

2009

Use of Kindergarten Screening Assessments for the Identification of At-risk Readers

Kathryn S. Gipe

Philadelphia College of Osteopathic Medicine, katiegipe@gmail.com

Follow this and additional works at: http://digitalcommons.pcom.edu/psychology_dissertations



Part of the [School Psychology Commons](#)

Recommended Citation

Gipe, Kathryn S., "Use of Kindergarten Screening Assessments for the Identification of At-risk Readers" (2009). *PCOM Psychology Dissertations*. Paper 51.

This Dissertation is brought to you for free and open access by the Student Dissertations, Theses and Papers at DigitalCommons@PCOM. It has been accepted for inclusion in PCOM Psychology Dissertations by an authorized administrator of DigitalCommons@PCOM. For more information, please contact library@pcom.edu.

Philadelphia College of Osteopathic Medicine

Department of Psychology

THE USE OF KINDERGARTEN SCREENING ASSESSMENTS FOR THE
IDENTIFICATION OF AT-RISK READERS

By Kathryn S. Gipe

Submitted in Partial Fulfillment of the Requirements of the Degree of

Doctor of Psychology

December 2009

PHILADELPHIA COLLEGE OF OSTEOPATHIC MEDICINE
DEPARTMENT OF PSYCHOLOGY

Dissertation Approval

This is to certify that the thesis presented to us by Kathryn Gipe
on the 18th day of June, 2009, in partial fulfillment of the
requirements for the degree of Doctor of Psychology, has been examined and is
acceptable in both scholarship and literary quality.

Committee Members' Signatures:

George McCloskey, Ph.D., Chairperson

Diane L. Smallwood, Psy.D., NCSP

Barbara B. Williams, Ph.D.

Robert A. DiTomasso, Ph.D., ABPP, Chair, Department of Psychology

Acknowledgments

To Dr. George McCloskey for his wisdom, perspective, and encouragement throughout the dissertation process. He has shared valuable knowledge about program evaluation, assessment and the fundamentals of reading. His guidance has been instrumental throughout each aspect of this project. I am grateful for his patient approach and endless hours of assistance. The positive nature with which he interacts with students is commendable and appreciated.

To Dr. Diane Smallwood for her attention to detail and her flexibility throughout the process. I admire the high standards that she sets for her students and have benefited from her encouragement to produce high quality work.

To Dr. Barbara Bole Williams who has been an inspiration from the early days of my School Psychology career. Her dedication to the field of School Psychology and the students with whom she works is extraordinary.

To Gail Kulick of the East Stroudsburg Public Schools for her dedication to the field of education. Without her, this study would not have been possible. I appreciate her generosity of time and the openness to sharing the amazing work of the students and faculty at Resica Elementary School. Her assistance with data collection was invaluable. I respect her progressive approach to program analysis as well as her collegial approach to sharing knowledge.

To my family and friends who have been by my side throughout this endeavor. You have kept me motivated, encouraged me to have perspective, and enabled me to continue with this project when times were difficult.

To my parents, Robert and Susan Gipe, who have been sources of great strength. To my dad who has inspired me to set my standards high as a person, as a student, and as a professional. He has led by example and his amazing determination, and unrelenting work ethic have been a phenomenal force in shaping my life. To my mom; her unconditional love has supported me through challenging times. Her faith in me has enabled me to pursue my dreams, believe in myself, and to accept challenges with grace. I have learned from her that much of life's joy comes from putting others before oneself. Together, my parents have instilled in me the value of education, the importance of staying true to oneself, and the belief that the harder one works the sweeter the success. I love you both and owe much of my accomplishments to you.

To A.J. Talone, whose patience, understanding, and sense of humor have been a gift. Through many stages of schooling he has graciously shared me with my books, research, and evaluations. I have deep appreciation for his endless love and encouragement.

To my friends, and fellow School Psychologists, Mary Heim Elberson and Julie Boyle, who have provided unrelenting support, have cheered me on throughout this endeavor, and cajoled me into laughter along the way. Their friendship has been a bright spot throughout the quest for my

advanced degree. I admire their work as School Psychologists and I am proud to have them as colleagues and friends.

To my colleague, Dr. Barry Barbarasch, my gratitude for his interest in my doctoral pursuit. His clear and realistic vision of the dissertation process has helped me to stay focused and balanced. I admire his commitment to ethical practice, and I am grateful to have him as a colleague. To my teammates Gina DiCiurcio, Ella Kelly, and Joan Nagy, I appreciate the interest they have shown in my academic undertakings. Their friendships and encouragement have kept me motivated throughout this journey.

For the children with whom I work, and for those whose paths I will cross, I hope that this endeavor will enable me to better serve you now and in the future. The gift of education is most precious, and I am honored to be part of the educational process of our youth.

Table of Contents

Chapter	Page
1. Introduction	1
Statement of the Problem	3
Purpose of the Study	3
Related Research	4
Kindergarten Reading Skills	4
Use of Screening Instruments	10
Research Questions	21
2. Methods	22
Participants	22
Measures	22
School-made Kindergarten Screening Instrument	22
DIBELS	23
Measures of Academic Performance	23
Procedures	26
Statistical Analyses	26
3. Results	31
Research Question 1	31
Research Question 2	32
Research Question 3	36
Research Question 4	41

Research Question 4a	41
Research Question 4b	43
Research Question 4c	48
4. Discussion	57
Summary of Results	57
Research Question 1	57
Research Question 2	59
Research Question 3	61
Research Question 4	71
Research Question 4a	71
Research Question 4b	72
Research Question 4c	75
Contributions to the Field	93
Limitations of the Study	93
Future Directions	94
References	94
Appendix A	102

Tables

Tables	Page
1 Relationship between Kindergarten Screening Measure and Fall DIBELS Assessment	31
2 Relationship between Kindergarten Screening Measure and Kindergarten DIBELS with Winter Grade 1 Oral Reading Fluency	33
3 Relationship between Kindergarten Screening Measure and Kindergarten DIBELS with Grade 2 Winter Oral Reading Fluency	35
4 Percentage of Students At-Risk based on Winter Oral Reading Fluency	36
5 Percentage of Students At-Risk based on Grade 3 MAP Scores	37
6 Kindergarten DIBELS Summary Score Compared with Winter Grades 2 and 3 ORF and Grade 3 Fall and Spring MAP Scores	39
7 Kindergarten Screening Measure and Kindergarten DIBELS with Grade 1 Winter ORF with SES	41
8 Kindergarten Screening Measure and Kindergarten DIBELS with Grade 1 Winter ORF with SES	44
9 Kindergarten Screening Measure and Kindergarten DIBELS Compared with Grade 2 Winter ORF by SES Group	46

10	Percentage of At-Risk Students, by SES Group based on DIBELS Measures	48
11	Percentage of At-Risk Students, by SES Group based on Grade 3 MAP Scores	50
12	Kindergarten DIBELS Summary Score Compared with Winter Grade 2 and Grade 3 DIBELS ORF by SES Group	52
13	Kindergarten DIBELS Summary Scores and MAP Scores by SES	53
14	Grade 3 MAP Comprehension Domain Data for the Low SES Group of the 2003-2004 Cohort	89

Figures

Figure		Page
1	Indices used in Statistical Analyses of Data	27
2	Test Score Variables Used in Analyses for Each Kindergarten Cohort	30

Abstract

Early identification of students at-risk for reading problems has become a national priority. At the present time, the most commonly used kindergarten screening methods are self-made by local districts and are not considered effective methods of early detection for at-risk readers. This retrospective study involved third, fourth and fifth grade students enrolled in a suburban elementary school during the 2007-2008 school year. The first research question examined the relationship between at-risk status determined with the Kindergarten Screening measure prior to entry into kindergarten and at-risk status determined by Dynamic Indicators of Basic Early Literacy Skills measures administered in the fall of kindergarten. The comparison between the category classifications derived from the Kindergarten Screening Summary Scores and the category classifications derived from the Dynamic Indicators of Basic Early Literacy Skills Summary Scores produced the best combination of sensitivity and specificity values. The second research question was designed to examine the predictive relationship between the Kindergarten Screening measure and the Kindergarten Dynamic Indicators of Basic Early Literacy Skills measures with Winter Oral Reading Fluency in Grades 1 and 2. The Dynamic Indicators of Basic Early Literacy Skills Summary score demonstrated greater levels of sensitivity and greater improvement over chance in predicting Grade 1 Winter Oral Reading Fluency than did the Kindergarten Screening measure for two of the three cohorts and greater sensitivity than Letter Naming Fluency or Initial Sound Fluency for all cohorts. The third research question investigated the effectiveness of the reading instruction based on the results of the Kindergarten Summary Scores, Oral Reading Fluency and Measures of Academic Performance scores. Improvements in at-risk status were noted as for Oral Reading Fluency as

cohorts progressed through grades 1 and 2; this was followed by a slight decrease in grade 3, and improvements in subsequent cohorts were consistently reported. Improvement between the fall and spring Measures of Academic Performance scores was also reported for each cohort as well as with successive cohorts, suggesting positive effects of general education instruction and remedial efforts. The effect of Socioeconomic Status was investigated through re-examination of each research question, with the inclusion of Disadvantaged and Not Disadvantaged status. Results indicated that the inclusion of Socioeconomic Status is important in the prediction of at-risk status, and in the investigation of effectiveness of reading instruction.

Chapter 1

Introduction

The necessity for early identification of children with reading difficulties has become a national priority. The U. S. Department of Education, National Center for Education Statistics, sponsored a national longitudinal study of kindergartners, schools, teachers and families, titled the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K), (Denton & Germino-Hausken, 2000). This study began to draw attention to concerns surrounding kindergarten aged children and learning. ECLS-K examined factors associated with student learning, addressed the skill levels with which kindergarten aged children enter school, and investigated the implications of particular skill levels. ECLS-K was broad in scope and addressed Reading, Mathematics and General Knowledge. Congress raised concerns regarding the academic performances of young children. In 1997 Congress requested that the National Institute of Child Health and Human Development, in cooperation with the Secretary of Education, investigate the prevention of reading difficulties (National Reading Panel, 2000). The National Reading Panel (NRP) was established, and it set forth to summarize research critical to skills, environments, and developmental components that are essential to developing reading skills. The NRP published a report highlighting major findings and recommendations for the prevention of reading difficulties through the administration of sound teaching techniques which target key components of the reading process.

Subsequently, The No Child Left Behind Act (NCLB) of 2001 addressed illiteracy, outlining the early identification of children who are at-risk for reading difficulties as an essential feature in the prevention of illiteracy (NCLB, 2001). In fact, early intervention has been found to remedy the majority of reading problems (Snow, Burns & Griffin, 1998).

Furthermore, the earlier that reading interventions are implemented for at-risk readers, the more likely it is that the remediation of reading difficulties will occur (Bishop, 2003; Snow et al., 1998). Therefore it is essential to identify at-risk readers as early as possible. In terms of kindergarten screening, it is preferable that screenings take place in mid-Kindergarten to minimize over-identification of children who have not been exposed to early literacy skills prior to Kindergarten (O'Connor & Jenkins, 1999).

Statement of the Problem

Historically, Kindergarten screening procedures have involved primarily district made assessments, which demonstrate modest predictive validity in terms of reading achievement (Rafoth, 1997). Consequently, such screenings have done little to identify children entering Kindergarten with higher levels of risk for later reading problems. Current research suggests that specific pre-reading skills correlate with later reading achievement. Such skills include letter naming, rapid automatic naming, phonological awareness, phonemic awareness, phoneme segmentation and concepts of print (NRP, 2000). However, the ideal combination of measures has not been determined that will best predict children who are likely to be most at risk of reading problems (O'Conner & Jenkins, 1999, Snow et al., 1998, Torgesen, 1998).

Purpose of the Study

The increased interest in skill levels at kindergarten entry, the national call to focus on prevention, and findings from the NRP have highlighted the need for further investigation of how best to identify children at risk of developing reading problems and how best to prevent reading difficulties at a young age. This paper will investigate the predicative validity of standardized and self-made kindergarten screening instruments as related to later reading achievement in an effort to contribute to the literature base in this important area of study.

Related Research

Kindergarten Reading Skills

Reading ability, and conversely disability, is a long debated topic ; this includes varied explanations of how the ability is acquired (Levine, 2001). Likewise, variation exists in the literature regarding the necessary kindergarten age reading skills, as well as the screening instruments to assess such areas accurately(Kurdek & Sinclair, 2001). General cognitive skills and knowledge are often regarded as main components in academic success (Denton & Germino-Hausken, 2000). More specifically,however, letter naming and phonemic awareness have consistently been identified as the strongest predictors of how well children will learn to read (NRP, 2000; O'Conner & Jenkins, 1999; Snow et al., 1998). Additionally, concepts of print has been identified as a moderate predictor of later reading ability (Bishop, 2003; Snow et al., 1998). In order to select appropriate assessment tools for the identification of at-risk readers, it is essential to understand the early components and skills associated with the reading process.

Letter Naming and Rapid Automatic Naming

Letter identification has traditionally been included in the majority of readiness screening instruments, and has proved to be as strong a predictor for later reading as an entire readiness test (Snow et al., 1998). The NRP concluded that letter knowledge was one of the best predictors of how well students will learn to read (2000). A letter naming task typically measures how many letters a student can identify when he or she is presented with these letters in a random order (Bishop, 2003). This skill leads to the alphabet principle, which entails the connection that words are composed of letters that represent sounds (Kurdek & Sinclair, 2001; Snow et al.), a central feature in the acquisition of reading skills. Theoretically considered the first step in the

reading process, emergent readers must focus on letter patterns, the linking of letters, and the linking of letters to form words in order to begin reading.

Significant findings have been reported for letter naming as a predictor of reading achievement (Lennon & Slesinski, 1999; Share, Jorm, Maclean & Matthews, 1984). The use of letter naming as a predictor of reading achievement was found to be more useful than a traditional discrepancy model (Lennon & Slesinski). Lennon and Slesinski examined early intervention in reading and reported significant findings regarding the use of letter-naming and other measures of reading as predictors of later reading performance. Letter naming was used as a basis to group the 330 participants into three sections, low scoring participants, mid-scoring participants, and high scoring participants. The high scoring participants did not receive tutoring, but the other two groups did receive tutoring. Progress monitoring occurred at the ten and twenty week marks, and letter naming was found to correlate positively ($p < .001$) with six measures of reading performance including: letter sounds, phoneme segmentation, decoding, sight words, and concepts of print. A similar study conducted in Australia investigated factors associated with reading achievement, and letter names were used as one factor in the prediction of reading performance (Share et al., 1984). This longitudinal study of 543 Kindergarten children individually assessed participants by asking them the names of nine letters. Results of letter naming were compared with the Neale Analysis of Reading Ability (Neale, 1966) to investigate a predictive relationship. The letter naming task was found to predict performance consistently on the reading tasks, and demonstrated stronger predictive validity than all other measures. Simple correlations between letter naming and reading performance were found to be .68 and .58 at the end of Kindergarten and First Grade, respectively. Based on several studies, kindergarten measures of letter identification have been strong predictors of future reading

performance (Lennon & Slesinski; Share et al.), and have accounted for up to one-third of the variance in reading during grades 1-3 (Snow et al., 1998).

The link between letter identification and later reading performance has been established, and subsequently the speed of identification has been given attention (Blachman, 1984). The relationship between the speed of letter naming and later reading performance has been well investigated (Blachman; Denckla & Rudel, 1976). One of the first studies addressing letter naming and rapid automatic naming compared the performance of dyslexic students, non-dyslexic students and a control group (Denckla & Rudel). The study found that the accuracy of color, object, letter and number naming did not differ between or among groups. However, the response time for students with dyslexia was significantly longer than response time for the other two groups. Likewise, a similar study of second graders examined the response rate of good readers and of less skilled readers. Results suggested significant differences in the speed of naming, with more skilled readers responding more quickly (Blachman, 1984). The results of these early studies, and subsequent studies replicating the findings, lend credence to the argument that speedier identification of letters is correlated to reading performance (Blachman; Felton, 1992).

Phonological awareness

The Florida Center for Reading Research (FCRR) defines phonological awareness as a broad term that represents sensitivity to, and awareness of, the structure of words (2006a). Phonological awareness in young children has been established as a strong predictor of later reading success (Adams, 1990; NRP, 2000; Snow et al., 1998). Students with weak phonological awareness are likely to fall behind their peers and have difficulty learning to read (Lennon & Slesinski, 1999; Snow et al.). Phonological awareness encompasses phonemic

awareness, which is characterized by the abilities to recognize, think about, and manipulate phonemes, the smallest sound units in language (FCRR 2006b; Snow et al.; Stahl & Murray, 1994). Phonemic awareness typically involves isolation and manipulation of phonemes, as well as syllables (NRP). Therefore, identification of students with weak phonological awareness, and specifically phonemic awareness, becomes essential for early intervention.

Phonemic awareness and phoneme segmentation

A key component of phonemic awareness is phoneme segmentation, in which the individual sounds of a word can be separated into units (FCRR, 2006b). Phoneme segmentation is key to understanding the alphabetic principle; because phonemes are represented by letters, it is essential to recognize phonemes in order to master the concept of the alphabetic principle (Snow et al., 1998). The performance on phoneme segmentation tasks has been established as a predictor of reading success (Blachman, 1984; Share et al., 1984; Stahl & Murray). In addition to letter naming, the study by Share et al. investigated phonemic awareness as a predictor of reading ability. Specifically, the study examined phoneme segmentation. The method of the study required the student to isolate the beginning sounds and then to say the remainder of the word. Additionally, the student was required to break words into beginning, middle and ending sounds. Phonemic awareness measures were taken during the first term of kindergarten ;subsequently, reading measures were taken at the end of the kindergarten year and at the end of first grade. Reading performance was measured by the Neale Analysis of Reading Ability (Neale, 1966). Findings suggested that phoneme segmentation, as well as letter identification, consistently predicted reading outcomes more accurately than any oral language abilities or motor skills. Phoneme segmentation demonstrated a significant correlate to reading performance at the end of kindergarten ($r = .66$), and at the end of first grade ($r = .68$). Results clearly support

the use of phonemic awareness as a screening tool for later reading performance. However, phoneme segmentation can be taught, and consequently may improve future reading performance (Fuchs & Fuchs, 1999). Therefore, early identification of students who lack the skill of phonemic segmentation is vital to intervention and to the later success of the reader. In order to deliver intervention, the students must be identified as having difficulty with phoneme segmentation.

