Abstract

LAUGLE, KELLY MARIE. Measuring Opportunities to Learn from Scientifically Based Reading Instruction. (Under the direction of Jeffery P. Braden.)

This study was an effort to determine the psychometric properties of an instrument designed to measure a student's previous opportunities to learn from reading instruction. This instrument (Teachers' Implementation of Scientifically Based Reading Instruction (TISBRI)) is based on Porter and Smithson's (2001) framework for systematically evaluating opportunities to learn, and its content stems from *Put Reading First: The Research Building Blocks for Teaching Children to Read* (Armbruster, Lehr & Osborn, 2001). The pilot version of the TISBRI was administered to 19 teachers who rated the reading instruction they had delivered to a single student in their class since the beginning of the school year. Item analyses, reliability, and validity studies were conducted to determine the merits of the instrument and refine it for future use. The instrument showed promise for having strong internal consistency but no conclusions could be drawn for the stability of the instrument, the relationships between scales in the instrument, and its ability to detect grade level differences in response patterns.

Measuring Opportunities to Learn

from Scientifically Based Reading Instruction

by

Kelly M. Laugle

A thesis submitted to the Graduate Faculty of

North Carolina State University

in partial fulfillment of the

requirements for the Degree of

Master of Science

Psychology

Raleigh, North Carolina

2006

APPROVED BY:

_____ _____

_____

Chair of Advisory Committee

Biography

An Oregonian at heart, Kelly Marie Laugle spent most of the first eighteen years of her life enjoying the beautiful outdoors of the Pacific Northwest with her parents Thomas and Jeannette and her older brother Neal. From 1997 to 2001 she attended Gonzaga University in Spokane, Washington where she majored in Psychology and Art and continued to pursue her love for the outdoors. During her junior year, Kelly lived in Florence, Italy where she studied art, ate countless bowls of pasta, and learned to love *vino*. Upon graduating from college, Kelly joined Teach For America and spent five memorable weeks living in Houston, Texas enduring a "crash-course" in teaching. After recovering from the course, Kelly moved to Henderson, North Carolina where she taught one year of fourth grade and one year of second grade at Clark Street Elementary School. In 2003, Kelly entered the School Psychology program at NC State, and focused her studies on early elementary reading instruction. She plans to continue this focus when she returns home in the fall of 2006 to complete a doctoral degree in School Psychology at the University of Oregon.

Acknowledgments

The completion of this thesis was made possible by the support and guidance I received from numerous people along the way. First, I would like to thank my advisor Jeff Braden for challenging me to think critically and to express my thoughts clearly and concisely. I would also like to thank Ann Schulte for her thought-provoking comments and stimulating chats in the hallway and Susan Osborne for her curriculum resources and assistance with data collection.

I am extremely grateful to the kindergarten through third grade teachers who took time out of their busy days to complete my instrument and for the assistance I received from Ann Brady of Rockingham Public Schools as well as Lynne Myers, Carolyn Patton, and Lois Huffman of Wake County Public Schools. I also want to thank Sandye, Hailey, Michelle, Athena, Kendrea, Katie, Kristina, Lauren, and John for helping me develop my instrument, score numerous student work samples, and prepare for my defense. Additionally, the completion of this thesis would not have been possible with out the assistance of the 7th floor printer and my Kinkos discount card.

I am deeply indebted to my friends and family members who have provided support to me over the past three years. To my mentor Susan Robinson: you are an amazing school psychologist and friend; thank you for keeping me motivated and caffeinated. To Jenny: you are the best running partner ever; thanks for all the laughs and good times on the trails. To my dad: you taught me how to challenge myself; thank you for reminding me to dream big and make time for fishing. Finally, to my mom: you are the hardest working, most altruistic person I know; thank you for inspiring and supporting me every step of the way.

