty	√	dog	√	what
said	√	school	√	ride	√	what	√	have
and	√	car	√	has	√	there	√	got
to	√	Mr.	√	him	√	that	√	there
a	√	children		find	√	his	√	that
is	√	help	√	after	√	some	√	some
I	√	ready		let	√		√	
in	√	let	√	live	√		√	
he	√	meet		may	√		√	
here	√	away	√	away	√		√	
up	√	please		there	√		√	
little		boys	√	some	√		√	
on	√			put	√		√	
went	√			over	√		√	
at	√		√	make				
look	√		√	did				
she				do	√		√	
Dad				eat	√		√	
you	√		√	what	√		√	
for	√		√	who	√		√	
are	√		√	an	√		√	
they	√		√	them	√		√	
Mom				then	√		√	
come	√			walk	√		√	
it			√	one	√		√	
my				play	√		√	
down			√	red			√	
big	√		√	again	√		√	
we	√		√	now	√		√	
can			√	could	√		√	
go	√			give	√		√	
bear				from				
was				yes	√		√	
with	√		√	saw			√	
looked				have	√			
ran			√					
am	√		√				√	
me	√				√		√	
of					√		√	
baby								
out							√	
all					√		√	
will	√				√			
not	√				√		√	
going	√						√	
I'm								
Mother	√							
like	√		√					
where	√		√		√			
comes								
get	√				√			
but					√		√	
too	√							
see			√		√		√	
back								
no			√		√		√	
this	√		√					
her			√					
came							√	
shouted	√							
day								
61	36	12	28	35	40	6	42	6

Appendix D:
Word Reading Task - Score Sheet

WORD TEST SCORE SHEET

Use any one list of words

Name: _____ TEST SCORE: _____ Date: _____

Age: _____ Date of Birth: _____ STANINE GROUP:

Record incorrect response beside word. Please circle Boy or Girl and Rural or Urban

LIST A	LIST B	LIST C
said	and	my
I	is	a
in	he	here
on	little	up
Mother	at	go
you	Dad	she
for	are	they
it	come	Mom
to	down	big
look	can	we
bear	was	with
am	ran	looked
me	of	baby
will	all	out
not	going	I'm
where	like	went
comes	get	but
back	see	too
no	this	her
day	shouted	came

COMMENT:

LIST A	LIST B	LIST C
Practice Word	Practice Word	Practice Word
the	the	the

said

and

my

I

is

a

in

he

here

on

little

up

Mother

at

go

you

Dad

she

for

are

they

it

come

Mom

to

down

big

look

can

we

bear

was

with

am

ran

looked

baby

of

me

will

all

out

not

going

I'm

where

like

went

comes

get

but

back

see

too

no

this

her

day

shouted

came

**Appendix E: -
Expected Levels of Text Reading**

APPROXIMATE READING RECOVERY™ LEVEL	READY-TO READ COLOUR WHEEL	GRADE LEVEL	EQUIVALENT AGE BAND
Level 23	Journals	Grade 3-4	8y 0m -8y 6m
Level 21-22	Dark Yellow	Grade 3	7y 3m - 8y 3m
Level 19-20	Purple	Grade 2	7y 0m -7y 6m
Level 17-18	Blue	Grade 2	6y 6m - 7y 3m
Level 15-16	Orange	Grade 1	6y 0m -6y 9m
Level 12-14	Green	Grade1	5y 9m -6y 3m
Level 9-11	Dark Blue	Grade 1	5y 6m - 6y 0m
Level 6-8	Yellow	Grade1	5y 3m - 5y 9m
Level 3-5	Red	K-1	5y 3m - 5y 6m
Level 1-2	Magenta	KGN	5y 0m- 5y 6m

Appendix F:

Protocol for Administering Running Records of Text Reading

1. Use the books from the attached list. Notice the list does not include the PM Plus series.
2. The administration may begin with a familiar book to put the student at ease and establish rapport.
3. The tester should always read the title to the student before taking the running record.
4. The student may want to look through the book first. This would be acceptable if the student initiated.
5. Student responses should be accepted as given. Responses should not be added to or clarified. You may also tell the student the word if necessary. Remember to use two responses - "You try it" or "Try that again."
6. Testing can be stopped after one level below 90%, and/or if the student is showing signs of frustration with the loss of meaning and little or no self-correction is occurring (this will probably be below 90% accuracy). You may record on this running record, "too hard."
7. You will need to calculate the accuracy rate but you do not have to analyze the Running Records. We only need an indication of the highest level of text the student can read at 90% or better.
8. Attached the running record forms to the other task sheets.

Appendix G:
List of Titles For Assessment

PM Starters Level 1 and 2 - Magenta

Me	The go-karts
Mom	In the shopping cart
Dad	Climbing
A house	The shopping mall
Big things	Playing
Little things	Look at me
Dressing up	Pets
Time for Dinner	The way I go to school
At the zoo	We go out
Moms and Dads	The skier
Out in the Weather	Ball games
Cat and Mouse	We can run
Where are the babies?	Ben's red car
Packing my bag	My little dog
The rock pools	The pencil
Stop!	Looking down
Sally's new shoes	Fishing