The use of phonological and phonemic awareness as a predictor of reading performance has been substantiated (Blachman, 1984; Share et al., 1984; Snow et al., 1998; Stahl & Murray, 1994). However, some critics suggest that at the kindergarten level it is a better predictor of superior reading ability, as opposed to a predictor of those who will struggle with reading (Felton, 1992; Snow et al.). Snow et al. investigated the predictive correlation of phonological awareness and reading. Results of 27 research samples were studied, and findings suggested that phonological awareness demonstrated the same predictive validity as memory for sentences and stories, confrontation naming and general language measures ($r = .46$). These findings led Snow et al. to conclude that at the Kindergarten level, phonological awareness is not necessarily a clear predictor of later reading achievement, and that early measures of phonological awareness may not accurately identify those who will have later reading difficulties from those who will be successful.

Conversely, one unique study selected those students who were predicted to be above-average or superior in reading skills, and examined the relationship between phonological awareness and reading performance (Felton, 1992). Based on information provided by their kindergarten teachers, students identified as potentially strong readers were eliminated from the study to examine the predictive validity for poorer readers. With the remaining participants,

Felton examined several factors associated with reading, including phonological awareness. In this longitudinal study, kindergarten participants were administered measures of phonological awareness and subsequently reading achievement was measured in the third grade. Results suggested that phonological awareness was one of three variables that contributed to the prediction of reading success in third grade. These results are encouraging and suggest that measures of phonological awareness can, in fact, be successful in the identification of students who are at-risk of developing reading difficulties.

Concepts of print

Concepts of print is considered the basic awareness of how print can be used, as opposed to knowledge regarding specific letters (Snow et al., 1998) and includes the awareness of the semantic and the visual arrangement of text (Gunn, Simmons & Kameenui, 1995). Throughout the preschool and kindergarten years children learn about the conventions, purpose and function of the printed word. This awareness is often gained through experience with storybook reading and daily living routines. Concepts of print awareness has been proven to have a moderate correlation with reading performance in the early grades (Snow et al.), and lack of the skill has been linked to later reading difficulty (Gunn et al.).

Concepts of print can be further distinguished as conventions of print, and purpose and function of print (Gunn et al., 1995). Conventions of print is associated with the physical structure of print and later, with the notion that text progresses from top to bottom and from left to right. The purpose and function of print encompass the knowledge that print differs from speech, although, similar to speech, print carries a message. Consequently, the connection is made that the printed word carries the message, and that the picture does not. Finally, the

development of concepts of print can also be detected in the acquisition of writing skills, as children begin to create messages through scribbling.

By the time children who will become poor readers have entered school, Hildebrand and Bader (1992) found that they have had significantly less exposure to print than children who develop into better readers (as cited in Gunn et al.). Parents of preschoolers were interviewed regarding their children's exposure to reading. Topics such as frequency of adult-child reading, adult reading and children's independent use of books were questioned. This longitudinal study conducted by Scarborough et al. (1991) followed the participants to second grade, at which time it was determined that children who became poor readers had substantially less exposure to print than those who became strong readers (as cited in Gunn et al.).

Use of Screening Instruments

The knowledge of those skills that are linked to later reading performance is important in the identification of students who are at-risk of developing reading problems. When such skills have been identified, the next step becomes the recognition of screening practices and tools that assist with the identification of potentially at-risk students. Screening practices have been commonplace in the United States for at least the last 50 years. This practice began primarily as a method for identification of first grade children who may have difficulty learning to read (Gredler, 1997). However, the development and implementation of accurate measures to identify at-risk learners has been a challenge (Bishop, 2003). Standardized measures that target the key components of the reading process have been lacking (Bishop), and accurate timing of identification has been problematic (O'Connor & Jenkins, 1999).

A survey of state assessment practices revealed that 16 of 48 states mandate some sort of screening procedure for Kindergarten students (as cited in Rafoth, 1997). The goals of such

screening procedures may vary from identification of students whose entry to school should be delayed, to forming groups for instructional purposes. However, the importance of early identification of at-risk students has been established (NCLB, 2001) and therefore, the goal should be to identify at-risk learners and assist teachers in the planning of appropriate curricular modifications (Rafoth). Yet the most efficient and effective method of early identification remains controversial.

The selection of specific tools for screening should be made with the objective of identifying at-risk readers. Tools to assess readiness are not appropriate for use if the goal is to identify at-risk students (Rafoth, 1997). Such readiness tools are typically criterion-referenced and describe specific attributes of the child that are not directly related to success in reading. Ideally, screening tests should be norm-referenced and assess the potential of a student in terms of the probability of reading skill acquisition.

Countless options exist for screening purposes, and are published or are locally made by school districts. Costenbader, Rohrer & Difonzo, (2000) report that of the 358 New York State schools surveyed, 30% (112 schools) used locally developed measures; 26% (97 schools) used the Diagnostic Inventory of Early Development—Revised (DIAL-R) (Mardell- Czudnowski & Goldenberg, 1990); 16% (60 schools) administered the Brigance K &1 Screen (Brigance, 1992), and 13% (50 schools) utilized the Gesell Test of Early Learning (Ilg & Ames, 1972). Although these measures have some components that address literacy, they are more broad based, and screen for more than literacy skills.

Conventional kindergarten screening batteries, which typically involve measures of fine and gross motor skills, visual-perceptual skills, knowledge of the alphabet, counting skills, vocabulary and oral expression, are reported to be only moderately accurate in identifying

students who are at-risk of developing reading difficulties (Rafoth, 1997). Most predictor batteries overlook phonological awareness (Share et al., 1984), which may explain the moderate success of identification. Likewise, discrepancy models that focus on differences between achievement and intelligence have not been particularly helpful in the early identification of at-risk readers (Felton, 1992). Despite the vast amount of knowledge surrounding the acquisition of reading readiness skills, an ideal combination of measures has not yet been perfected in the early identification of at-risk readers (O'Conner & Jenkins, 1999; Snow et al., 1998).

A summary provided by the FCRR delineates several tools for the assessment of emerging literacy skills. Several screening instruments from the FCRR's recommended list (2006a) will be reviewed here. Of those notated by the FCRR, the Phonological Awareness and Literacy Screening-Pre Kindergarten (PALS-PreK) (Invernizzi, Sullivan, & Meier, 2002), Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good & Kaminski, 2002), Developing Skills Checklist (DSC) (CTB Macmillan McGraw Hill, 1990) the Lindamood Auditory Conceptualization Test (LAC) (Lindamood & Lindamood, 1979) are considered screeners, rather than diagnostic tools. These screeners assess at least one of the major components associated with early reading skills. To target the kindergarten population most accurately, the PALS-Kindergarten (PALS-K) (Invernizzi, Swank, Juel & Meier, 2003) will be reviewed as opposed to the PALS-PreK. Additionally, the DIAL-III (Mardell-Czudnowski & Goldenberg, 1990), Brigance K & 1 (Brigance, 1992), Gesell School Readiness Test (Ilg & Ames, 1972) and a general overview of locally made screenings will be reviewed, based on their frequent use (Costenbader, et al., 2000).

Phonological Awareness and Literacy Screening-Kindergarten (PALS-K)

The PALS-K (Invernizzi et al., 2003) was established in response to increased interest from educators, legislators, and policymakers to identify children at-risk for developing reading difficulties early in their educational careers. Through the year 2003, more than 430,000 kindergarten students were screened using the PALS-K (Invernizzi, 2004), which was designed with the knowledge that the two best predictors of later reading achievement include phonological awareness and letter recognition. Therefore the instrument includes subtests to address these two areas as well as tasks to screen Grapheme-Phoneme Correspondence, the relationship between letters and sounds, Spelling and Concept of Word.

A longitudinal analysis was completed through comparison of scores recorded on the fall administration of the PALS-K with results of a Standards of Learning test (SOL), a measure based on state standards (Invernizzi, 2004). At the end of the third grade, results of discriminate analysis accurately predicted pass/fail results on the SOL test in reading for 80% of the children screened. The longitudinal study lends support to the validity of the PALS-K for predicting reading achievement.

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

DIBELS was created with the intention of the early identification of children who were not making progress toward the mastery of the necessary early reading skills and the ongoing monitoring of skill acquisition (Kaminski & Good, 1996). The authors of the DIBELS designed the measure to correlate with early literacy scores, specifically phonological awareness, knowledge of letter names and language skills. Scores on the DIBELS can be interpreted, based on published guidelines pertaining to local norms, allowing for easy identification of at-risk students.

Depending on the grade level, DIBELS is composed of different subtests; however, all subtests are administered individually. At the kindergarten level, a typical DIBELS screening includes Initial Sound Fluency (ISF), Letter Naming Fluency (LNF), Phoneme Segmentation Fluency (PSF) and Nonsense Word Fluency (NWF) (Good & Kaminski, 2002). The LNF test presents the child with upper and lower case letters in a randomized fashion. The student is permitted 1 minute to name as many letters as possible. The predictive validity of kindergarten LNF scores with subsequent first grade scores on the Woodcock-Johnson Psycho Educational Battery—Revised Reading Cluster (Woodcock & Johnson, 1989) was reported to be .65, and correlations with first grade curriculum based measures oral reading fluency was .71 (as cited in Good & Kaminski, 2002).

ISF is targeted to assess the student's skill level in recognizing and producing initial sounds when words are orally presented (Good & Kaminski, 2002). Students are shown four pictures and the items are named by the examiner. The student is then asked to identify the picture that starts with the same sound that is produced by the examiner. Additionally, the student is asked to produce the initial sound orally in a word that the examiner supplies orally. Specific information regarding predictive validity of ISF with later reading performance has not been published. ISF takes approximately three minutes to administer.

PSF is recommended for kindergarten students later in the year and requires the student to segment three and four phoneme words into their individual phonemes (Good & Kaminski, 2002). Administration time for PSF is approximately 2 minutes. Predictive validity for PSF scores in the spring of kindergarten in correlation with spring scores on the Woodcock-Johnson Psycho Educational Battery—Revised (Woodcock & Johnson, 1989) total Reading Cluster is .68 (as cited in Good & Kaminski).

Finally, the NWF assesses the student's performance in regard to letter sound correspondence, and the ability to blend letters into words. The student is given a paper with nonsense words and is asked to supply, verbally, either an individual letter sound or the entire word. The subject is permitted 1 minute to complete as many nonsense words as possible. Predictive validity of late kindergarten scores on the NWF to WJ-III-R total Reading Cluster is reported as .60 (as cited in Good & Kaminski).

The upper grades utilize a different set of measures. Beginning in first grade, an Oral Reading Fluency (ORF) measure is individually administered and is associated with accuracy and fluency of text (Good & Kaminski, 2002). The ORF measure was designed not only to identify students who may need more support, but also to monitor progress toward reading goals. The measure is administered by an individual student reading a passage aloud for one minute. Word omissions, word substitutions and hesitations of more than three seconds are scored as errors, but self-corrections within three seconds are scored correctly. The overall number of correctly read words becomes the score. Test-retest reliability ranged from .92 to .97, and criterion-related validity ranged from .52-.91 (as cited in Good & Kaminski, 2002). The link between statewide assessments and CBM measures of reading, specifically ORF, has been established by numerous states and across many studies (Shapiro, Keller, Lutz, Santoro & Hintze, 2006).

The reliability and validity of the measure has been investigated and has yielded encouraging results. Based on the Spearman-Brown formula, DIBELS were found to be a stable measure, with reliability estimates ranging from .97-.99 (Kaminski & Good, 1996). Furthermore, concurrent, criterion related validity was investigated through correlations on the DIBELS measures with several criterion based tools. Correlations for the Kindergarten

measures were positive, (.43-.90, $p < 0.1$), suggesting that DIBELS does measure what it purports to measure (Good & Kaminski 2002; Kaminski & Good, 1996; VanDerHyden, Witt, Naquin & Noell, 2001). Finally, a discriminate function was conducted to investigate if the probes accurately predicated retention. Results suggested that scores on the probe measure accurately predicted retention in 71.4% of the cases, and accurately predicted students that would not be retained in 94.4% of the cases (VanDerHyden et al.).

Numerical DIBELS scores are recorded based on a student's performance for each of the subtests administered. These numerical scores can be easily translated into percentiles, based on the charts provided by the scoring manual. Percentiles are provided for each subtest based on grade and the time of year that the probe was given. Percentiles are reported for the "beginning," "middle" and "end," of each grade and for each subtest and are then associated with "level of risk." Levels of Risk are reported as High Risk (HR), Moderate Risk (MR), Low Risk (LR) and Above Average (AA). HR is defined as seriously below grade level and in need of substantial intervention; MR is defined as moderately below grade level and in need of additional intervention; LR is considered at grade level and AA is described as at or above the 60th percentile.

Developing Skills Checklist (DSC)

The purpose of the DSC is to measure skills and behaviors that are typically developed between the ages of pre-kindergarten and the end of kindergarten (Clark & Earnhart, 1995). This individually administered test is reported to be closely aligned with Kindergarten classroom activities. The DSC addresses prereading skills, as well as mathematics skills, fine and gross motor development, printing, and writing, and social and emotional development. Adequate reliability and internal consistency are reported for most scales (.81-.95) except for the Visual

scale (.69). Test-retest reliability was not reported, and data regarding predictive validity are not included in the technical manual.

Lindamood Auditory Conceptualization Test (LAC)

The LAC is an individually administered assessment tool (Lindamood & Lindamood, 1979). The purpose of the LAC is to identify individual speech sounds, determine the number of speech sounds within a spoken pattern and order the sequence of sounds (Bountress & Cox, 2007). Administration time is approximately 10 minutes (FCRR, 2006a). Sound segmentation and identification are assessed through Category I-A, which requires the individual to listen to 2- or 3-phoneme sequences and then place colored blocks to correspond to the phoneme sequences heard (Bountress & Cox). Category I-B mirrors Category I-A, but presents segments of six phonemes and requires the child to place the blocks in order of sound presentation, through the addition, substitution, omission, shift or repetition of block placements. In terms of predictive validity, correlations of the LAC with the reading and spelling portions of the Wide Range Achievement Test (WRAT) (Wilkinson, 1993) yielded a range of values from .66 to .81 (Bountress & Cox). Results suggest that the LAC demonstrates value in the prediction of spelling and reading performance. However, this caution is noted: children below the age of seven may have difficulty with the skills necessary to respond to the task demands. Yet the LAC may be suitable for any age, if the subject understands the concepts of sameness, difference, numbers up to four, and left-right orientation (FCRR).

Diagnostic Inventory of Early Development—Third Edition (DIAL-3)

The DIAL-3, the most recent version of the instrument, is an individually administered screening test, which aims to identify children who are in need of further assessment (Cizek & Fairbank, 1998; Mardell-Czudnowski & Goldenberg, 1990). This instrument was the most

widely used standardized measure in a survey of New York State schools (Costenbader, et al., 2000). The DIAL-3 comprises five major subtest areas (Cizek & Fairbank; Mardell-Czudnowski & Goldenberg). Two of the subtests relate to the skills that have been identified as correlating with later reading performance. The Language subtest assesses both receptive and expressive language, and targets phonemic awareness. The Concepts subtest assesses knowledge of basic concepts such as counting and colors, and involves rapid automatic naming. However, the Concepts subscales do not require letter-naming, which has been more closely associated with later reading performance than is color or object naming (Blachman, 1984).

Internal consistency of the Language and Concepts subtests has been reported to be adequate (.70 and .84 respectively). No predictive validity has been reported (Cizek & Faribank, 2001; Mardell-Czudnowski & Goldenberg, 1998), making it difficult to judge the adequacy of this instrument for identification of at-risk readers.

Brigance K & 1

The Brigance K & 1 Screen is a criterion-referenced assessment designed for kindergarten and first grade students (Brigance, 1992), and was used by 13% of the New York State schools surveyed (Costenbader, 2000). The kindergarten level comprises of 13 skill tests (Brigance). Broad skill areas are assessed, including: motor, body awareness, number, language, and auditory and visual discrimination. The language component of the screening tool addresses both receptive and expressive language.

The Brigance K & 1 Screen technical manual does not report any information regarding reliability or validity (Berk & Watson, 1995). Because of the deficiencies in reported psychometric data, Berk and Watson suggest that educational decisions should not be made

based on information from the Brigance. Furthermore, it is recommended that the tool be used only as an informal screening measure.

Gesell School Readiness Test

It was reported that the Gesell School Readiness Test (Ilg & Ames, 1972), was used by 13% of New York State schools in the screening process (Costenbader et al., 2000). This assessment tool is a developmental/process instrument that aims to measure a child's developmental levels as a method to predict skill acquisition (Ilg & Ames). Graue & Shepard (1989) found predictive validity of the Gesell Readiness Test to be low (as cited in Costenbader et al.). The test, which was published in 1972, is outdated and the psychometrics are poor. Given the age of the Gesell School Readiness Test, as well as the poor psychometrics, the use of this instrument in screening Kindergarten children for potential reading difficulties is problematic.

Locally Made Screenings

Typical screening batteries are only moderately successful in predicting those children who will develop reading problems (Rafoth, 1997). Therefore it may not be a surprise that 30% of the schools surveyed in New York State used locally made screening batteries; this is a figure exceeding that of any of the commercially available measures (Costenbader et al., 2000). After a thorough investigation, The Committee on the Prevention of Reading Difficulties in Young Children made the recommendation that government agencies and private organizations sponsor research to develop screening instruments that are more efficient and sensitive to the identification of children at risk for developing reading difficulties (Snow et al., 1998).

Limited English Proficiency

The number of children in U.S. schools who speak other languages and who have Limited English Proficiency (LEP) has risen dramatically from the 1980's and continues to increase (Capps, Fix, Murray, Ost, Passel, & Herwanto, 2005; Snow et al., 1998). In 2000, the highest number of LEP students was found in kindergarten; it was estimated to be 10% of the total kindergarten population (Capps et al.). Additionally, 8% of kindergarten students lived in linguistically isolated homes, where individuals over 14 demonstrated LEP. Children with limited English proficiency are likely to begin school with weaknesses in areas associated with literacy knowledge, and consequently are likely to fall behind in the primary grades (Snow et al.).