Table of Contents

List of Tables

Chapter 1

Introduction and Review of Research Literature

The Individuals with Disabilities Education Improvement Act of 2004 includes a unique rule for determining a student's eligibility for services. This rule mandates that a student shall not be identified as having a disability if his or her insufficient progress is rooted in a lack of scientifically based reading instruction (PL No: 108-446). The S*pecial Rule for Eligibility Determination* states:

> ` (5) In making a determination of eligibility under paragraph (4)(A), a child shall not be determined to be a child with a disability if the determinant factor for such determination is –
>
> > ` (A) lack of appropriate instruction in reading, including in the essential components of reading instruction (as defined in section 1208(3) of the Elementary and Secondary Education Act of 1965) (Sec. 614(b)(5)(A)).

The *Special Rule for Eligibility Determination* implies that a student's previous opportunities to learn from appropriate reading instruction can and should be measured as part of determining eligibility for special education services. Several factors such as student attendance, teacher training, administrative support, and instructional resources can influence the reading instruction a student receives. Therefore, to fully determine if a student is lacking in appropriate instruction, previous learning opportunities need to be assessed in a systematic way.

Porter and Smithson (2001) developed a framework for systematically evaluating opportunities to learn. They identified four curricular areas that influence learning: the

intended curriculum, e*nacted curriculum, assessed curriculum,* and *learned curriculum*. The intended curriculum refers to policy documents and other publications that outline objectives for instruction in each grade level and content area. The enacted curriculum refers to opportunities teachers provide within their classrooms. The enacted curriculum includes the content that is covered, as well as the way that content is delivered. Those objectives outlined in the intended curriculum that are represented on measures of student achievement comprise the assessed curriculum. Finally, the learned curriculum includes the students' demonstrations of learning and levels of proficiency reflected in their performance on the assessments.

Using Porter and Smithson's (2001) systems model, opportunities can be assessed within each curricular area as well as across areas using measures of alignment. When alignment exists across all curricular areas, all levels of the educational system are working together to maximize opportunities for student learning. Misalignment is likely to reduce students' opportunities to learn the intended curriculum. For example, teachers may only cover half of the objectives presented on a state-wide assessment, yielding misalignment between the enacted and assessed curricula resulting in less opportunity to learn omitted content. A misalignment could also exist between the intended and assessed curricula if teachers are drawing their instruction from a particular set of standards (while ignoring others), or using assessment tools that do not capture what they taught. In addition, students may receive learning opportunities that are not outlined in the standards for a particular content area. Given that instructional time is finite, attention to standards not included in the intended curriculum necessarily reduces the time available to cover topics in the intended

curriculum. Misalignment, however, generally limits the learning opportunities students receive relative to the intended curriculum outlined by their district or state.

The purpose of this study was to describe the psychometric properties of an instrument designed to measure a student's previous opportunities to learn from reading instruction. This instrument is based on Porter and Smithson's (2001) model and includes measures of the enacted, assessed, and learned curricula. Although Porter and Smithson's work allowed the intended curriculum to vary by subject matter and educational agency (e.g., district or state standards), the intended curriculum for this study is assumed. [The intended curriculum is presumed to be the five essential components of scientifically based reading instruction identified in federal legislation (PL 107-110, Sec. 1208 (3)).]

In the following section, I will provide definitions for each of the five components of scientifically based reading instruction and review the literature that pertains to their development. Then, using a top-down approach, I will discuss how opportunities to learn have previously been measured at various levels of the educational system beginning with measures at the national/state level and concluding with measures at the individual level. Each method for assessing opportunities to learn will be discussed in the context of the intended, enacted, assessed, and learned curricula.

Definitions for the Essential Components of Reading Instruction

Put Reading First (Armbruster, Lehr, & Osborn, 2001) identifies instruction in the areas of *phonemic awareness*, *phonics*, *fluency*, *vocabulary*, and *text comprehension* as essential to student success in reading. The National Reading Panel (2000a, 2000b) states that each of these components must be addressed through systematic and explicit instruction.