PM Library Red Level Set A and B

The Photo Book	The little Snowman
Hedgehog is hungry	A birthday cake for Ben
Wake up, Dad	The baby owls
Tiger, Tiger	The bumper cars
The lazy pig	The flower girl
The merry-go-round	Hide and seek
Baby lamb's first drink	A home for little Teddy
Sally and the daisies	Where is Hannah?
The big kick	Hot dogs
Kitty and the birds	Ben's teddy bear
Ben's treasure hunt	Lizard loses his tail
Father Bear goes fishing	Tom is brave

PM Library Yellow Level Pack A & B

Where are the sunhats?	Blackberries
Brave Father Mouse	Mumps
The hungry kitten	Sally's beans
Lucky goes to dog school	Ben's Dad

The new baby
 Hermit Crab
 Little Bulldozer
 Friend for little white rabbit
 Baby Hippo
 Soccer at the park
 Tiny and the big wave

Baby Bear goes fishing
 Seagull is clever
 Sally's red bucket
 Fire! Fire!
 Choosing a puppy
 A lucky day for little dinosaur
 Snowy gets a wash

PM Library Blue level Pack A & B

Magpie's baking day
 Sally's friends
 Baby Bear's present
 Jane's car
 The house in the tree
 Mushrooms for dinner
 Cows in the garden
 Locked out
 Little Bulldozer helps again
 Lost at the fun park
 Tim's favorite toy
 Teasing Dad

The lion and the rabbit
 Honey for Baby Bear
 The best cake
 Tabby in the tree
 The Christmas tree
 Come on, Tim
 Late for soccer
 The lion and the mouse
 The duck with the broken wing
 Birthday balloons
 Chug the tractor
 Tiger runs away

PM Library Green Level Pack A & B

The Naughty Ann
 The clever penguins
 House-hunting
 Mrs. Spider's beautiful web
 Ten little garden snails
 The little red bus
 The flood
 The waving sheep
 The flying fish
 Snow on the hill
 Father Bear's surprise
 Joey

Brave Triceratops
 Pete Little
 Candle-light
 Ben's tooth
 The fox who foxed
 The island picnic
 Pepper's adventure
 The cross-country race
 The rescue
 The babysitter
 After the flood
 Try again, Hananh

PM Library Orange Fiction Pack
 Traditional Tales and Plays

The dinosaur chase
 Jack and Chug
 Toby and the Big Tree
 Sarah and the Barking Dog

The Biggest Fish
 Toby and BJ
 The toy farm
 Pterosaur's Long Flight

Just One Guinea Pig
 Toby and the Big Red Van
 Rebecca and the Concert
 The Careful Crocodile
 The Busy Beavers
 Chicken-Licken
 The Three Little Pigs
 The Tale of the Turnip

Jessica in the Dark
 Mitch to the Rescue
 Lost in the Forest
 Skates for Luke
 Two Little Goldfish
 The Gingerbread Man
 The Little Red Hen
 The Three Billy Goats Gruff

PM Library Turquoise Level Fiction Pack
 Traditional Tales

When the Volcano Erupted
 The Cabin in the Hills
 Nelson the Baby Elephant
 Little Dinosaur's Narrow Escape
 License Plates
 Bird's Eye view
 Grandpa's Mask
 The Nesting Place
 Riding to Craggy Rock
 Goldilocks and the Three Bears
 Stone Soup
 The Ugly Duckling

Monkey Tricks
 Jonathan Buys a Present
 Toby and the Accident
 Rescuing Nelson
 The Seatbelt Song
 The Hailstorm
 Ant City
 Jordan's Lucky Day
 The Race to Green End
 Little Red Riding Hood
 The Elves and the Shoemaker
 The Brave Little Tailor

PM Library Purple Level
 Traditional Tales and Plays

The Skating Trail
 A Dog Called Bear
 Moppet on the Run
 The Green Dragons
 A New School for Megan
 A Bike for Brad
 Gorgo Meets her Match
 Two Red Tugs
 Riding High
 Sly Fox and the Little Red Hen
 The Hare and the Tortoise
 Town Mouse and Country Mouse

Nelson Gets A Fright
 Zala Runs for her Life
 Caps for Sale
 A Troop of Little Dinosaurs
 Muffin is Trapped
 The Surf Carnival
 Jordan's Catch
 Toby at Stony Bay
 The Marble Patch
 The Boy Who Cried Wolf
 The Animal Band
 Puss-in-Boots

Appendix H: Overall Stanine Scores

Running Records of Text Reading

Stanine Group	1	2	3	4	5	6	7	8	9
Raw Score	0	1-2	3-5	6-9	10-13	14-19	20		

Concepts About Print

Stanine Group	1	2	3	4	5	6	7	8	9
Raw score	0-12	13-14	15	16-17	18	19	20	21-22	23-24

Letter Identification

Stanine group	1	2	3	4	5	6	7	8	9
Raw Score	0-47	48-50	51	52	53	54			

Clay Word Reading Task

Stanine Group	1	2	3	4	5	6	7	8	9
Raw Score	0-2	3-4	5-8	9-11	12-13	14	15		

Writing Vocabulary

Stanine Group	1	2	3	4	5	6	7	8	9
Raw Score	0-7	8-15	16-22	23-29	30-39	40-46	47-55	56-65	66-81

Hearing and Recording Sounds in Words

Stanine Group	1	2	3	4	5	6	7	8	9
Raw Score	0-15	16-26	27-31	32-33	34-35	36	37		

Manitoba Word Reading Task

Stanine Group	1	2	3	4	5	6	7	8	9
Raw Score	0-3	4-8	9-12	13-16	17-18	19	20		