Socioeconomic Status

Discrepancies in children's literacy accomplishments that are a result of Socioeconomic status (SES) are prominent (Snow et al., 1998). Children from economically disadvantaged neighborhoods are likely to begin school lacking literacy related skills and knowledge of print, increasing the chances that they will fall behind from the start. The National Assessment of Educational Progress noted that the SES-related effects on achievement continue to accrue, and that literacy achievement of students in affluent suburbs is significantly and routinely higher than the literacy achievement of students in underprivileged urban settings (as cited in Snow et al.).

It is clear that current practices in the detection of students who are at-risk for developing reading difficulties are problematic. The use of published batteries has proved to be especially problematic (Rafoth, 1997); this has led many school districts to develop their own screening batteries. However, the effectiveness of such locally developed screening batteries

must be validated for the early detection of students who are at-risk of later reading difficulties. Therefore it is essential to investigate the components of locally made screening instruments to ensure the inclusion of skills that have been shown to have high predictive validity for the identification of students who may be at-risk for later reading problems . The proposed study will investigate the relationship between skills measured in kindergarten and later reading performance to determine if a significant relationship exists, and consequently determine if the screening tool is adequate for use as a predictor of later reading achievement.

Research Questions

The first research question will examine the relationship between Kindergarten Screening Summary Score levels and Fall DIBELS measures of ISF, LNF and the DIBELS Summary Score levels in order to investigate the prediction of at-risk status. It is hypothesized that the Kindergarten Summary Score levels will demonstrate some degree of effective prediction of at-risk status as reflected in the fall DIBELS score levels.

The second research question will examine the relationship between Kindergarten Screening Decisions and Kindergarten DIBELS measures with Grades 1 and 2 Winter DIBELS ORF levels. It is hypothesized that the Kindergarten DIBELS measures will demonstrate a stronger relationship to Grades 1 and 2 Winter DIBELS ORF levels than will the Kindergarten Screening Decision.

The third research question will examine intervention efficiency by comparing Grades 2 and 3 Winter DIBELS ORF levels and fall and spring MAP score levels with DIBELS Summary Score levels. It is hypothesized that intervention efficiency will be demonstrated by changes in students' at-risk status between fall and spring in each grade level and between grades 2 and 3.

The fourth research question is designed to examine the effect of SES on prediction of at-risk status, and on intervention effectiveness. It is hypothesized that prediction of at-risk status and the demonstration of intervention effectiveness will vary based on SES level.

Chapter 2

Methods

Participants

This retrospective study will involve male and female students currently enrolled in the third, fourth and fifth grades during the 2007-2008 school year. Students attend a suburban school in eastern Pennsylvania. According to the 2000 United States Census, the estimated population of the borough was 10,476 with 87.8% reported to be White Non-Hispanic, 6.3% Black or African American, 5.6% Hispanic; 2.4% reported to be two or more races, 1.8% other races, 1.6% Asian, 0.6% American Indian and Alaska Native (United States Census Bureau, 2000). Homes reporting a primary language other than English were reported at 10.2%. Median household income was not reported during the 2000 census; however, 9.1% of families were reported to be living below the poverty level. According to the district, more than 8,000 students are served per year in grades Kindergarten through 12 (East Stroudsburg Area School District Report Card).

*Measures**School-made Kindergarten Screening Instrument*

The school-made screening instrument consists of a likert type scale and is administered individually by a certified teacher. Each child was assessed on his or her performance in the following categories: a discussion with the examiner, demonstration of counting to ten, identification of presented colors, number recognition for numerals 1-10, generation of his or her printed name, identification of letter names, replication of a design, concepts of print, generation of a human figure drawing, and attention span. The examiner assigns a score for each category based on prescribed criteria. Minor variations in the assessment occur between administration

years, and yield overall scores ranging from 46-58 points. The numerical scores were collapsed by the examiner and corresponded to categories of High, Middle, Low and At-Risk performances. For the purposes of statistical analyses for this study, scores were further collapsed into At-Risk (aggregating scores in the High and Middle categories) and Not at-Risk (aggregating scores in the Low and At-Risk categories) categories. Additional information collected but not used directly in the statistical analysis of data included gender, handedness, recommendation for Speech, Occupational or Physical Therapy, or English as a Second Language services.

DIBELS

Fall, winter and spring DIBELS scores were obtained for each student included in the study. A DIBELS Summary Score was reported and was derived by the Resica elementary School staff for use as a screening decision-making tool. The DIBELS Kindergarten Summary Score (DIBELS KSUM) represents a single score based on the combination of all of the DIBELS subtests that were administered in the fall of the Kindergarten school year. KSUM Scores were initially recorded as HR, MR, LR and AA as per the DIBELS administration and scoring guidelines. For the purposes of this study, scores were further collapsed into dichotomous categories of At-risk and Not At-Risk. Subjects who achieved HR, MR, and LR category scores were considered at-risk, but students who achieved AA scores were considered to be not at-risk. See Chapter I: Use of Instruments for operational definitions. The relationship between ORF and reading achievement has been consistently demonstrated in studies typically using the winter ORF scores for analysis (Shapiro, et al., 2006). Therefore, for this study the DIBELS winter ORF score was used as a measure of reading achievement in grades 1 through 5.

Measures of Academic Performance

The Measures of Academic Performance (MAP) is a computerized adaptive assessment tool that dynamically measures the performance of the student by calibrating item selection for the individual student to determine the performance level (NWEA, 2003). If a student answers a question incorrectly, the subsequent question is slightly easier, or conversely, if a student answers a question correctly, the subsequent question is slightly harder. This process continues until the completion of the test, allowing for a specific measure of the student's actual achievement level. The MAP produces error bands which correspond to categories identified as low, high or advanced. As the test progresses these error bands are narrowed and upon completion, highly specific information is provided regarding the student's performance.

The process of utilizing the MAP involves several steps and includes test design, definition of content, item selection and test production (NWEA, 2003). MAP tests are designed specifically for an agency or school district, allowing for unique goals to be assessed. Most MAP assessments include roughly four to eight goals, each with five to six sub-goals and are typically curriculum driven.

No time limit is set for completion of the MAP; students are not permitted to skip any items and are unable to return to previously administered items. The systems allows for four administrations per student per year. Upon completion of the test, the student's score and individualized goals appear on the screen. Reports can be generated for individual students, for classes, for grade levels or for entire districts. Scores are reported as Rasch unit equal interval scores (RIT) typically ranging between 150 and 300. Standard error of measurement (SEM) is reported between 2.5-3.5 RIT points. In addition to RIT scores, percentile ranks are provided and

collapsed into categories of high, average and low. High scores are categorized by performances at or above the 67th percentile; average performances fall between the 34th and 66th percentiles and low performances are at or below the 33rd percentile.

The MAP demonstrates acceptable concurrent validity when compared with the Pennsylvania System of School Assessment (PSSA) for Grades 5 and 8 and demonstrates acceptable concurrent validity, $r = .84$ (NWEA, 2003). Additionally, the MAP test is reported as being highly and consistently correlated with other measures of academic achievement used by a variety of states. Studies regarding reliability for the MAP demonstrated strong findings with test-retest reliability for reading in the spring of 2002, ranging from .84-.91 for grades 2-10.

For the purpose of this study the MAP is aligned with the Pennsylvania State Standards Assessment (PSSA). At each grade level the MAP assesses four areas of Reading. Text Structure and Vocabulary, Comprehension Strategies, Read Critically in Content Areas, and Read, Analyze and Interpret Literature compose the four areas assessed by the MAP.

Pennsylvania State Standard 1.1, Learning to Read Independently, is broad in nature and includes the specific skills of Comprehension and Interpretation as well as Vocabulary Development. Comprehension and Interpretation and Vocabulary Development are assessed as two distinct categories by the MAP. Pennsylvania State Standard 1.2, Reading Critically in All Content Areas, is composed of five smaller units. Standard 1.2 includes Detail, Inferences, Fact from Opinion, Comparison, and Analysis and Evaluation. The MAP questions in the category of Reading Critically in the Content Areas correspond to each of the target areas delineated by the State Standards. Pennsylvania State Standard 1.3, Reading, Analyzing and Interpreting Literature, comprises four smaller units. Standard 1.3 includes, Literary Elements, Literary

Devices, Poetry and Drama. The MAP questions in the category of Reading, Analyzing and Interpreting Literature correspond to each of the target areas as outlined by the State Standards. Scores are earned for each of these categories and are recorded as low, average or high.

Scoring for the MAP is as follows: Low scores correspond to scores below the 34th percentile; average scores fall between the 34th and 66th percentile and high scores are above the 66th percentile. For purposes of statistical analyses, the MAP scores were further collapsed into At-Risk (scores in the MAP low category) and Not At-Risk (scores from the average and high MAP categories).

Procedures

Student names were removed from the data file and replaced with identification numbers to ensure confidentiality. Demographic data in the student files included age, gender, SES, and student status related to participation in remedial services, participation in special education, speech and language services and English as a Second Language services. Scores from the district, school-made Kindergarten screening instrument included in the data file were collapsed into the categories of High, Middle, Low and At-Risk, based on performance and were further collapsed into the At-Risk and Not At-Risk categories described in the previous section. DIBELS scores for the fall season of Kindergarten grade and for the winter season of each successive grade were retained for statistical analyses and collapsed into categories of High Risk, Moderate Risk, Low Risk and Above Average (Kaminski & Good, 1996) and further collapsed into At-Risk and Not At-Risk categories as described in the previous section. MAP data for fall and spring in grade 3 and each successive grade level were retained for statistical analyses and collapsed into categories of At-Risk, Low, Proficient and High Proficient and further collapsed into At-Risk and Not At-Risk categories as described in the previous section.

Statistical Analyses

Statistical analyses addressing Research Questions 1 through 4 involved the construction of 2 x 2 crosstabulation tables as shown in Figure 1 and the calculation of one or more of the following indices (also shown in Figure 1): Sensitivity, Specificity, False Positives, False Negatives, Kappa, Intervention Efficiency, and Instructional Stability.

Figure 1

Indices used in Statistical Analyses of Data

		DIBELS or MAPS Criterion Variable Category	
		At-Risk	Not At-Risk
K Screen or DIBELS Fall K Summary Score	At-Risk	A	B
	Not At-Risk	C	D

$$\text{Sensitivity} = (A/(A+C)) \times 100$$

$$\text{Specificity} = (D/(B+D)) \times 100$$

$$\text{Kappa} = ((p_o - p_e)/(1 - e)) \times 100 \text{ where:}$$

$$p_o = p_A + p_D$$

$$p_e = ((p_A+p_C)(p_A+p_B)) + ((p_B+p_D)(p_C+p_D))$$

$$p_A = A/\text{Total N} \quad p_B = B/\text{Total N} \quad p_C = C/\text{Total N} \quad p_D = D/\text{Total N}$$

$$\text{Overall Consistency} = ((A + D)/\text{Total N}) \times 100$$

$$\text{Intervention Efficiency} = (B/(A+B)) \times 100$$

$$\text{Instructional Stability} = (D/(C+D)) \times 100$$

Operational definitions for the indices used to analyze data and interpret findings are as follows:

Sensitivity: Sensitivity was operationally defined as the percent of students categorized as At-Risk based on Kindergarten Screening Summary Scores or DIBELS Kindergarten Summary Scores who were also categorized as At-Risk, based on scores from a comparison measure (i.e., with a DIBELS ORF score or a MAP score).

Specificity: Specificity was operationally defined as the percent of students categorized as Not At-Risk based on Kindergarten Screening Summary Scores or DIBELS Kindergarten Summary scores who were also categorized as Not At-Risk, based on scores from a comparison measure (i.e., with a DIBELS ORF score or a Map score).

Kappa: The Kappa statistic indicates the percentage of increase over chance level represented by the sensitivity and specificity values obtained from the same data.

Overall Consistency: Overall Consistency was operationally defined as the percentage of students whose category classification on one measure was consistent with their category classification on the comparison measure. Overall Consistency reflects a combination of the percentage of students categorized as at-risk on both measures, with the percentage of students categorized as not at-risk on both measures.

Intervention Efficiency: Intervention Efficiency was operationally defined as the percent of students categorized as At-Risk based on Kindergarten Screening Summary Scores who, after intervention efforts to improve reading skill development, were categorized as Not At-Risk, based on scores from achievement criterion measures (i.e., DIBELS winter ORF scores or MAP fall and spring scores in grades 3, 4, and/or 5). Intervention Efficiency represents the success of

intervention efforts in later grades with students identified in kindergarten as At-Risk of developing reading problems.

Instructional Stability: Instructional Stability was operationally defined as the percent of students categorized as Not At-Risk based on Kindergarten Screening Summary Scores who, after general education instructional efforts to develop reading skills, continued to be categorized as Not At-Risk, based on scores from achievement criterion measures (i.e., DIBELS winter ORF scores and MAP fall and spring scores in grades 3, 4, and/or 5). Instructional Stability represents the success of general education instructional efforts in ensuring that students identified in kindergarten as not at risk of developing reading problems remain successful in later grades.

Analyses were conducted only with the data from students who had complete data sets (i.e., no missing data in any school year). Although this inclusionary criterion eliminated a sizeable number of students from the data set, it enabled meaningful comparisons of changes in category membership across time for the remaining students because the test scores being analyzed in successive years were derived from the same group of students.

Some of the analyses were completed using all of the students in the data base with complete data, i.e., aggregating across the entire sample at a specific grade level. All of the analyses were completed using separate Kindergarten class cohorts labeled Kindergarten 2002-2003, Kindergarten 2003-2004, and Kindergarten 2004-2005. The specific test score variables used in analyses for each cohort are shown in Figure 2. All cohorts were assessed with the DIBELS at all grade levels and two of the three cohorts were assessed with the MAP in the fall and spring of grade 3.

Figure 2

Test Score Variables Used in Analyses for Each Kindergarten Cohort

	By Grade Test Scores Used in Analyses			
Year Cohort Entered Kindergarten	K	1	2	3
2002-2003	K Screen DIBELS Fall	DIBELS ORF Winter	DIBELS ORF Winter	DIBELS ORF Winter
2003-2004	K Screen DIBELS Fall	DIBELS ORF Winter	DIBELS ORF Winter	DIBELS ORF Winter MAP Fall & Spring
2004-2005	K Screen DIBELS Fall	DIBELS ORF Winter	DIBELS ORF Winter	DIBELS ORF Winter MAP Fall & Spring
All Cohorts	K Screen DIBELS Fall			

Chapter 3

Results

*Research Question 1**Kindergarten Screening Measure and DIBELS: Prediction of At-risk readers*

To investigate the relationship between Kindergarten Screening Summary Scores obtained prior to entry into Kindergarten and DIBELS ISF, LNF and Summary Scores obtained in fall of Kindergarten year, crosstabulation tables were constructed. The data from the crosstabulation tables were used to calculate Sensitivity, Specificity, Kappa and Overall Consistency values. Results of the analyses are reported in Table 1 for all three cohorts combined, as well as for the three separate Kindergarten cohorts. As noted in Chapter 2, only students without missing data were used in these and all subsequent analyses.

*Table 1**Relationship between Kindergarten Screening Measure and Fall DIBELS Assessment*

Variable	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Kappa	Overall Consistency
All Groups Combined				
Initial Sound Fluency (n = 176)	79%	52%	27%	65%
Letter Naming Fluency (n = 176)	84%	53%	32%	68%
DIBELS Sum Decision (n = 174)	80%	65%	46%	74%

Kindergarten 2002-2003 (n = 41)

Initial Sound Fluency	69%	50%	16%	59%
Letter Naming Fluency	71%	29%	24%	62%
DIBELS Sum Decision	70%	61%	31%	66%

Kindergarten 2003-2004 (n = 60)

Initial Sound Fluency	74%	46%	18%	60%
Letter Naming Fluency	95%	54%	39%	73%
DIBELS Sum Decision				

Kindergarten 2004-2005 (n = 75)

Initial Sound Fluency	89%	58%	42%	73%
Letter Naming Fluency	86%	53%	30%	67%
DIBELS Sum Decision	86%	72%	59%	80%

*Research Question 2**Kindergarten Screening Decision and Grades 1 and 2 Winter Oral Reading Fluency: Prediction of At-Risk Readers*

To investigate the relationships between Kindergarten Screening Summary Scores with Grades 1 and 2 Winter ORF scores and Kindergarten DIBELS measures with Grades 1 and 2 Winter ORF scores cross tabulations were constructed and values were calculated for Sensitivity, Specificity, Kappa and Overall Consistency. Analyses were conducted separately for each

kindergarten cohort. Results of the analyses are presented in Table 2 (Grade 1 ORF comparisons) and Table 3 (Grade 2 ORF comparisons).

Table 2

Relationship between Kindergarten Screening Measure and Kindergarten DIBELS Measures with Winter Grade 1 Oral Reading Fluency

Variable	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Kappa	Overall Consistency
Kindergarten 2002-2003 (n = 40)				
K Screening Sum	67%	55%	21%	61%
Letter Naming Fluency	67%	77%	44%	73%
Initial Sound Fluency	44%	77%	22%	63%
DIBELS Sum Decision	78%	64%	41%	36%
Kindergarten 2003-2004 (n = 60)				
Kindergarten Decision	91%	57%	43%	74%
Letter Naming Fluency	61%	86%	49%	41%
Initial Sound Fluency	48%	68%	15%	58%
DIBELS Sum Decision	74%	62%	34%	67%
Kindergarten 2004-2005 (n = 73)				
Kindergarten Decision	83%	54%	30%	76%
Letter Naming Fluency	74%	92%	67%	85%
Initial Sound Fluency	48%	70%	17%	58%
DIBELS Sum Decision	91%	69%	53%	76%

Table 3

Relationship between Kindergarten Screening Measure and Kindergarten DIBELS Measures with Grade 2 Winter Oral Reading Fluency

Variable	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Kappa	Overall Consistency
Kindergarten 2002-2003 (n = 40)				
Kindergarten Decision	86%	62%	42%	72%
Letter Naming Fluency	71%	73%	42%	71%
Initial Sound Fluency	43%	73%	16%	58%
DIBELS Sum Decision	71%	54%	22%	62%
Kindergarten 2003-2004 (n = 60)				
Kindergarten Decision	88%	49%	27%	66%
Letter Naming Fluency	65%	81%	45%	73%
Initial Sound Fluency	47%	65%	11%	56%
DIBELS Sum Decision	71%	56%	21%	61%
Kindergarten 2004-2005 (n = 73)				
Kindergarten Decision	86%	54%	30%	67%
Letter Naming Fluency	57%	83%	40%	70%
Initial Sound Fluency	52%	71%	22%	61%
DIBELS Sum Decision	81%	63%	36%	68%

*Research Question 3**Effectiveness of Reading Instruction*

Percentages of students identified as at-risk in Kindergarten with the DIBELS K Sum score, at-risk in Grades 1, 2, and 3 with the DIBELS Oral Reading Fluency scores, and in Grade 3 Fall and Spring with MAP domain scores were calculated to provide a frame of reference for the additional analyses conducted for the purposes of answering Research Question 3. At-risk percentages for each cohort are presented in Table 4 for the DIBELS criterion measures and in Table 5 for the MAP criterion measures.