Phonemic awareness instruction includes activities that involve hearing, identifying, and/or manipulating sounds in spoken words. Phonics instruction focuses on learning the relationships between letters of written language and the sounds of spoken language. Activities that improve the accuracy, speed, and expressiveness of oral reading are identified as fluency instruction. Vocabulary instruction involves both direct and indirect development of oral vocabulary and reading vocabulary, whereas text comprehension instruction focuses on direct instruction of comprehension strategies and the application of those strategies to text. Literature pertaining to the development of the essential components of reading instruction will be reviewed next.

Development of the Essential Components of Reading Instruction

In the past decade two important committees were charged with evaluating the research base on reading and translating this research into guidelines for a national curriculum. The National Research Council Committee on the Prevention of Reading Difficulties in Young Children outlined the skills, environments, and early developmental interactions that are essential to learning to read (Snow, Burns & Griffin, 1998). The report also provided recommendations for practice and for future research. In 1997, in response to a Congressional request, the National Reading Panel (NRP) convened to continue the previous committee's examination of the research on reading. The NRP focused primarily on identifying "'how' critical reading skills are most effectively taught and what instructional methods, materials, and approaches are most beneficial for students of varying abilities," (NRP, 2000b, p. 1). The effort of these two committees to ensure the quality of the intended

curriculum was an important initial step in maximizing students' opportunities to learn from reading instruction.

Of approximately 100,000 research studies on reading that were published since 1966, 418 studies met criteria for sound research and were critically reviewed by the Panel (Shanahan, 2003). The NRP's findings were published in the *Reports of the Subgroups* (2000a) and then consolidated in a more user-friendly document referred to as the *Summary* (2000b). Although federal legislation provides support for the five essential components of reading instruction identified by the NRP, controversies related to some of the Panel's findings should be highlighted. A majority of these controversies focused on the process by which the Panel conducted its reviews, rather than its specific findings (Shanahan, 2003).

Garan (2001), Cunningham (2001), and Krashen (2001) attacked the scientific rigor with which the Panel selected and reviewed articles. Garan (2001) focused her criticism on the subgroups' report on phonics instruction. She noted that studies were combined in meta-analyses that had very different outcome measures. For example, studies measuring growth based on the ability to decode nonwords were compared to studies measuring growth on more authentic measures of reading. Garan suggested that because of this methodological flaw, the Panel's conclusions on phonics instruction may have been misguided. Garan also pointed out discrepancies between what was stated in the *Reports of the Subgroups* (NRP, 2000a) and the conclusions drawn in the *Summary* (NRP, 2000b). Other critics of the phonics subgroup report argued that the importance of phonics instruction was overemphasized at the expense of other elements of reading instruction (Cunningham, 2001; Zimmerman & Brown, 2003).

Krashen (2001) criticized the Panel for excluding correlational studies that support independent, silent reading as part of daily instruction. The Panel stated that there was insufficient evidence to conclude that independent, silent reading improves reading achievement and fluency. Krashen contended that studies demonstrating no difference between students engaged in independent reading and those engaged in traditional programs lend support to silent reading's effectiveness as an instructional tool.

Although objections do exist regarding the process by which the NRP came to its conclusions, much of the Panel's findings continue to be supported by current research. Foorman et al. (2003) conducted a large scale study of reading instruction in kindergarten and first grade classrooms and found support for the Panel's conclusions regarding the importance of phonemic awareness instruction. The RAND Reading Study Group (2002) summarized the research base for comprehension instruction and supported many of the NRP's findings. Specifically, the Study Group concluded that explicit instruction in the use of metacognitive strategies improves reading comprehension, especially for low achieving students. In addition, the Study Group supported the NRP's identification of the following set of strategies to improve reading comprehension: concept mapping, question generating, question answering, summarizing, and story mapping. The Study Group also concluded that the use of reading fluency techniques such as repeated oral reading improves word recognition and fluency and produces moderate gains in comprehension.