*Table 4**Percentage of Students At-Risk based on Winter Oral Reading Fluency*

Cohort	Grade Level	Percent At-Risk
2002-2003	Kindergarten	55%
2002-2003	Grade 1	45%
2002-2003	Grade 2	35%
2002-2003	Grade 3	53%
2003-2004	Kindergarten	52%
2003-2004	Grade 1	38%
2003-2004	Grade 2	28%
2003-2004	Grade 3	32%
2004-2005	Kindergarten	50%
2004-2005	Grade 1	32%
2004-2005	Grade 2	29%
2004-2005	Grade 3	26%

Table 5

Percentage of Students At-Risk based on Grade 3 MAP Scores

Cohort	MAP Category	Percentage At-Risk
2002-2003	Fall Vocabulary	85%
	Fall Comprehension	75%
	Fall Analyze	75%
2003-2004	Fall Vocabulary	75%
	Spring Vocabulary	43%
	Fall Comprehension	72%
	Spring Comprehension	45%
	Fall Critical Content	65%
	Spring Critical Content	38%
	Fall Analyze	72%
	Spring Analyze	42%
2004-2005	Fall Vocabulary	38%
	Spring Vocabulary	26%
	Fall Comprehension	29%
	Spring Comprehension	19%
	Fall Critical Content	19%
	Spring Critical Content	15%
	Fall Analyze	22%
	Spring Analyze	17%

Additional analyses were conducted to determine the effectiveness of reading interventions implemented after students were identified in kindergarten as at-risk and to determine the stability of instruction offered to students who were identified in kindergarten as not at-risk. Analyses included the calculation of values for intervention efficiency, instructional stability, sensitivity, specificity, and kappa levels.

As noted in Chapter 2, Intervention Efficiency was operationally defined as the percentage of students identified as at-risk in Kindergarten (in this case, based on the category assignment from the DIBELS Summary Score obtained in the fall of the kindergarten year) who received remedial instruction and who, in subsequent years, were able to earn a score in the proficient range on a criterion measure reflecting reading skill development. Instructional Stability was operationally defined as the percentage of students identified as not at-risk in kindergarten who, in subsequent years, were able to earn a score in the proficient range on a criterion measure reflecting reading skill development. In these analyses, two separate criterion measures were used to reflect reading skill development: DIBELS Oral Reading Fluency scores obtained in winter of the school year in grades 2 and 3, and MAP scores obtained in the fall and the spring of the grade 3 school year.

Crosstabulation tables were constructed for each comparison and values were calculated for Intervention Efficiency, Instructional Stability, Screening Sensitivity, Screening Specificity, and Kappa. Analysis results are shown in Table 6 for comparisons of DIBELS K Sum decisions with DIBELS Oral Reading Fluency measures in Grades 2, and 3 and MAP domain scores in the fall and winter of Grade 3.

Table 6

Kindergarten DIBELS Summary Score Compared with Winter Grades 2 and 3 ORF and Grade 3 Fall and Spring MAP Scores

Variable	Intervention Efficiency	Instructional Stability	Screening Sensitivity	Screening Specificity	Kappa
Kindergarten 2002-2003 (n = 40)					
<u>DIBELS Scores in Grade 2</u>					
Winter Oral Reading Fluency	55%	78%	71%	54%	22%
<u>DIBELS Scores in Grade 3</u>					
Winter Oral Reading Fluency	27%	72%	76%	68%	45%
<u>MAP Scores in Grade 3</u>					
Fall Vocabulary	64%	83%	73%	52%	19%
Fall Comprehension	14%	39%	63%	70%	26%
Fall Analyze	14%	39%	63%	70%	26%
Kindergarten 2003-2004 (n = 60)					
<u>DIBELS Scores in Grade 2</u>					
Winter Oral Reading Fluency	61%	83%	71%	56%	21%
<u>DIBELS Scores in Grade 3</u>					
Winter Oral Reading Fluency	42%	76%	72%	63%	34%

MAP Scores in Grade 3

Fall Vocabulary	13%	38%	60%	73%	25%
Spring Vocabulary	42%	72%	69%	62%	30%
Fall Comprehension	13%	45%	63%	76%	32%
Spring Comprehension	35%	76%	74%	67%	40%
Fall Critical Content	19%	52%	64%	71%	33%
Spring Critical Content	52%	72%	65%	57%	21%
Fall Analyze	13%	45%	63%	76%	32%
Spring Analyze	42%	76%	72%	63%	34%

Kindergarten 2004-2005 (n = 73)

DIBELS Scores in Grade 2

Winter Oral Reading Fluency	53%	47%	81%	63%	36%
-----------------------------	-----	-----	-----	-----	-----

DIBELS Scores in Grade 3

Winter Oral Reading Fluency	44%	56%	23%	83%	39%
-----------------------------	-----	-----	-----	-----	-----

MAP Scores in Grade 3

Fall Vocabulary	44%	81%	74%	64%	36%
Spring Vocabulary	67%	81%	63%	55%	14%
Fall Comprehension	53%	89%	81%	63%	36%
Spring Comprehension	74%	89%	71%	53%	15%
Fall Critical Content	69%	92%	79%	57%	22%
Spring Critical Content	78%	92%	73%	54%	14%
Fall Analyze	64%	92%	73%	54%	28%
Spring Analyze	78%	89%	67%	53%	11%

*Research Question 4**Socioeconomic Status*

To investigate the relationship of SES with prediction of at-risk readers, and the effectiveness of intervention, the analyses described in Research Questions 1 through 3 were repeated for two separate SES groups labeled Low SES and High SES. The Low SES group was operationally defined as those students identified as receiving free or reduced lunch during 2004-2005 school year. The High SES group was operationally defined as those students identified as not receiving free or reduced lunch during the 2004-2005 school year.

*Research Question 4a**Kindergarten Screening Measure and Fall DIBELS Assessment with SES*

To investigate the relationships between the Kindergarten Screening score and Kindergarten DIBELS measures with SES as a moderating factor, cross tabulations were constructed and values were calculated for Sensitivity, Specificity, Kappa and Overall Consistency. Analyses were conducted separately for each Kindergarten cohort. Results of the analyses are presented in Table 7.

*Table 7**Kindergarten Screening Measure Compared with Fall DIBELS Measures by SES Group*

	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Overall Consistency	Kappa
<u>All Groups Combined</u>				
Initial Sound Fluency				
High SES (n = 115)	69%	58%	62%	24%
Low SES (n = 55)	93%	39%	68%	30%

Letter Naming Fluency

High SES (n = 115) 81% 61% 67% 34%

Low SES (n = 60) 88% 34% 65% 20%

DIBELS Sum Decision

High SES (n = 113) 73% 69% 71% 42%

Low SES (n = 60) 90% 52% 76% 45%

Kindergarten 2002-2003

Initial Sound Fluency

High SES (n = 28) 43% 57% 50% 0%

Low SES (n = 13) 100% 29% 78% 27%

Letter Naming Fluency

High SES (n = 28) 60% 67% 63% 26%

Low SES (n = 13) 86% 17% 53% 2%

DIBELS Sum Decision

High SES (n = 28) 54% 67% 61% 21%

Low SES (n = 13) 90% 33% 66% 26%

Kindergarten 2003-2004

Initial Sound Fluency

High SES (n = 40) 67% 48% 57% 13%

Low SES (n = 20) 88% 42% 67% 26%

Letter Naming Fluency

High SES (n = 40) 100% 57% 72% 40%

Low SES (n = 20) 89% 45% 70% 33%

DIBELS Sum Decision

High SES (n = 40)	75%	60%	68%	35%
Low SES (n = 28)	91%	56%	77%	48%

Kindergarten 2004-2005**Initial Sound Fluency**

High SES (n = 47)	86%	67%	72%	44%
Low SES (n = 27)	92%	43%	73%	35%

Letter Naming Fluency

High SES (n = 47)	83%	63%	68%	35%
Low SES (n = 27)	89%	33%	63%	17%

DIBELS Sum Decision

High SES (n = 45)	83%	78%	80%	59%
Low SES (n = 27)	89%	56%	76%	47%

Research Question 4b

Kindergarten Screening Measures and Kindergarten DIBELS scores compared with Grade 1 and Grade 2 Winter ORF scores by SES group

To investigate the relationships between Kindergarten Screening Summary Scores with Grades 1 and 2 Winter ORF scores, having SES as a moderating factor and Kindergarten DIBELS measures with Grades 1 and 2 Winter ORF scores, having SES as a moderating factor, cross tabulations were constructed and values were calculated for Sensitivity, Specificity, Kappa and Overall Consistency. Analyses were conducted separately for each Kindergarten cohort.

Results of the analyses are presented in Table 8 (Grade 1 ORF comparisons) and Table 9 (Grade 2 ORF comparisons).

Table 8

Kindergarten Screening Measure and Kindergarten DIBELS Compared with Grade 1 Winter ORF by SES Group

Variable	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Overall Consistency	Kappa
Kindergarten 2002-2003				
Kindergarten Screener Decision				
High SES (n = 28)	55%	65%	60%	19%
Low SES (n = 12)	86%	20%	55%	6%
Letter Naming Fluency				
High SES (n = 28)	55%	76%	66%	32%
Low SES (n = 12)	86%	80%	83%	66%
Initial Sound Fluency				
High SES (n = 28)	36%	82%	62%	20%
Low SES (n = 12)	57%	60%	59%	17%
DIBELS Sum Decision				
High SES (n = 28)	64%	65%	64%	27%
Low SES (n = 12)	100%	60%	89%	64%

Kindergarten 2003-2004

Kindergarten Screener Decision

High SES (n = 40)	91%	55%	69%	34%
Low SES (n = 20)	92%	63%	81%	57%

Letter Naming Fluency

High SES (n = 40)	45%	83%	65%	29%
Low SES (n = 20)	75%	100%	87%	71%

Initial Sound Fluency

High SES (n = 40)	45%	66%	55%	10%
Low SES (n = 20)	50%	75%	63%	23%

DIBELS Sum Decision

High SES (n = 40)	73%	59%	63%	25%
Low SES (n = 20)	75%	75%	75%	49%

Kindergarten 2004-2005

Kindergarten Screener Decision

High SES (n = 45)	79%	66%	69%	38%
Low SES (n = 27)	57%	8%	27%	-36%

Letter Naming Fluency

High SES (n = 45)	71%	94%	86%	68%
Low SES (n = 27)	78%	89%	84%	67%

Initial Sound Fluency

High SES (n = 45)	50%	81%	67%	32%
Low SES (n = 27)	44%	50%	48%	-5%

DIBELS Sum Decision

High SES (n = 45)	86%	81%	80%	62%
Low SES (n = 27)	100%	50%	75%	40%

Table 9

Kindergarten Screening Measure and Kindergarten DIBELS Compared with Grade 2 Winter ORF by SES Group

Variable	Sensitivity At-Risk Agreement	Specificity Not At-Risk Agreement	Overall Consistency	Kappa
Kindergarten 2002-2003				
Kindergarten Screener Decision				
High SES (n = 28)	75%	70%	69%	39%
Low SES (n = 12)	100%	33%	80%	33%
Letter Naming Fluency				
High SES (n = 28)	63%	75%	67%	35%
Low SES (n = 12)	83%	67%	76%	50%
Initial Sound Fluency				
High SES (n = 28)	25%	75%	50%	0%
Low SES (n = 12)	67%	67%	67%	33%
DIBELS Sum Decision				
High SES (n = 28)	63%	60%	59%	19%
Low SES (n = 12)	83%	33%	62%	17%
Kindergarten 2003-2004				
Kindergarten Screener Decision				
High SES (n = 40)	88%	50%	62%	22%
Low SES (n = 20)	89%	45%	70%	33%
Letter Naming Fluency				
High SES (n = 40)	50%	81%	64%	29%
Low SES (n = 20)	78%	82%	80%	60%

Initial Sound Fluency

High SES (n = 40)	38%	63%	50%	0%
Low SES (n = 20)	56%	73%	65%	29%

DIBELS Sum Decision

High SES (n = 40)	63%	53%	55%	10%
Low SES (n = 20)	78%	64%	71%	41%

 Kindergarten 2004-2005
Kindergarten Screener Decision

High SES (n = 45)	77%	64%	67%	34%
Low SES (n = 27)	100%	37%	70%	26%

Letter Naming Fluency

High SES (n = 45)	54%	85%	70%	40%
Low SES (n = 27)	63%	79%	70%	40%

Initial Sound Fluency

High SES (n = 45)	46%	79%	63%	25%
Low SES (n = 27)	63%	58%	70%	17%

DIBELS Sum Decision

High SES (n = 45)	77%	75%	73%	47%
Low SES (n = 27)	88%	42%	59%	22%

*Research Question 4c**Effectiveness of Reading Instruction by SES*

Research question 3 was re-examined with SES as a moderating factor. As was the case with Question 3, percentages of students identified as at-risk in Kindergarten with the DIBELS K Sum score, at-risk in Grades 1, 2, and 3 with the DIBELS Oral Reading Fluency scores, and in Grade 3 Fall and Spring with MAP domain scores were calculated separately for the High SES and Low SES groups of each cohort to provide a frame of reference for the additional analyses conducted for the purposes of answering Research Question 4. At-risk percentages for each cohort are presented in Table 10 for the DIBELS criterion measures and in Table 11 for the MAP criterion measures.

*Table 10**Percentage of At-Risk Students, by SES Group based on DIBELS Measures*

Cohort	Grade Level	Percent At-Risk
2002-2003		
DIBELS K Sum Score		
High SES	Kindergarten	46%
Low SES	Kindergarten	75%
DIBELS Oral Reading Fluency Score		
High SES	First Grade	39%
Low SES	First Grade	58%
High SES	Second Grade	29%
Low SES	Second Grade	50%
High SES	Third Grade	46%
Low SES	Third Grade	67%

2003-2004**DIBELS K Sum Score**

High SES	Kindergarten	50%
Low SES	Kindergarten	55%

DIBELS Oral Reading Fluency Score

High SES	First Grade	28%
Low SES	First Grade	60%
High SES	Second Grade	20%
Low SES	Second Grade	45%
High SES	Third Grade	38%
Low SES	Third Grade	50%

2004-2005**DIBELS K Sum Score**

High SES	Kindergarten	40%
Low SES	Kindergarten	67%

DIBELS Oral Reading Fluency Score

High SES	First Grade	31%
Low SES	First Grade	33%
High SES	Second Grade	29%
Low SES	Second Grade	30%
High SES	Third Grade	31%
Low SES	Third Grade	44%

Table 11

Percentage of At-Risk Students, by SES Group based on Grade 3 MAP Scores

Cohort	Percentage At-Risk High SES	Percentage At-Risk Low SES
<u>Kindergarten 2002-2003</u>		
Fall Vocabulary	82%	92%
Fall Comprehension	68%	92%
Fall Analyze	68%	92%
<u>Kindergarten 2003-2004</u>		
Fall Vocabulary	70%	85%
Spring Vocabulary	38%	55%
Fall Comprehension	65%	85%
Spring Comprehension	43%	50%
Fall Critical Content	55%	85%
Spring Critical Content	35%	45%
Fall Analyze	63%	90%
Spring Analyze	45%	35%

Kindergarten 2004-2005

Fall Vocabulary	33%	41%
Spring Vocabulary	29%	30%
Fall Comprehension	29%	30%
Spring Comprehension	20%	50%
Fall Critical Content	24%	11%
Spring Critical Content	13%	18%
Fall Analyze	20%	26%
Spring Analyze	13%	22%

Analyses were conducted for the LowSES and High SES groups separately to determine how the effectiveness varied by SES, relative to the reading interventions implemented after students were identified in kindergarten as at-risk. The analyses also examined by SES the stability of instruction offered to students who were identified in kindergarten as not at-risk. The analyses compared Kindergarten DIBELS Summary scores to DIBELS Grades 2 and 3 winter ORF scores and MAP scores in fall and spring of Grade 3.

Crosstabulation tables were constructed for each comparison and values were calculated for Intervention Efficiency, Instructional Stability, Screening Sensitivity, Screening Specificity, and Kappa. Results of Kindergarten DIBELS Summary Score compared with DIBELS winter ORF scores for grades 2 and 3 by SES are reported in Table 12. Results of the Kindergarten DIBELS Summary Scores compared with MAP scores by SES are presented in Table 13.

Table 12

Kindergarten DIBELS Summary Score Compared with Winter Grade 2 and Grade 3 DIBELS ORF by SES Group

Variable	Intervention Efficiency	Instructional Stability	Screening Sensitivity	Screening Specificity	Kappa
Kindergarten 2002-2003					
<u>Scores in Grade 2</u>					
Winter Oral Reading Fluency					
High SES (n = 28)	62%	80%	63%	60%	19%
Low SES (n = 12)	44%	67%	83%	33%	17%
<u>Scores in Grade 3</u>					
Winter Oral Reading Fluency					
High SES (n = 28)	31%	73%	69%	73%	43%
Low SES (n = 12)	22%	67%	88%	50%	40%
Kindergarten 2003-2004					
<u>Scores in Grade 2</u>					
Winter Oral Reading Fluency					
High SES (n = 40)	75%	85%	63%	53%	10%
Low SES (n = 20)	36%	78%	78%	64%	41%
<u>Scores in Grade 3</u>					
Winter Oral Reading Fluency					
High SES (n = 40)	45%	80%	73%	64%	35%
Low SES (n = 20)	36%	67%	70%	60%	30%

Kindergarten 2004-2005

Scores in Grade 2**Winter Oral Reading Fluency**

High SES (n = 45)	44%	89%	77%	75%	47%
Low SES (n = 27)	61%	89%	88%	42%	22%

Scores in Grade 3**Winter Oral Reading Fluency**

High SES (n = 45)	44%	85%	71%	74%	42%
Low SES (n = 27)	44%	78%	83%	47%	29%

*Table 13**Kindergarten DIBELS Summary Score and MAP Scores by SES*

Variable	Intervention Efficiency	Instructional Stability	Screening Sensitivity	Screening Specificity	Kappa
Kindergarten 2002-2003					
<u>Scores in Grade 3</u>					
Fall Vocabulary					
High SES (n = 28)	8%	27%	52%	80%	18%
Low SES (n = 12)	0%	33%	82%	100%	43%
Fall Comprehension					
High SES (n = 28)	23%	40%	53%	67%	16%
Low SES (n = 12)	0%	33%	82%	100%	43%
Fall Analyze					
High SES (n = 28)	23%	40%	53%	67%	16%
Low SES (n = 12)	0%	33%	82%	100%	43%

Kindergarten 2003-2004

Scores in Grade 3**Fall Vocabulary**

High SES (n = 40)	20%	40%	57%	67%	20%
Low SES (n = 20)	0%	33%	65%	100%	35%

Fall Comprehension

High SES (n = 40)	15%	55%	65%	79%	40%
Low SES (n = 20)	9%	22%	59%	67%	14%

Fall Critical Content

High SES (n = 40)	25%	65%	68%	72%	40%
Low SES (n = 20)	9%	22%	59%	67%	14%

Fall Analyze

High SES (n = 40)	20%	55%	64%	73%	35%
Low SES (n = 20)	0%	22%	61%	100%	24%

Spring Vocabulary

High SES (n = 40)	45%	80%	73%	64%	35%
Low SES (n = 20)	36%	56%	64%	56%	19%

Spring Comprehension

High SES (n = 40)	35%	80%	76%	70%	45%
Low SES (n = 20)	36%	67%	70%	60%	30%

Spring Critical Content

High SES (n = 40)	55%	75%	64%	58%	20%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 20)	45%	67%	67%	55%	21%
------------------	-----	-----	-----	-----	-----

Spring Analyze

High SES (n = 40)	35%	75%	72%	68%	40%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 20)	55%	78%	71%	54%	22%
------------------	-----	-----	-----	-----	-----

Kindergarten 2004-2005**Scores in Grade 3****Fall Vocabulary**

High SES (n = 45)	39%	80%	69%	76%	43%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 27)	50%	78%	82%	44%	23%
------------------	-----	-----	-----	-----	-----

Fall Comprehension

High SES (n = 45)	50%	85%	69%	72%	37%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 27)	56%	100%	100%	47%	35%
------------------	-----	------	------	-----	-----

Fall Critical Content

High SES (n = 45)	36%	89%	73%	71%	36%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 27)	83%	100%	100%	38%	12%
------------------	-----	------	------	-----	-----

Fall Analyze

High SES (n = 45)	61%	93%	78%	69%	34%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 27)	67%	89%	86%	40%	17%
------------------	-----	-----	-----	-----	-----

Spring Vocabulary

High SES (n = 45)	72%	85%	56%	64%	14%
-------------------	-----	-----	-----	-----	-----

Low SES (n = 27)	61%	67%	70%	35%	5%
------------------	-----	-----	-----	-----	----

Spring Comprehension

High SES (n = 45)	67%	89%	67%	67%	24%
Low SES (n = 27)	78%	89%	80%	36%	8%

Spring Critical Content

High SES (n = 45)	78%	95%	67%	71%	20%
Low SES (n = 27)	78%	89%	80%	36%	8%

Spring Analyze

High SES (n = 45)	83%	89%	50%	62%	6%
Low SES (n = 27)	72%	89%	83%	38%	13%

Chapter 4

Discussion

This chapter presents a summary and discussion of the results of the analyses and the limitations of the study; included are conclusions and recommendations for future research.

Summary of Results

In this section, the results of the research questions are reviewed and the significance of the findings discussed. The first two research questions were designed from the perspective of predicting reading achievement, specifically detecting students at-risk of developing reading problems through the second grade. For the third research question, the emphasis of investigation shifted to intervention efficiency and instructional stability in order to assess the effectiveness of reading instruction. Finally, the research questions are re-examined to determine the relationship of SES with at-risk status and reading skill development.

Research Question 1

Kindergarten Screening Measure and DIBELS: Prediction of At-risk readers

The first research question was designed to examine the relationship between at-risk status determined with the Kindergarten Screening measure prior to entry into kindergarten and at-risk status determined by DIBELS measures administered in the fall of kindergarten.

As shown in Table 1, sensitivity values for the total group ranged from 79% to 84% and specificity values ranged from 52% to 65%. The comparison between the category classifications derived from the Kindergarten Screening Summary Scores and the category classifications derived from the DIBELS Summary Scores produced the best combination of sensitivity (80%) and specificity (65%) values, with a kappa value reflecting a 46% increase over

chance level assignment of at-risk and not at-risk status and an Overall Consistency percentage of 74%.

For the separate kindergarten cohorts, each successive year of entry into kindergarten reflected an increase in agreement between the category classifications derived from the Kindergarten Screening Summary Scores and the category classifications derived from the DIBELS measures. For example, sensitivity values for the K-Screen/DIBELS Summary comparisons were 70% for the 2002-2003 cohort, 81% for the 2003-2004 cohort, and 86% for the 2004-2005 cohort with related kappa values of 31%, 40%, and 59% and overall consistency values of 66%, 71%, and 80% respectively.

Overall, the Kindergarten Screening measure and the DIBELS Summary Score demonstrated the highest levels of agreement in terms of at-risk status. The strength of the relationship between the Kindergarten Screening score and the DIBELS Summary score is not unexpected, because the DIBELS Summary score incorporates all components of the DIBELS screener, and is therefore a more comprehensive score than the individual DIBELS scores. The high level of agreement between the Kindergarten Screening score and the DIBELS Summary Score suggested little change in at-risk status between the time at which the Kindergarten screener was administered in the spring prior to school entry and the time at which the DIBELS data was gathered in the fall of the kindergarten year. The degree of consistency between the Kindergarten Screening score and the DIBELS Summary score could be attributed to the type of items included on the Kindergarten Screener. Although it has been established that locally made screening instruments typically are not the most efficient method of screening the (Rafoth, 1997), efficiency of such screeners improves when items linked to reading readiness skills are included, as was the case with the ESASD Kindergarten Screener used in this study. The

inclusion of reading readiness items likely increased the rate of agreement between the Kindergarten Screening score and the more empirically supported DIBLES measure.

Research Question Two

Kindergarten Screening Decision and Grades 1 and 2 Winter Oral Reading Fluency: Prediction of At-Risk Readers

The second research question was designed to examine the predictive relationship between the Kindergarten Screening measure and Kindergarten DIBELS measures with winter ORF in Grades 1 and 2.

The DIBELS Summary score demonstrated greater levels of sensitivity and greater improvement over chance in predicting Grade 1 winter ORF than did the Kindergarten Screening measure for two of the three cohorts and greater sensitivity than LNF or ISF for all cohorts. Likewise, the DIBELS Summary score achieved greater specificity for all cohorts than did the Kindergarten Screening measure. DIBELS is an empirically supported screening method that has been consistently linked to reading achievement (Good & Kaminski, 2002; Kaminski & Good, 1996; VanDerHyden, Witt, Naquin & Noell, 2001). The comprehensive nature of the DIBELS Summary score and the empirically supported content of the DIBELS screener are the most likely sources for the higher levels of sensitivity and specificity noted. Overall, the DIBELS Summary score performed better as a predictor of risk status related to Grade 1 ORF than the Kindergarten Screening measure. The highest level of overall consistency was demonstrated between DIBELS Grade 1 ORF and DIBELS Kindergarten LNF for two of the three groups. The highest kappa values were obtained with the comparison of the DIBELS Grade 1 Winter ORF and the DIBELS Kindergarten LNF. Consequently, for the children in this study, the Kindergarten DIBELS LNF may be considered the most accurate predictor of DIBELS Grade 1

ORF performance. This relationship makes sense, given the fact that the rapid letter naming task is the Kindergarten DIBELS measure that is the most similar to oral reading fluency (i.e., rapid word reading).

The predictive relationship between the Kindergarten Screening measure and Kindergarten DIBELS measures and winter ORF in Grade 2 was investigated. Comparisons were made between the Kindergarten Screening measure and the DIBELS Summary score and Grade 2 winter ORF. Both the Kindergarten Screening score and the DIBELS Summary score demonstrated strong specificity and sensitivity. However, because of the empirically supported nature of the DIBELS measures, the DIBELS Summary score is considered a better choice for prediction of at-risk status than the Kindergarten Screening measure. Additionally, the highest kappa values were reflected in the prediction of the DIBELS Grade 2 ORF with the DIBELS LNF obtained in the fall of the kindergarten year. For two of the three cohorts the highest overall consistency was reflected in the prediction of the DIBELS Grade 2 ORF with the DIBELS Kindergarten LNF. Therefore the DIBELS LNF demonstrated the strongest predictive relationship with DIBELS Grade 2 winter ORF. This finding is consistent with that of Good & Kaminski's study,(2002) that found that LNF correlated highly with the Woodcock-Johnson Psycho Educational Battery—Revised Reading cluster. These findings suggested that DIBELS LNF boasted a stronger relationship with DIBELS ORF than did DIBELS ISF.

As in the case of the relationship between grade 1 ORF and DIBELS Kindergarten LNF, these findings are not surprising, given the similarity of the DIBELS rapid letter naming and rapid word reading tasks. Theoretically speaking, both of these tasks require similar neuropsychological processes for effective performance, with letter naming being a simpler

precursor to word naming. This theoretical relationship is supported in the empirical evidence of this study.

Research Question Three

Effectiveness of Reading Instruction

The third research question investigated the effectiveness of the reading instruction based on the results of the Kindergarten Summary Scores. This research question was developed, based on the results of research questions 1 and 2. As the research progressed it became evident that the relationships between the Kindergarten Screening measure and the scores on the DIBELS and MAP was likely affected by interventions provided from Kindergarten through grades 2 and 3. Given the high level of progress monitoring, interventions were provided to address academic needs that were evident through progress monitoring. Although the effectiveness of reading instruction will be investigated to some degree here, a detailed discussion of the reading instruction and intervention programs provided at Resica Elementary and the ways in which these programs and interventions were adjusted to meet the needs of at-risk students is beyond the scope of this study. However, it would remiss to not provide comment on the effectiveness of the reading instruction and intervention based on the data collected in this study.

At-Risk status based on DIBELS Measures. At-risk status percentages based on oral reading fluency were reported for each cohort during kindergarten and grades 1, 2, and 3. In all three cohorts, more than 50% of the students were classified as at-risk, based on DIBELS Sum scores obtained in the fall of Kindergarten. (2002-2003 55%; 2003-2004 52%; 2004-2005 50%). Examination of a cohort as it progresses through school years enables the analysis of change over time for the total number of students who are at-risk and provides insight about how

intervention efforts affect the specific cohort. As the individual cohorts progressed from kindergarten through grades 1 and 2, the number of students at-risk decreased considerably, (2002-2003 decreased from 55% in kindergarten to 45% in grade 1 to 35% in grade 2; 2003-2004 decreased from 52% in kindergarten to 38% in grade 1 to 28% in grade 2; 2004-2005 decreased from 50% in kindergarten to 32% in grade 1 to 29% in grade 2). The reduction in at-risk status between grade levels for each cohort shows that each year fewer students remained at-risk, suggesting that intervention and instructional efforts were effective. All cohorts demonstrated an increase in at-risk status at grade 3 (2002-2003, 18% increase; 2003-2004, 14% increase; 2004-2005 cohort from 2% increase), suggesting that beyond grade 2 instructional efforts and intervention efforts did not keep pace with grade level expectations for oral reading fluency. However, it should be noted that these increases in at-risk status were smaller for younger cohorts, suggesting a narrowing in the gap between expectations and intervention outcomes in Grade 3 with each passing year.

At-Risk (not proficient) status based on MAP scores. Percentages of students earning scores in the not proficient range, based on the 4 domains of the MAP were reported for each cohort for fall testing in Grade 3 and for the 2003-2004 and 2004-2005 cohorts for spring testing in Grade 3. Significant changes in proficiency status were noted between the fall and spring MAP results for the 2003-2004 and for the 2004-2005 cohorts. The 2003-2004 cohort reported a 32% reduction in at-risk status for the Vocabulary domain, 72% reduction for the Comprehension domain, 27% reduction for the Critical Content domain and 30% reduction for the Analyze domain, suggesting that general education classroom instruction and remedial intervention efforts were effective in improving the reading skills of the students of this cohort between the fall and spring of Grade 3. The 2004-2005 cohort followed a similar pattern with a

12% reduction in at-risk status for the Vocabulary domain, 10% for the Comprehension domain, 4% for the Critical Content domain, and 5% for the Analyze domain. Although the reductions in at-risk status were smaller in magnitude for the 2004-2005 cohort than for the 2003-2004 cohort, it is important to note that the 2004-2005 cohort demonstrated much higher levels of proficiency on the fall MAP testing than did the 2003-2004 cohort.

Also noteworthy was the reduction in at-risk status between cohort years, suggesting an improvement in instructional efforts for each subsequent cohort. For all the MAP domains, the highest percentage of students scoring in the not proficient range in the fall of Grade 3 was reported for the 2002-2003 cohort, with each subsequent cohort reporting smaller percentages of students scoring in the not proficient range. The changes in the percentages of students scoring in the not proficient range in the fall between the 2002-2003 cohort and the 2003-2004 cohort were small (a 10% reduction for the Vocabulary domain; a 3% reduction for the Comprehension domain; Analyze 3% reduction). The reductions in the not proficient status between the fall 2003-2004 cohort and the fall 2004-2005 cohort were much greater, suggesting a greater impact of intervention efforts for the youngest cohort. The number of students earning not proficient ratings on the fall administration of the MAP was reduced from 75% for the 2003-2004 cohort to 28% for the 2004-2005 cohort. MAP Comprehension domain not proficient ratings were reduced from 72% to 29%, Critical Content domain not proficient ratings were reduced from 65% to 19%, and Analyze domain not proficient ratings were reduced from 72% to 22%. The decrease in at-risk status between subsequent cohorts suggests that as years progressed, adjustments to instructional efforts reduced the numbers of students who were at-risk. The trend of at-risk reduction was also observed in the spring administration of the MAP across all 4 domains. The lowest at-risk rates were reported by the 2004-2005 cohort in the spring of Grade

3, reflecting the culmination of a consistent trend of subsequent cohorts receiving increasingly greater benefits from remedial reading instruction.

Intervention Efficiency based on DIBELS Oral Reading Fluency Measures. For each of the three cohorts, more than 50% of the students in the cohort identified as at-risk in kindergarten, based on the DIBELS K summary score, were identified as not at-risk with the DIBELS oral reading fluency measure in Grade 2, suggesting a substantial degree of effectiveness resulting from interventions offered between the fall of kindergarten and winter of grade 2. Intervention efforts from the winter of Grade 2 to the winter of Grade 3 produced a reversal of some of the gains in improved oral reading fluency that were made from K to Grade 2, as Intervention Efficiency based on the DIBELS oral reading fluency score dropped from Grade 2 to Grade 3 for all cohorts (55% to 27% for 2002-2003; 61% to 42% for 2003-2004; 55% to 44% for 2004-2005).

It is important to note, however, that the size of the drop in intervention efficiency was consistently smaller in each successive cohort (28% for 2002-2003; 19% for 2003-2004; 9% for 2004-2005), suggesting that adjustments to intervention efforts between grades 2 and 3 that were made for the 2003-2004 and 2004-2005 cohorts were having positive results, but that the intervention gains being made in these cohorts were not keeping pace with grade level expectations as much as they had been up to grade 2. The result was a drop in the number of students judged proficient in oral reading fluency in the winter of Grade 3.

Instructional Stability based on DIBELS Oral Reading Fluency Measures. For each of the three cohorts, instructional stability in Grade 2 was more than 75%, meaning a large majority of each cohort that was identified as not at-risk in kindergarten with the DIBELS K summary score earned scores in the not at-risk range on the DIBELS oral reading fluency measure in

Grade 2. These results suggest that instruction offered in the general education classrooms between the fall of kindergarten and winter of grade 2 enabled many students to maintain expected levels of proficiency in oral reading fluency.

General classroom instruction efforts from the winter of Grade 2 to the winter of Grade 3 produced a slight decline for all three cohorts in the number of students that remained not at-risk, based on the DIBELS oral reading fluency score (78% to 72% for 2002-2003; 83% to 76% for 2003-2004; 89% to 83% for 2004-2005). Once again, it is important to note that even with the slight decline in instructional stability, a large majority of each cohort continued to demonstrate proficient oral reading fluency.

As was the case with intervention efficiency, instructional stability increased with each successive cohort in Grade 2 and in Grade 3. In grade 2 increases were as follows: 78% for 2002-2003 to 83% for 2003-2004 to 89% for 2004-2005. In grade 3 increases were as follows: 72% for 2002-2003 to 76% for 2003-2004 to 83% for 2004-2005. These increases indicated that adjustments in the general education curriculum to improve instructional stability in Grade 2 and in Grade 3 were being effective to some degree.

Sensitivity based on DIBELS Oral Reading Fluency Measures. Sensitivity values for the three cohorts were moderately high in Grade 2 (71% for both 2002-2003 and 2003-2004 and 81% for 2004-2005) and remained in the moderately high range for all three cohorts in Grade 3 (76% for 2002-2003; 72% for 2003-2004; 77% for 2004-2005). These moderately high sensitivity values indicate that of the students performing in the at-risk range on oral fluency measures in Grades 2 and 3, a majority of those students also had been identified as at-risk in Kindergarten with the DIBELS Summary score.

Sensitivity of the DIBELS Kindergarten Summary score for predicting at-risk status, based on the DIBELS oral reading fluency score as the criterion measure, increased by 5% from Grade 2 to Grade 3 for the 2002-2003 cohort, reflecting a drop back into the at-risk range for some students from Grade 2 to Grade 3. Sensitivity of the DIBELS Kindergarten Summary score as a predictor of DIBELS oral reading fluency was essentially the same in Grades 2 and 3 for the 2003-2004 cohort, 71% and 72% respectively. In the case of the 2004-2005 cohort, sensitivity dropped from 81% to 77%. This drop in sensitivity reflects the increased closing of the gaps between Grade 2 and Grade 3 both in intervention efficiency and in instructional stability for this youngest of the cohorts.