To bridge the research to practice gap, the Center for the Improvement of Early Reading Achievement used the NRP's *Reports of the Subgroups* (2000a) to develop *Put Reading First: The Research Building Blocks for Teaching Children to Read* (Armbruster,

Lehr, & Osborn, 2001). This publication describes the five essential components of reading instruction, summarizes the research related to each component, and provides learning objectives, student strategies, and teacher strategies to improve student outcomes in reading for students in kindergarten through third grade.

Shanahan (2003), a member of the Panel, supported *Put Reading First* as a meaningful and useful extension of the *Reports of the Subgroups* (NRP, 2000a). Shanahan identified only one aspect of *Put Reading First* that was questionable based on the NRP's findings: that adding phonics workbooks or phonics activities to non-systematic reading programs such as literature-based, basal reading, or sight-word programs "has not been effective" and "confuse[s] rather than help[s] children to read" (Armbruster, Lehr, & Osborn, 2001, p. 17). Overall *Put Reading First* has been well accepted and widely distributed (Shanahan, 2003). All item content for the instrument used in the present study stems directly from *Put Reading First* and does not address the area of contention noted by Shanahan.

The five essential components of reading instruction, as identified and described in *Put Reading First*, represent the intended curriculum guiding the development of the instrument used in the present study. Including a well-defined intended curriculum in a measure of opportunities to learn facilitates analyses of alignments across curricular areas. The following section will address several methods that have been developed to measure opportunities within and across the four curricular areas identified by Porter and Smithson (2001).

Methods for Assessing Opportunities to Learn

Measures of opportunities to learn address the quality of instruction, time, and resources provided to students in any academic area. Methods for evaluating opportunities to learn will be reviewed at the national/state, school, classroom, and individual levels.

Measuring opportunities to learn at the national and state level. Methods at the national and state level have focused primarily on the alignment between the intended and assessed curricula. Webb (1999) addressed the alignment of these curricula by comparing state standards for instruction in mathematics and science to yearly state assessments in those areas. Webb defined alignment as "the degree to which the standards and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what they are expected to know and do" (p. 2). After measuring the alignment of sixteen different assessments with their corresponding state's standards, Webb concluded that, in general, most assessments demonstrated categorical concurrence (all items on the tests reflected state objectives). However, Webb found that the assessments required a lower level of knowledge and exhibited less breadth of coverage than was mandated by the standards for instruction in each state. Porter (2002) also developed content analysis tools that have been used to analyze the content of seventh grade mathematics standards and tests.

Both Webb's (1999) and Porter's (2002) measures of alignment help evaluate the degree to which educational agencies send a clear and consistent message to teachers on what they are expected to teach and what the agency will assess. Misalignment of expectations (i.e., the intended curriculum) and the test used to evaluate student achievement (i.e., the

assessed curriculum) may encourage teachers to distort or reduce students' learning opportunities (e.g., teach to the test).

Analyzing the intended curriculum's alignment with relevant policy documents and with the latest research findings is another important step in ensuring opportunities to learn (Porter & Smithson, 2001). Text book analysis provides one way to assess the intended curriculum (Porter, 2002). Our current national policy for special education reform assumes that by ensuring reading texts are based on sound research, schools can improve opportunities for student success (Individuals with Disabilities Education Improvement Act of 2004).

Recently, many states receiving Reading First grant money reviewed the degree to which specific reading programs were aligned with scientifically based reading research so as to make specific recommendations to districts and schools for program selection (Oregon Reading First Curriculum Review Panel, 2004). Oregon, Washington, Arizona, and Colorado used *The Consumer's Guide to Evaluating a Core Reading Program Grades K-3: A Critical Analysis* (Simmons & Kame'enui, 2003) to provide quantitative data on the design and delivery features of comprehensive reading programs. Using this guide, a panel of experts rated the degree to which the program under review met the criterion for specific items related to the essential components of reading instruction. This method of analysis provided objective data to help schools choose research-based programs that would meet the needs of their students and hopefully facilitate opportunities to learn.