Specificity based on DIBELS Oral Reading Fluency measures. Specificity values of the DIBELS Kindergarten Summary score based on the DIBELS Grade 2 oral reading fluency score were in the moderate range for all three cohorts (54% for 2002-2003; 56% for 2003-2004; 63% for 2004-2005). Specificity values increased from grade 2 to grade 3 for all three cohorts, but the increase was consistently smaller with each successive cohort (14% increase for the 2002-2003 cohort; 7% increase for the 2003-2004 cohort; 2% increase for the 2004-2005 cohort). The increases in specificity values reflected the drops in intervention efficiency from Grade 2 to Grade 3 realized by each cohort as the number of students identified because not at-risk in Kindergarten made up a larger proportion of the total number of students identified as proficient in Grade 3 compared with Grade 2. The near stable specificity percentage of the 2004-2005 cohort, however, reflects the improvements in intervention efficiency affected for this youngest cohort.

Kappa values based on DIBELS Oral Reading Fluency measures. Similar to specificity values, the percent of increase over chance when using the DIBELS Kindergarten Summary

score to predict at-risk oral reading fluency in later grades increased from Grade 2 to Grade 3 for all three cohorts (Kappa increase from 22% to 45% for 2002-2003; from 21% to 34% for 2003-2004; from 36% to 39% for 2004-2005). Kappa values in Grade 2 and in Grade 3 decreased with each successive cohort with the exception of a slight rise in kappa for the 2004-2005 cohort in Grade 3. These continued improvements in kappa values from Grade 2 to Grade 3 reflect in a general way the positive trend of results obtained with the Intervention Efficiency, Instructional Stability, Sensitivity, and Specificity data. Ideally, when intervention efforts are highly productive, kappa values should decrease from one year to the next, reflecting the fact, that in each successive year, a greater percentage of students thought to be at-risk in kindergarten are performing in the proficient or not at-risk range.

Intervention Efficiency based on MAP domain scores. Intervention efficiency based on MAP domain scores was relatively low in fall for the 2003-2004 cohort, but much higher for the 2004-2005 cohort. This reflects improvement in intervention efforts from 2003-2004 to 2004-2005. For 2003-2004 and 2004-2005 cohorts, intervention efficiency increased greatly from the fall to spring, with both cohorts showing similar percentages of increase across all MAP domains. The number of students earning scores in the proficient range in the MAP Vocabulary domain increased from 13% to 42% for 2003-2004 cohort and from 44% to 67% for 2004-2005 cohort. The percentage of students earning scores in the proficient range in the MAP Comprehension domain increased from 13% to 35% for the 2003-2004 cohort and from 53% to 74% for the 2004-2005 cohort, and the percentage of proficient range scores in the MAP Critical Content domain increased from 19% to 52% for the 2003-2004 cohort and from 69% to 78% for the 2004-2005 cohort. Finally, proficient scores in the Analyze domain reflected an increase from 13% to 42% for the 2003-2004 cohort and from 44% to 67% for the 2004-2005 cohort.

Instructional Stability based on MAP domain scores. Instructional stability values were relatively low for the 2003-2004 cohort in the fall, (38% for Vocabulary; 43% for Comprehension; 52% for Critical Content and 45% for Analyze), but large increases were realized by the spring administration of the MAP, (72% for Vocabulary; 76% for Comprehension; 72% for Critical Content and 76% for Analyze), reflecting effective outcomes from general education instruction efforts from fall to spring for this cohort across all MAP domains. For the 2004-2005 cohort, stability was at a very high level initially, (81% for Vocabulary; 89% for Comprehension; 92% for Critical Content; 92% for Analyze), and was maintained at that high level in the spring, (81% for Vocabulary; 89% for Comprehension; 92% for Critical Content; and 89% for Analyze), reflecting continued effective general instruction efforts across all MAP Domains.

Sensitivity based on MAP domain scores. For 2003-2004 cohort, sensitivity increased from fall to spring across all MAP domains (60% to 69% for Vocabulary; 63% to 74% for Comprehension; 64% to 65% for Critical Content and 63% to 72% for Analyze). The 2004-2005 cohort reported a decrease in sensitivity across all domains (74% to 63% for Vocabulary; 81% to 71% for Comprehension; 79% to 73% for Critical Content and 73% to 67% for Analyze). The difference in the pattern of sensitivity values directly reflects the pattern of changes in instructional efficiency and instructional stability for the two cohorts. For the 2003-2004 cohort, although both intervention efficiency and instructional stability increased from fall to spring, greater gains were realized for instructional stability than for intervention efficiency, thereby increasing the ratio of the number of students at-risk in kindergarten that remained at-risk (not proficient on MAP) in Grade 3, relative to the number of students identified as not-at risk in kindergarten that were not at-risk (not proficient on MAP) in Grade 3. For the 2004-2005

cohort, intervention efficiency increased from fall to spring but instructional stability, already at a very high level in the fall, remained the same. As a result, the ratio of the number of students at-risk in kindergarten that remained at-risk (not proficient on MAP) in Grade 3, relative to the number of students identified as not-at risk in kindergarten that were not at-risk (proficient on MAP) in Grade 3 decreased, because a number of students at-risk in kindergarten were classified as proficient in Grade 3; the result was a lowered sensitivity value. The pattern demonstrated here by the 2004-2005 cohort reflects the ideal situation wherein instructional stability starts out very high because of effective general education instruction, and intervention efficiency increases greatly because of good early intervention efforts. The better the instruction and intervention, the greater the decrease in sensitivity from kindergarten to later grades.

Specificity based on MAP domain scores. For both cohorts, specificity values decreased from fall to spring. The 2003-2004 cohort reported specificity decreases from 73% to 62% for Vocabulary; from 76% to 67% for Comprehension; from 71% to 57% for Critical Content and from 76% to 63% for Analyze. Specificity decreases for the 2004-2005 cohort were from 62% to 55% for Vocabulary, from 63% to 53% for Comprehension, from 57% to 54% for Critical Content and from 54% to 53% for Analyze. As was the case with sensitivity values, the pattern of decreases in specificity values reflects the pattern of changes in instructional efficiency and instructional stability for the two cohorts and across all MAP domains. For the 2003-2004 cohort, both intervention efficiency and instructional stability increased from fall to spring, thereby altering the ratio of the number of students not at-risk in kindergarten that remained not at-risk (proficient on MAP) in Grade 3 relative to the number of students identified as at-risk in kindergarten that were not at-risk (proficient on MAP) in Grade 3. For the 2004-2005 cohort, intervention efficiency increased from fall to spring but instructional stability, already at a very

high level in the fall, remained the same. As a result, the ratio of the number of students not at-risk in kindergarten that remained not at-risk (proficient on MAP) in Grade 3 relative to the number of students identified as at-risk in kindergarten that were not at-risk (proficient on MAP) in Grade 3 decreased, resulting in a lowered specificity value.

Once again, the pattern demonstrated here by the 2004-2005 cohort reflects the ideal situation wherein instructional stability starts out very high because of effective general education instruction and intervention efficiency increases greatly because of good early intervention efforts. The better the instruction and intervention, the greater the decrease in specificity from kindergarten to later grades.

Kappa based on MAP domain scores. Kappa values tended to increase from fall to spring for the 2003-2004 cohort across all MAP domains. Vocabulary increased from 25% to 39%; Comprehension from 32% to 40%; Critical Content from 33% to 21% and Analyze from 32% to 34%. Kappa values decreased from fall to spring for the 2004-2005 cohort for all MAP domains. Vocabulary decreased from 36% to 14%; Comprehension from 36% to 15%; Critical Content from 22% to 14% and Analyze from 28% to 11%. Because kappa represents the percentage of accuracy greater than chance assignment to categories, this value is most sensitive to the overall consistency of at-risk and not at-risk classifications. Therefore the greater the number of students who change their status from kindergarten to a later grade, the greater the decrease in kappa. As reflected in the data for the 2004-2005 cohort, the higher intervention efficiency and instructional stability values, and the lower sensitivity and specificity values resulted in relatively large decreases in kappa values from fall to spring. In the ideal situation, in a manner similar to sensitivity and specificity, kappa values would decrease from kindergarten to later grades, reflecting high levels of intervention efficiency and instructional stability.

Research Question 4: Socioeconomic Status

The fourth research question was designed to examine the effects of SES on prediction of at-risk readers, as well as on program effectiveness. Each research question was re-examined with SES.

*Research Question 4a**Kindergarten Screening Measure and Fall DIBELS Assessment with Socioeconomic Status*

The effect of SES on the relationship between the Kindergarten Screening measure and Fall DIBELS measures was investigated. Analyses were completed for individual cohorts and for all cohorts combined. For the Low SES groups, the comparison of the Kindergarten Screening measure with the DIBELS Summary Score produced the highest kappa values for all groups combined as well as for two of the three cohorts. It is likely that the disadvantaged group was identified with greater accuracy because students with low SES backgrounds often present with multiple risk factors, increasing the likelihood that a screening instrument would detect concerns (Snow et al., 1998). For the High SES group, high kappa values were also obtained for the comparison of the Kindergarten Screening measure with the DIBELS LNF. The relationship between DIBELS LNF and reading achievement was established and increased the likelihood over chance for identification of at-risk status (Good & Kaminski, 2002). Overall, the ISF reported the highest levels of sensitivity for two of the three High SES groups and for two of the three Low SES groups. Research has shown a strong link between phonemic awareness and phoneme segmentation (Blachman, 1984; Share et al., 1984, Stahl & Murray, 1994) and later reading achievement, suggesting that ISF identifies students with weak phonemic awareness and the potential for reading difficulties. Additionally, the Kindergarten Screening score with the

DIBELS LNF score produced the highest specificity values for two of the three High SES cohorts, as well as for the combined High SES group. The automaticity associated with the DIBELS LNF task has been proven to correlate with reading achievement (Good & Kaminski, 2002). The specificity was stronger for the High SES group who were likely to enter Kindergarten with stronger readiness skills than for students from Low SES backgrounds (Rafoth, 1997). Overall consistency was high, and relatively comparable, for the High SES group and for the Low SES group. The high level of overall consistency for both groups demonstrated that little change took place between the early Kindergarten Screening and the DIBELS measures during the fall of Kindergarten.

Research Question 4b

Kindergarten Screening Decision and Grades 1 and 2 Winter Oral Reading Fluency and Socioeconomic Status: Prediction of At-Risk Readers

The effect of SES on the relationship between the Kindergarten Screening decision and Fall DIBELS Scores with Grade 1 and Grade 2 Winter ORF was investigated. Results of Grade 1 Winter ORF comparisons are discussed, followed by results of Grade 2 Winter ORF comparisons. The DIBELS Summary score consistently demonstrated the strongest sensitivity for the disadvantaged group across all cohorts when looking at Grade 1 Winter ORF. For both cohorts, the highest kappa values were reported for the Low SES group when the Kindergarten Screening measure was compared with DIBELS LNF. The 2002-2003 cohort reported a 66% kappa value; the 2003-2004 cohort reported a 71% kappa value, and the 2004-2005 cohort reported a 67% kappa value. The high kappa levels for the Low SES group was attributed to the many risk factors that students from disadvantaged backgrounds often present (Snow et al., 2002); this made the at-risk status easier to identify. Two of the three cohorts reported 100%

sensitivity for the Low SES groups in the comparison of the DIBELS Summary score and Grade 1 Winter ORF. The high level of sensitivity for the Low SES group indicated that that SES status improved the detection of at-risk status. For two of the three cohorts, and for both the Low SES and High SES groups, the DIBELS Summary score consistently reported stronger sensitivity than the Kindergarten Screening measure. Research has stated that DIBELS measures correlate with reading success, in contrast to the Kindergarten Screening measure, which is not empirically supported. The DIBELS ISF and DIBELS LNF subtests also reported consistently stronger sensitivity for the Low SES groups than for the High SES groups. The Low SES group consistently reported higher levels of sensitivity than the High SES group, providing evidence that the inclusion of SES is valuable in the detection of at-risk status. The highest level of specificity was characterized by the comparison of the Kindergarten Screening measure and the DIBELS LNF for Low SES group. Specificity for the DIBELS LNF Low SES group was consistently higher than the specificity for the High SES group. Therefore, SES increased the detection of students who were not at-risk and lent support to the importance of examination of SES when investigating at-risk status.

Overall consistency for the Low SES group was highest for two of the three groups when the Kindergarten Screening measure was compared with the DIBELS LNF. For the High SES group Overall Consistency was highest for all three groups when the Kindergarten Screening measure was compared with the DIBELS LNF. The strong Overall Consistency that was characterized by the comparison of the Kindergarten Screening measure with the DIBELS LNF suggested that, for both Low SES and High SES students, that the skill of LNF remained the most stable over time.

The effect of SES on the relationship between the Kindergarten Screening measure and DIBELS with Grade 2 Winter ORF was investigated. Two of the three cohorts reported the highest kappa values for the Low SES groups when the DIBELS LNF was compared with DIBELS Grade 2 Winter ORF. The research-based nature of the DIBELS LNF and the reality that the Low SES group likely presented with more risk factors accounted for the strong kappa levels. The Kindergarten Screening Decision reported 100% sensitivity for the Low SES group for two of three cohorts. The Kindergarten Screening measure demonstrated stronger sensitivity for the High SES and Low SES groups than did the DIBELS Summary score, or the DIBELS ISF or DIBELS LNF. Although greater sensitivity was reported by the Kindergarten Screening measure, it is important to take a close and comprehensive look at the fundamentals of the different screening measures. As previously discussed, the DIBELS Summary score and the DIBELS ISF and DIBELS LNF are linked to reading research, but the Kindergarten Screening measure is not. For both the Low SES and the High SES groups the comparison between the DIBELS LNF and Grade 2 Winter ORF demonstrated the strongest specificity for all cohorts. This finding lends support to the predictive validity of the DIBELS LNF in the determination of reading status. High levels of Overall Consistency were reported both for the Low SES and for High SES cohorts across variables. The Low SES groups consistently reported higher Overall Consistency than the High SES group. The high level of agreement between the Kindergarten Screening measures and DIBELS Grade 2 Winter ORF occurred because students from disadvantaged backgrounds likely presented with more severe at-risk status and were more easily identified on screening instruments than were students who presented with more mild concerns.

*Research Question 4c**Effectiveness of Reading Instruction and Socioeconomic Status*

At-Risk status by SES group based on DIBELS measures. Percentages of High SES and Low SES students classified as at-risk, based on DIBELS scores, were reported for each cohort during kindergarten and grades 1, 2, and 3. In all three cohorts, less than 50% of the students in the High SES group were classified as at-risk based on DIBELS Sum scores obtained in the fall of Kindergarten (2002-2003 46%; 2003-2004, 50%; 2004-2005 40%). As the High SES students of all cohorts progressed through grades 1 and 2, at-risk status decreased, suggesting that modifications to general classroom instruction and specialized intervention efforts implemented for successive cohorts had a positive effect. A slight upturn in at-risk status of High SES students was noted in all three cohorts between the grade 2 and grade 3 years (2002-2003, 17% increase; 2003-2004, 18% increase; 2004-2005, 2% increase), suggesting that instructional efforts produced a decrease in at-risk status between kindergarten and grades 1 and 2; however, the positive effects on instruction and interventions did not enable some students in the High SES group to keep pace with grade 3 expectations in oral reading fluency. For the 2002-2003 cohort, 39% of High SES students earned oral reading fluency scores in the at-risk range. A decrease in at-risk status in Grade 1 was noted for the High SES students of subsequent cohorts, however (2003-2004, 28%; 2004-2005, 31%), suggesting that subsequent modifications to general classroom instruction and specialized intervention efforts employed in Grade 1 benefited the High SES students of the 2003-2004 and 2004-2005 cohorts. The trend of lower oral reading fluency at-risk percentages for the High SES students of subsequent cohorts was also demonstrated in grade 2 between the 2002-2003 cohort (29% at-risk) and the 2003-2004 cohort (20% at-risk), suggesting that the 2003-2004 cohort benefited from modification of instructional

efforts aimed at reducing at-risk status. Increases in at-risk percentages for High SES students were noted between the 2003-2004 and the 2004-2005 cohorts for grades 1 and 2 oral reading fluency (from 28 to 31% in Grade 1; from 20 to 29% in Grade 2). The 2004-2005 at-risk percentages, however, did not surpass the 2002-2003 at-risk percentages, suggesting that effects of intervention efforts were maintained by the High SES students in the 2004-2005 cohort. Analysis of grade 3 data showed that at-risk status decreased for the High SES students of each subsequent cohort (2002-2003, 46%; 2003-2004, 38%; 2004-2005, 31%), suggesting that efforts to improve general classroom instruction and specialized interventions benefited the High SES students of each subsequent younger cohort.

In all three cohorts, more than 50% of the Low SES students were classified as at-risk based on DIBELS Sum scores obtained in the fall of Kindergarten (2002-2003, 75%; 2003-2004, 55%; 2004-2005, 67%). As each cohort progressed through grades 1 and 2, decreases were noted in the number of Low SES students classified as at-risk based on DIBELS oral reading fluency measures. These consistent decreases suggest that instructional approaches and specialized intervention efforts were effective in reducing the at-risk status of Low SES students in Grades 1 and 2. A slight upturn in at-risk percentages in Grade 3 were noted, however, for the Low SES students of all three cohorts (2002-2003, 17% increase; 2003-2004, 15% increase; 2004-2005, 14% increase), suggesting that beyond grade 2, classroom instructional efforts and specialized intervention efforts did not keep pace with grade level expectations for oral reading fluency.

The overall trend between the oldest cohort and those subsequently entering the same grade level reflected a decrease in at-risk status of Low SES students across cohorts at the same grade level. The trend in kindergarten was a reduction in at-risk status of Low SES students

between the 2002-2003 and 2003-2004 cohorts (from 75% down to 55%), suggesting that the younger cohort benefited from alterations of instructional methods that were used with the oldest cohort. Contrary to the general trend, an increase of 12% in at-risk status for Low SES students was noted between the 2003-2004 and 2004-2005 cohorts at the kindergarten level, a minor increase of 2% between the 2002-2003 and 2003-2004 cohorts at the Grade 1 level. A 27% decrease in at-risk status, however, was reported between the 2003-2004 and 2004-2005 cohorts at the Grade 1 level, suggesting that the 2004-2005 cohort benefited from modifications to instructional and intervention methods used with the previous cohorts in Grade 1.

Additionally, reductions in at-risk status of Low SES students from older to younger cohorts were consistently noted in grades 2 and 3, suggesting that later cohorts continued to benefit from the modifications made to instructional strategies and specialized intervention efforts aimed at improving oral reading fluency skills.

Across all the cohorts, similar trends in the pattern of reduction of at-risk status were recognized between the High SES and Low SES student groups. Greater percentages of at-risk status students were reported consistently for the Low SES student group than for the High SES group. This finding is not surprising because literature has indicated that students from disadvantaged backgrounds enter school lacking literacy skills (Snow et. al, 1998). The Low SES students, however, demonstrated greater overall reductions of at-risk status between cohort years of the same grade levels than did their High SES counterparts, suggesting that instructional approaches and remedial intervention efforts reduced the at-risk status for more Low SES students than for more High SES students. It is important to note that the percentages of at-risk Low SES students in the 2004-2005 cohort began much higher than the percentage of High SES students but were reduced to the point at which they matched those of the High SES students in

grades 1 and 2 before showing a slight rise above the level of the High SES students in grade 3. After careful monitoring of progress across three cohorts of students, the school's instructional staff had managed to eradicate the effects of Low SES typically seen in elementary school populations.

Comparison of DIBELS Kindergarten Summary Scores with Grades 2 and 3 DIBELS Oral Reading Fluency Scores by SES Group

Intervention Efficiency. In the two oldest cohorts, Intervention Efficiency was better than 50% (62% for the 02-03 Cohort; 75% for the 03-04 cohort) for High SES students identified as at-risk in kindergarten, based on the DIBELS Kindergarten summary score and identified as not at-risk with the DIBELS ORF measure in Grade 2, suggesting a considerable degree of effectiveness that resulted from interventions offered between the fall of kindergarten and the winter of Grade 2. For the youngest cohort, intervention efficiency with High SES students dropped below 50% in Grade 2 (44% for the 04-05 cohort). Intervention efficiency percentages in Grade 2 were much lower for the disadvantaged (Low SES) students in the two oldest cohorts (44% for the 02-03 cohort and 36% for the 03-04 cohort), but much higher for the Low SES students in the youngest cohort (61% for the 04-05 cohort). Although the Low SES students of only one cohort demonstrated more than 50% intervention efficiency from instructional efforts between K and 2nd grade, it is important to recognize that this cohort was the youngest of the three, i.e., the one entering kindergarten most recently, suggesting that changes in the interventions in successive years helped to improve the outcomes for these newest disadvantaged students. For both the High SES and Low SES students, intervention efforts from the winter of Grade 2 to the winter of Grade 3 failed to show the kinds of improvements obtained between the fall of kindergarten and the winter of Grade 2. Both the Low SES and the High SES

groups demonstrated lower intervention efficiency percentages in Grade 3, for two of three cohorts. There were exceptions to this pattern; these included the High SES students of the 04-05 cohort, whose percentage remained stable at the already lower level of 44%, and the disadvantaged students in the 03-04 cohort, whose percentage remained stable at the already low level of 36%. Although intervention efficiency decreased or remained at a lower level from Grade 2 to Grade 3 for both High SES and Low SES student groups, it is important to recognize that the youngest cohort (04-05) demonstrated an upturn in intervention efficiency percentages in both Grade 2 and Grade 3 for the disadvantaged students, suggesting a trend toward greater success with these disadvantaged students than with their High SES counterparts.

Instructional Stability. For the High SES students in all three cohorts, and for Low SES students of two of the three cohorts, instructional stability in Grade 2 was greater than 75%, reflecting the fact that a large majority of each cohort identified as not at-risk in kindergarten with the DIBELS, earned scores on the DIBELS ORF in the not-at risk range in Grade 2. These results suggest that general education instruction between the fall of kindergarten and the winter of Grade 2 allowed large proportions both of High SES and of Low SES students to maintain proficiency in oral reading fluency. Across the three cohorts, improvements in instructional stability rates in Grade 2 were evident for both the High SES and the Low SES student groups, suggesting that subsequent cohorts benefited from modifications to general education instructional practices made in each successive year. A slight loss of instructional stability was noted between grade 2 and grade 3 for both the High SES and Low SES student groups across all three cohorts, but the size of the losses decreased with each successive cohort, suggesting that modifications made to instructional practices in Grade 3 in each subsequent year were helping to improve stability across Grades 2 and 3.

Sensitivity. Grade 2 Sensitivity values for the High SES student groups were moderately high (63% for 2002-2003, 63% for 2003-2004 and 77% for 2004-2005), indicating that a majority of the students that were identified as at risk based on DIBELS oral reading fluency scores in Grade 2, were also identified as at-risk in Kindergarten. Grade 2 and Grade 3 sensitivity values for the Low SES student groups were somewhat higher (83% and 88% for the 2002-2003 cohort, 78% and 70% for the 2003-2004 cohort, and 88% and 83% for the 2004-2005 cohort), indicating that a greater number of students in the Low SES group that were identified as at risk based on DIBELS oral reading fluency scores in Grade 2 were also identified as at-risk in Kindergarten.

Specificity. Specificity values comparing the DIBELS Kindergarten Summary Scores to the DIBELS Grade 2 ORF scores were in the moderate range for both the High SES and Low SES groups. Specificity rates increased between grade 2 and grade 3 for the 2002-2003, and 2003-2004 High SES groups and remained relatively consistent for the 2004-2005 High SES group. For the Low SES group, increases were noted for the 2002-2003 and the 2003-2004 groups, but the 2004-2005 group remained relatively consistent. With each successive cohort, and for both SES groups, increases in specificity between grade 2 and grade 3 were consistently smaller for each cohort. The increases in specificity values reflected the drops in intervention efficiency and in instructional stability from grade 2 to grade 3; this was demonstrated by each cohort and SES group.

Kappa. The percent increase over chance (Kappa) obtained when using the DIBELS Kindergarten Summary score to predict at-risk status based on the DIBELS ORF increased for all cohorts between grade 2 and grade 3 regardless of SES. However, for the High SES groups,

kappa values decreased between the 2002-2003 and the 2003-2004 cohort in grade 2 and grade 3, and kappa values decreased for the Low SES groups between the 2003-2004 and 2004-2005 years for grade 2 and grade 3. The consistent decrease in kappa values between grade 2 and grade 3, for these specific groups, illustrated that a greater percentage of students thought to be at-risk in kindergarten were performing in the not at-risk range in grade 3, and suggested that modifications to instruction and intervention efforts were effective.

Comparison of DIBELS Kindergarten Summary Scores with Grade 3 Fall and Spring MAP Scores

MAP Performance by SES Level. Percentages of students earning scores in the not proficient range, based on the 4 domains of the MAP, were reported for the High SES and Low SES student groups for fall testing in Grade 3 for all three cohorts and for spring testing in grade 3 for the 2003-2004 and 2004-2005 cohorts.

The High SES students of both the 2003-2004 and the 2004-2005 cohorts showed a decrease in at-risk (i.e., not proficient) percentages between the fall and spring for all four MAP domains, suggesting that general education classroom instruction and remedial intervention efforts were effective in improving reading skills. The High SES students of the 2003-2004 cohort reported a 32% reduction in the not proficient status between the fall and spring for the Vocabulary domain, 22% reduction for the Comprehension domain, 20% reduction for the Critical Content domain, and an 18% reduction for the Analyze domain. The High SES students of the 2004-2005 cohort followed a similar pattern with a reported 13% reduction in not proficient status between the fall and spring for the Vocabulary domain, a 9% reduction for the Comprehension domain, a 9% reduction for the Critical Content domain and a 7% reduction for the Analyze domain. Although the reductions in not proficient status were smaller in magnitude

for the 2004-2005 cohort than for the 2003-2004 cohorts, it is necessary to recognize that at the outset, the High SES students of the 2004-2005 cohort demonstrated higher levels of proficiency on the fall MAP testing than did the 2003-2004 cohort, making the gains realized by the 2004-2005 cohort that much more impressive.

The reduction in at-risk status of High SES students across cohort years suggest that improvements in instructional efforts benefited the cohort that followed. The High SES students of the 2002-2003 cohort (the oldest cohort, reflected the greatest level of at-risk status in the fall for all four MAP domains (Vocabulary 82%; Comprehension 68%, Analyze 68%), but subsequent cohorts reflected reductions in the percentages of students scoring in the not proficient range . Across the 2002-2003 and 2003-2004 cohorts, Low SES students showed a 12% decrease in at-risk status for the Vocabulary domain; a 3% decrease was noted for the Comprehension domain, and a 5% decrease for the Analyze domain. Larger reductions in percentages of students earning scores in the not proficient range occurred for the High SES students across the 2003-2004 and 2004-2005 cohorts, with a reported 37% reduction for the Vocabulary domain, a 36% reduction for the Comprehension domain, a 31% reduction for the Critical Content domain, and a 44% reduction for the Analyze domain, suggesting that modifications to instructional efforts and specialized intervention across cohort years were effective in improving the reading skills of High SES students who performed poorly on the MAP in the fall of grade 3.

Reduction in percentage of High SES students scoring in the not proficient range on the spring administration of the MAP also was observed across the 2003-2004 and 2004-2005 cohorts, with a reported 18% reduction in at-risk status for the Vocabulary domain, a 23% reduction for the Comprehension domain, a 22% reduction for the Critical Content domain and

32% reduction for the Analyze domain. The lowest percentages of High SES students scoring in the not proficient range was reported by the 2004-2005 cohort in the spring of grade 3 reflecting the compounding effect of improvements in reading instruction made with each subsequent cohort.

The Low SES students of the 2003-2004 and 2004-2005 cohorts showed a decrease in not proficient status between the fall and spring across all four MAP domains, except for the 2004-2005 Critical Content score, suggesting that general education instruction and specialized intervention methods improved the reading skills of the Low SES students of both cohorts between the fall and spring of grade 3. The 2003-2004 cohort showed a 30% reduction in at-risk status for the Vocabulary domain, a 35% reduction for the Comprehension domain, a 40% reduction for the Critical Content domain and a 55% reduction for the Analyze domain. Low SES students in the 2004-2005 cohort reported a 7% increase of students earning scores in the not proficient range between the fall and spring on the Critical Content domain, suggesting that reading instruction efforts did not provide the level of support necessary to keep pace with grade 3 expectations for the Critical Content domain. The Low SES students of the 2004-2005 cohort experienced a 4% reduction of not proficient range scores for the Vocabulary domain, a 12% reduction for the Comprehension domain, and 4% reduction for the Analyze domain. Although the reductions in not proficient range scores were not as great for the 2004-2005 cohort as for the 2003-2004 cohort, it is important to note that the Low SES students of the 2004-2005 cohort demonstrated much higher levels of proficiency for all fall MAP domains than did the 2003-2004 cohort.

Reductions in percentages of Low SES students earning scores in the not proficient range were evident across cohort years, suggesting that improvements in instructional efforts with the

subsequent cohort were effective. The Low SES students of the oldest cohort (2002-2003) reflected the greatest percentage of not proficient MAP domain scores in the fall. Reductions in at-risk status were noted between the 2002-2003 and 2003-2004 cohorts, with a reported 7% reduction in the Vocabulary and Comprehension domains and a 2% reduction for the Analyze domain. Additionally, very substantial reductions in percentages of Low SES students earning scores in the not proficient range on all domains of the MAP in the fall were reported between the 2003-2004 and 2004-2005 cohort years (44% reduction in the Vocabulary domain, 55% reduction in the Comprehension domain, 74% reduction in the Critical Content domain, and 64% reduction in the Analyze domain), suggesting that modifications to instructional efforts and specialized intervention across cohort years were effective in improving the reading skills of Low SES students who performed poorly on the MAP in the fall of grade 3.

The trend of reduction of scores in the not proficient range for Low SES students continued into the spring across the 2003-2004 and 2004-2005 cohorts, during which an 18% reduction in at-risk status was reported for the Vocabulary domain, a 32% reduction for the Comprehension domain, a 27% reduction for the Critical Content domain and a 13% reduction for the Analyze domain. Although the magnitude of reductions was greater in the fall, it is important to note that the percentage of Low SES students scoring in the not proficient range in the spring was much lower than in the fall, making it much more difficult to affect large decreases in the spring. The lowest percentages of Low SES students scoring in the not proficient range on MAP domains were reported by the 2004-2005 cohort, reflecting the compounding effect of improvements in reading instruction made with each subsequent cohort.

Similar patterns of not proficient status, and reduction of not proficient status, for MAP results in both fall and spring were observed for the High SES and Low SES groups of all three

cohorts. Comparison of the changes in not proficient status between the fall and spring MAP domains for High SES and Low SES students in each cohort provided information regarding the effects of intervention efforts.

Across all the cohorts, similar trends in the pattern of reduction of students scoring in the not proficient range were recognized between the High SES and Low SES student groups. Greater percentages of not proficient status were consistently reported for the Low SES student group than were reported for the High SES group. The finding noted in the discussion related to oral reading fluency performance is not surprising because literature has indicated that students from disadvantaged backgrounds enter school lacking literacy skills (Snow et. al, 1998).

The Low SES students, however, demonstrated greater overall reductions of not proficient status between cohort years of the same grade levels than did their High SES counterparts, suggesting that instructional approaches and remedial intervention efforts reduced the at-risk status for a greater number of Low SES students than for High SES students. It is important to note that the percentages of not proficient Low SES students in all three cohorts, although somewhat higher, were very similar in magnitude to those of the High SES students for all four MAP domains in both the fall and the spring. The relative similarity in status of the High SES and Low SES groups in grade 3 strongly suggest that after careful monitoring of progress across three cohorts of students, Resica Elementary school's instructional staff had managed by grade 3 to eradicate the effects of Low SES typically seen in elementary school populations.

Intervention Efficiency. For the 2002-2003 and 2003-2004 cohorts, intervention efficiency for the Low SES student groups was extremely low across all MAP domains in the fall of Grade 3 (either 0% or 9%). The intervention efficiency percentages for the High SES student groups of these two cohorts were somewhat higher, ranging from 8% to 25%.

In contrast to the low percentages of the two older cohorts, intervention efficiency percentages for the 2004-2005 cohort were much higher in the fall of Grade 3 for the High SES and for Low SES groups. The High SES group's percentages ranged from 36% in the Critical Content domain to 61% in the Analyze domain. The Low SES group's percentages were even better, ranging from 50% in the Vocabulary domain to 83% in the Critical Content domain. The higher levels of intervention efficiency for both the High SES and Low SES groups of the 2004-2005 cohort suggest that modifications to both general education instructional methods and to specialized intervention approaches had a positive impact on the youngest cohort.

Between the fall and the spring, increases in the number of students earning scores in the proficient range were shown for both the 2003-2004 and 2004-2005 cohorts for the High SES and for the Low SES student groups in each cohort across all 4 MAP domains with the exception of the 2004-2005 Low SES group in which intervention efficiency decreased slightly from 83% to 78%. Fall to spring gains were much larger for the 2004-2005 cohort for both the High SES and the Low SES student groups. The substantial increases between fall and spring administrations of the MAP indicate that intervention efforts engaged between the fall and spring of Grade 3 were effective for both cohorts, with increased benefits derived from modifications to instructional methods in each successive year. The steady increases in intervention efficiency from the fall to spring of 2003-2004 and fall to spring of 2004-2005 suggest that efforts to improve remedial instruction were producing continuous incremental gains.

Instructional Stability. Instructional stability was relatively low for the High SES and for the Low SES groups of the 2002-2003 cohort across all MAP domains in the fall of Grade 3 (percentages ranging from 27% for Vocabulary to 40% for Comprehension and Analyze); approximately 60-70% of High SES and Low SES students identified as not at-risk in

Kindergarten earned scores in the not proficient range for the MAP domains. Instructional stability percentages in the fall of Grade 2 were somewhat better for the 2003-2004 High SES group across all four MAP domains, ranging from 40% for Vocabulary to 65% for Critical Content, but the percentages for the 2003-2004 Low SES group were at a lower level, similar to that of the 2004-2005 cohort, ranging from 22% for Comprehension, Critical Content and Analyze to 33% for Vocabulary.

In contrast, the High SES and Low SES groups of the 2004-2005 cohort demonstrated much higher instructional stability percentages in the fall of grade 2 across all four MAP domains, with percentages ranging from 80% in Vocabulary to 93% in Analyze for the High SES group and 78% in Vocabulary to 100% in Comprehension, Critical content and Analyze for the Low SES group. Significant improvements in instructional stability were noted between the fall and the spring for the High SES group of the 2003-2004 cohort, who produced percentages at or above 75% for each MAP domain (Vocabulary 80%, Comprehension 80%, Critical Content 75% and Analyze 75%). Gains of a similar magnitude were realized as well for the Low SES group of the 2003-2004 cohort, who produced percentages at or above 50% for each MAP domain (56% for Vocabulary, 67% for Comprehension, 67% for Critical Content and 78% for Analyze). The substantial increases in instructional stability suggests that the general education instructional efforts between the fall and spring of grade 3 had a positive impact for both the High SES and Low SES groups of the 2003-2004 cohort.

For the High SES and Low SES groups of the 2004-2005 cohort, instructional stability began above 75% in all MAP domains in the fall of Grade 3. For the High SES group, instructional stability percentages increased slightly for 3 of the four MAP domains (Vocabulary 5% increase, Comprehension 4% increase, Critical Content 6% increase) and decreased slightly

for the remaining domain (4% decrease), suggesting that instructional practices continued to strengthen between the fall and spring of Grade 3. For the Low SES group, instructional stability decreased somewhat for three of the 4 domains (Vocabulary 11% decrease, Comprehension and Critical Content 11% decrease) and remained at the same percentage (89%) for the Analyze domain, reflecting a slight decrease in effectiveness of instruction in the general education setting.

Instructional stability was moderately high, above 75%, for the Low SES group of the 2004-2005 cohort in the fall of grade 2, reflecting a substantial improvement in comparison with the Low SES student groups of the 2002-2003 and 2003-2004 cohorts, suggesting that instructional modifications made in successive cohorts benefited the youngest cohort. Although a decrease in instructional stability was reported between the fall and spring for the Low SES students of the 2004-2005 cohort, spring instructional stability levels still marked an improvement from the spring 2003-2004, suggesting that adjustments in instructional approaches in the general education that were made in the 2003-2004 year were beneficial to the subsequent cohort.

Sensitivity. Sensitivity percentages were in the moderate range for the High SES groups of all three cohorts for all four MAP domains in the fall of Grade 3, with values ranging from 52% to 78%. Sensitivity percentages were higher primarily for the Low SES groups of all three cohorts, with values ranging from 59% to 100%. For both the High SES and Low SES groups of the 2003-2004 cohort, sensitivity percentages increased from fall to spring in three of four MAP domains. For the High SES and Low SES groups of the 2004-2005 cohort, sensitivity percentages decreased from fall to spring for all four of the MAP domains.

It is important to note that the increases in sensitivity for the 2003-2004 cohort occurred in the presence of significant gains in all four MAP domains from fall to spring both in intervention efficiency and in instructional stability for High SES and Low SES group students in both cohorts. These increases in sensitivity simply mean that of the students identified as not proficient on a MAP domain in the spring of Grade 3, a greater proportion of that group of students was also identified as at-risk in Kindergarten, relative to similar values for MAP scores in the fall of Grade 3.

The table that follows illustrates this statistical outcome, using the data from the MAP fall and spring Comprehension domain assessment of the Low SES group from the 2003-2004 cohort.

Table 14

Grade 3 MAP Comprehension Domain Data for the Low SES Group of the 2003-2004 Cohort

Low SES Students

DIBEL K Sum	Fall MAP Comprehension		Spring MAP Comprehension	
	<u>Not Proficient</u>	<u>Proficient</u>	<u>Not Proficient</u>	<u>Proficient</u>
At-Risk	10	1	7	4
Not At-Risk	7	2	3	6

In the fall of Grade 3, a total of 17 students (85%) earned scores in the Not Proficient range in the MAP Comprehension domain. In the spring, that number who earned scores in the

Not Proficient range was reduced to 10 (50%). The overall reduction of the number of students in the Not Proficient range was accomplished through an increase in intervention efficiency (from 9% to 36%) and instructional stability (from 22% to 67%). The increase in intervention efficiency was achieved by moving 3 students who had been identified as at-risk in Kindergarten and who scored in the Not Proficient range in the MAP Comprehension domain in the fall into the Proficient range on the MAP Comprehension domain in the spring. The increase in instructional stability resulted when an additional 4 students who had been identified as not at-risk in Kindergarten but who earned MAP Comprehension in the Not Proficient range in the fall earned MAP Comprehension scores in the Proficient range in the spring. Despite these positive gains, the ability of the K Sum score to predict at-risk status (i.e., the sensitivity of the K Sum score) increased from the fall to the spring because of the manner in which Sensitivity values are calculated. In the fall, the sensitivity percentage was determined by dividing the number of students identified as at-risk in Kindergarten and who had earned not proficient scores in the fall (10) by the total number of students earning not proficient scores in the fall (17), yielding a value of 59% (10/17). In the spring, the sensitivity percentage was determined the same way, by dividing the number of students identified as at-risk in Kindergarten and who had earned not proficient scores in the spring (7) by the total number of students earning not proficient scores in the spring (10), yielding a value of 70% (7/10). These data illustrate the relatively restricted meaning of the sensitivity value and the potentially misleading nature of reporting this statistic as a metric for monitoring progress. The fact that the spring data represented a higher proportion of agreement between the K Sum score decision and the MAP Comprehension Score decision (70%) than the fall data (59%) does not mean that a larger number of students in the spring reverted to at-risk status that they had been assigned in Kindergarten, but rather that a greater

proportion of all the students earning not proficient scores on the MAP in the spring also earned at-risk scores on the DIBELS in Kindergarten.

In cases closer to the ideal, such as that presented in the data of the High SES and Low SES groups of the 2004-2005 cohort, instructional stability was maintained at a high rate and intervention efficiency increased substantially. In such cases, sensitivity percentages were influenced more heavily by the increases in intervention efficiency percentages, leading to decreased sensitivity proportions from fall to spring.

For purposes of progress monitoring, therefore, the more important statistics to focus on are intervention efficiency and instructional stability rather than sensitivity or specificity. The interrelationship of these variables is such that the sensitivity value is dependent on the degree of change in the intervention efficiency proportion and in the instructional stability proportion.

Specificity. All of the High SES and Low SES groups of the 2003-2004 and 2004-2005 cohorts demonstrated decreases in specificity values for three of four MAP domains between the fall and spring with the exception of the High SES Critical Content domain, which remained stable at 71%. As was evidenced with the pattern of changes in sensitivity values, the pattern of decreasing specificity values reflects the interrelationship between intervention efficiency and instructional stability. The specificity value is dependent on the degree of change in the intervention efficiency proportion and on the degree of change in the instructional stability proportion.

For the High SES and Low SES groups of both cohorts, the decreased specificity values indicated that students identified as not at-risk in Kindergarten made up a smaller proportion of the total number of students identified as proficient on MAP domains in the spring of Grade 3 than was the case for the MAP domains in the fall of Grade 3. As was the case with the

increased sensitivity, the decreased specificity reflected in changes in MAP scores from fall to spring occurred in the presence of reduced overall at-risk rates and increases in intervention efficiency and high rates of instructional stability, suggesting that specificity values are not a good metric for use in progress monitoring.

Kappa. Kappa values ranged from low (5%) to moderate (43%) for the High SES and Low SES groups of the 2003-2004 and 2004-2005 cohorts. For the High SES group of the 2003-2004 cohort, kappa values increased from fall to spring for 3 of the 4 MAP domains (Vocabulary, Comprehension and Analyze) and decreased for the other domain Critical Content. For the Low SES group of the 2003-2004 cohort, kappa values increased for the Comprehension and Critical Content domains, decreased for the Vocabulary domain and remained stable for the Analyze domain. For the High SES and for the Low SES groups of the 2004-2005 cohort, kappa values decreased from fall to spring for MAP domains.

As noted previously, kappa represents the percentage of accuracy greater than chance assignment to the categories of at-risk and not at-risk. It is important to note the disparity between the Kappa values of the 2003-2004 cohort and the 2004-2005 cohort. Kappa values are most sensitive to the overall consistency of at-risk and not at-risk classifications, but the complex interrelationship of sensitivity, specificity, instructional stability and intervention efficiency can make kappa a relatively unstable index for monitoring progress. Kappa values for the High SES and Low SES groups of the 2004-2005 cohort reflected a consistent pattern of decrease despite the fact that instructional stability decreased but intervention efficiency increased. In contrast, kappa values for the High SES and Low SES groups of the 2003-2004 cohort did not show a consistent pattern of decrease despite the consistent increases in intervention efficiency and in instructional stability demonstrated for all four MAP domains.

Contributions to the Field

The results of this study indicate that early and effective identification of at-risk readers is possible. The importance of using an empirically supported method of screening, as opposed to the more common, locally made screener, has been supported by the findings of the study. Additionally, the reality that early identification is possible leads to the importance of the next step, early intervention, and consequently supports the importance of early intervention for at-risk readers (NCLB, 2001). As research has shown, the earlier that effective intervention is provided, the greater the likelihood that remediation will take place (Bishop, 2003; Snow, Burns & Griffin, 1998). This study lends support to the need to provide early identification and early intervention for at-risk readers.

Results of this study also lend support to the field of RTI, which demonstrates the importance of progress monitoring. Students, identified at-risk in this study, who received intervention demonstrated progress and proficiency. Without progress monitoring, it would not have been possible to monitor these improvements and validate intervention efforts. Therefore this study supports the need for effective progress monitoring.

Limitations of the Study

The sample utilized in this study was one of convenience. Therefore it was not possible to control for external variables that may have interfered with results. External validity is limited by sample characteristics because all subjects attend the same elementary school. The narrow sample limits conclusion validity and weakens potential generalization of results.

The study did not control for involvement in special education or for students who received additional supports outside of school during the time of the study. Therefore the

possibility of additional interventions may have affected results and poses a threat to internal validity.

Because of the retrospective nature of the study, attrition is a limitation of the current design and is a potential threat to internal validity. The study involved examining variables over a several grade levels and attrition impacted the sample size; this consequently may have weakened the power of results.

Multiple examiners were used to administer the School-Made Kindergarten screening instrument as well as the DIBELS probes. Consequently, variability in administration procedures may present a threat to conclusion validity. Experimenter expectancy may also influence performance on the School-Made Kindergarten screening instrument and on DIBELS measures and could threaten construct validity.

Progress monitoring took place frequently throughout the school year. Based on progress monitoring information, interventions were provided to students determined to be at-risk. Information regarding the types of interventions and the frequency of delivery was not available. Therefore when examining the relationship between Kindergarten screening decisions and later reading achievement, validity may have been compromised.

Finally, data was collapsed into categories of at-risk and not at-risk. Collapsing data into dichotomous categories sacrificed sensitivity of the measures. Because of sensitivity was weakened, it is possible that more discrete trends among the data may have been overlooked.

Future Directions

Results supported the accurate early identification of at-risk readers. Currently, most schools utilize self-made screeners, which demonstrate weak predictive validity (Rafoth, 1997). Research has indicated that the best combination of measures to detect students who will be at-

risk of reading difficulties has not yet be established (O'Conner & Jenkins, 1999, Snow et al., 1998, Torgesen, 1998). Results of this study support the use of the DIBELS for early identification of at-risk readers. Future research should focus on the use of the DIBELS measures not only to identify at-risk readers, but also to provide early identification.

Significant findings regarding the successful early identification of at-risk students, and strong findings supporting the effectiveness of reading instruction and intervention were reported. Therefore it would be important to investigate the types of interventions employed, the rate and frequency of the interventions, and the success rates of each intervention. In order to ensure that at-risk students are receiving the most appropriate interventions, success rates of interventions should be investigated for at-risk students. Additionally, given the strong findings indicating that at-risk disadvantaged students be successfully identified, the relationship between the specific interventions utilized with disadvantaged students should also be explored. These investigations would enhance the current study by providing an RTI approach to the data and consequently would provide a more comprehensive program evaluation.

REFERENCES

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Berk, R. A. & Watson, T. S. (1995). Review of the Brigance K and 1 Screen for Kindergarten and First Grade. In *The twelfth mental measurements yearbook*. Retrieved June 1, 2007 from <http://exprozy.pcom.edu:2410/gwl/ovidweb.cgi>
- Bishop, A. G. (2003). Prediction of first-grade reading achievement: A comparison of fall and winter kindergarten screenings. *Learning Disabilities Quarterly*, 26, 189-200.
- Blachman, B. A. (1984). Relationship of rapid naming ability and language analysis skills to kindergarten and first-grade reading achievement. *Journal of Educational Psychology*, 79(4), 610-622.
- Bountress, N. G., & Cox, J. R. (1985). Review of the Lindamood Auditory Conceptualization Test. In *The ninth mental measurements yearbook*. Retrieved May 21, 2007 from <http://ezproxy.pcom.edu:2410/gwl/ovidweb.cgi>
- Brigance, A. (1992). Brigance K & 1 Screen for Kindergarten and First Grade—Revised. North Billerica, MA: Curriculum Associates.
- Capps, R., Fix, M. E., Murray, J., Ost, J., Passel, J. S., & Herwanto, S. H. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Retrieved May 29, 2007, from <http://www.urban.org/publications/311230.html>
- Cizek, G. J., & Fairbank, D. W. (2001). Review of the Developmental Indicators for the Assessment of Learning: Third Edition. In *The fourteenth mental measurement yearbook*. Retrieved May 21, 2007 from <http://ezproxy.pcom.edu:2410/gwl/ovidweb.cgi>

- Clark, E., & Earnhart, C. B. (1995). Review of the Developing Skills Checklist. In *The twelfth mental measurements yearbook*. Retrieved May 21, 2007 from <http://ezprozy.pcom.edu:2410/gwl/ovidweb.cgi>
- Costenbader, V., Rohrer, A. M. & Difonzo, N. (2000). Kindergarten screening: A survey of current practice. *Psychology in the schools*, 37(4), 323-332.
- CTB Macmillian McGraw Hill. (1990). *Developing Skills Checklist*. Monterey CA: CTB Macmillian-McGraw Hill.
- Denckla, M. B., & Rudel, R. G. (1976). Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471-479.
- Denton, K. & Germino-Hausken, E. (2000). *America's Kindergartners*, (NCES Publication No. 2000-070). Project Officer, Jerry West. Washington DC: U.S. Department of Education, National Center for Education Statistics.
- East Stroudsburg Area School District Report Card 2006-2007. Retrieved August 13, 2008 from: www.esad.net/curriculum/ayp_01_29_2008.pdf
- Felton, R. H. (1992). Early identification of children at risk for reading disabilities. *Topics in Early Childhood Special Education*, 12(2).
- Florida Center For Reading Research. (2006a). Retrieved May 14, 2007 from http://www.fcrr.org/assessment/pdf/prek_kassessments.pdf
- Florida Center For Reading Research. (2006b). Retrieved March 1, 2007, from <http://www.fcrr.org/Curriculum/glossary/glossaryOfReading.pdf>
- Fuchs, L. S. & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review*, 28(4), 659-671.

- Good, R. H. & Kaminski, R. A. (Eds.). (2002). Dynamic indicators of basic early literacy skills, (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- Gredler, G. R. (1997). Issues in early childhood screening and assessment. *Psychology in the schools, 34*(2), 99-106.
- Gunn, B. K., Simmons, D. C., & Kameenui, D. J. (1995). Emergent literacy: Synthesis of the research. National Center to Improve the Tools of Educators, funded by the U.S. Office of Special Education Programs.
- Ilg, F. & Ames, L., (1972). School readiness: Behavior tests used at the Gessell Institute. New York, NY: Harper and Row.
- Invernizzi, M. (2004). Early literacy screening in kindergarten: Widespread implementation in Virginia. *Journal of Literacy Research, 479-500*.
- Invernizzi, M. , Sullivan, A., & Meier, J. (2002). Phonological awareness literacy screening for preschool (PALS-PreK). Teacher's Manual. Charlottesville, VA: University Press.
- Invernizzi, M., Swank, L., Juel, C., & Meier, J. (2003). Phonological Awareness Literacy Screening—Kindergarten. Charlottesville, VA: University Printing.
- Kaminski, R. A. & Good III, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 259*(2), 215-228.
- Kurdek, L. A. & Sinclair, R. J. (2001). Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology, 93*(3), 451-455.
- Lennon, J. E., & Slesinski, C. (1999). Early intervention in reading: Results of a screening and intervention program for kindergarten students. *School Psychology Review, 28*(3), 353-364.

- Levine, M. D. & Reed, M. (2001). *Developmental variations and learning disorders* (2nd ed.). Cambridge, MA: Educators Publishing Inc.
- Lindamood, C. H. & Lindamood, P. C. (1979). Lindamood Auditory Conceptualization Test—Revised Edition. Riverside Publishing Co.
- Mardell-Czudnowski, C., & Goldenberg, D. S. (1990). DIAL-R: Developmental indicators for the assessment of learning manual (rev. ed.). Circle Pines, MN: American Guidance Service.
- National Reading Panel. (2000). *Teaching Children to read: An evidence-based assessment of the scientific research literature and its implications for reading instruction. Executive Summary*. Washington D.C. : U.S. Department of Health and Human Services, National Institutes of Health, Nation Institute of Child and Human Development.
- Neale, M. D. (1966). *Neale Analysis of Reading Ability*. London: Macmillan.
- No Child Left Behind Act of 2001. Washington, D.C. : U.S. Department of Education.
Retrieved May 2, 2007 from: <http://www.ed.gov/nclb/landing.jhtml>
- O’Conner, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, 3(2), 159-197.
- Rafoth, M. A. (1997). Guidelines for developing screening programs. *Psychology in the Schools*, 34(2), 129-142.
- Shaprio, E. S., Keller, M. A., Lutz, G. L., Santoro, L. E. & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment*, 24(1), 19-35.

- Share, D. L., Jorm, A. F., Maclean, R. & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology*, 76(6), 1309-1324.
- Snow, C. E., Burns, M. S. & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington DC: National Academy Press.
- Stahl, S. A. & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology*, 86(2), 221-234.
- Technical manual for the Northwest Evaluation Association (NWEA) measures of academic progress and achievement level tests (2003). Portland, OR: Northwest Evaluation Association.
- United States Census Bureau (2000). *East Stroudsburg borough, Pennsylvania American Fact Finder*. Retrieved March 2, 2008 from http://www.factfinder.census.gov/servlet/SAFFFacts?_event=Search&geo_id=&_geoContext=&_street=&_county=east+stroudsburg&_cityTown=east+stroudsburg&_state=04000US42&_zip=&_lang=en&_sse=on&pctxt=fph&pgsl=010&show_2003_tab=&redirect=Y
- VanDerHeyden, A. M., Witt, J. C., Naquin, G. & Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. *School Psychology Review*, 30(3), 363-382.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test—Revision 3*. Wilmington, DE: Jastak Associates.
- Woodcock, R. W., & Johnson, M. B. (1989). Woodcock-Johnson psychoeducational battery Revised. Allen, TX: DLM Teaching Resources.

Appendix A

Standardized Scoring for Kindergarten Registration
(Teacher-Given Test)

Rating: High Middle Low At-Risk
52-58 36-51 16-35 0-15

Score _____/58
Letter Recognition (Subtotal)
Upper Case: _____/26
Lower Case: _____/26

	Total Points	Rubric	Points Scores
1. Discussion	3	3- talks freely 2- responds to questions 3- need encouragement	
2. Counting	10	1 point for each number to 10	
3. Colors	8	1 point for each color	
4. Number Recognition	11	1 point for each numeral 1-10	
5. Name	3	3- perfect 2- mixed 1- letters	
6. Letter Names	10	10= 46-52 9= 41-45 8= 36-40 7= 43-35 6= 26-30 5= 21-24 4= 16-20 3= 11-15 2= 6-10 1= 1-5	
7. Designs	3	3-exact 2- disjointed 1- 1 shape	
8. Concepts of Print	3	3- reads 2- uses pictures to tell story 1- holds correctly	
9. Picture of a Person	4	4- detailed 3- stick 2- sun 1- attempt	
10. Attention Span	3	3- focuses 2- redirection 1- off task	

Comments _____

Speech _____ Occupational Therapy _____ Physical Therapy _____ ESL _____
Teacher's Name _____ Date Given _____