To maximize these opportunities, teachers must deliver the intended curriculum with high degrees of fidelity. Many publishing companies of reading texts have taken steps to

ensure that teachers understand the rationales behind program development and maximize the use of their programs for instruction. The *Teacher's Guide for Reading Mastery* (Engelmann & Bruner, 1995) provides rationales for each skill to be taught and includes explicit scripts for teachers to follow for each lesson within the *Reading Mastery* curriculum. Guidelines are provided for how to group students, how to schedule daily lessons, and how to use the presentation books and the teacher's guide. Houghton Mifflin Company (2002) provides detailed lesson plans that outline the amount of time, objectives, and resources for each day of instruction for their reading curriculum. In addition, the company offers two-day professional development workshops for each of the essential components of reading instruction.

Steps taken by publishing companies to ensure that the enacted curriculum aligns with the intended curriculum can be important to promoting student learning opportunities. However, the utility of these efforts is limited by the degree to which the programs themselves align with research findings on reading instruction, the degree to which teacher's actually implement the programs, and the degree to which these programs address students' individual instructional needs.

Measuring opportunities to learn within and across schools. Opportunity to learn may explain between-group differences in performance within and across schools. Many studies have used high school course transcripts to measure the learning opportunities of groups of students within schools, specifically Limited English proficient students, low-income students, and minority students (Mickelson, 2001; Oakes, 1985; Wang & Goldschmidt, 1999). This research has called attention to the common practice of tracking

students into lower level classes consequently limiting their access to portions of the enacted curriculum. Herman and Klein (1996) used teacher interviews, student surveys, student retrospective think-aloud interviews, and classroom observations to evaluate differences in learning opportunities for three broad categories of schools: affluent suburban, low socio-economic status urban, and remote, mixed socio-economic status rural. This area of research has provided unique ways to measure and define opportunities to learn and has suggested that factors such as the personal characteristics of the student (e.g., race, language proficiency, and socio-economic status) and the type of school a student attends can enhance or limit access to these opportunities.

Measuring opportunities to learn within the classroom. Focusing on the enacted curriculum (what the teacher is actually doing in the classroom) provides the most direct measure of students' opportunities to learn. Porter and Smithson (2001) stated "the *enacted* curriculum is arguably the single most important feature of any curricular indicator system" [emphasis in the original] (p.2). Stevens and Grymes (1993) identified four distinct variables related to the enacted curriculum: *content coverage*, *content exposure*, *content emphasis*, and *quality of instructional delivery*. *Content coverage* analysis involves identifying which objectives in the curriculum actually get addressed in instruction, where as *content exposure* analysis involves investigating the amount of time allocated for instruction. Analysis of c*ontent emphasis* focuses on the relative attention given to objectives within the curriculum as well as to students with lower and higher level skills. Finally, analysis of the *quality of instructional delivery* involves evaluating how teaching practices influence student achievement.

Several methods have been developed to evaluate these variables within the enacted curriculum. One method, direct observation, may provide the most objective and rich information; however, it also requires more resources than are commonly available for conducting most research. Some of the more promising methods of data collection involving teacher self-report include the Surveys of the Enacted Curriculum (SEC) (Blank, 2002), teacher interviews and student surveys combined with the collection of classroom artifacts (Burstein et al., 1995; Herman, Klein & Abedi, 2000), and web-based teacher logs (Ball, Camburn, Correnti, Phelps & Wallace, 1999). Each of these methods will be reviewed as it pertains to measuring opportunities to learn within the classroom. Research on the reliability and validity of using teacher self-report measures will also be discussed.

The Surveys of the Enacted Curriculum (SEC), developed by the Council of Chief State School Officers and the Wisconsin Center for Education Research, are a set of tools for data collection, data analysis, and reporting that provide information on the alignment of teachers' instruction with national and state standards and assessments (Blank, 2002). Good alignment suggests that students are receiving adequate opportunities to learn. Until recently the SEC have focused exclusively on math and science education. The Surveys now include a single generic measure of language arts and reading instruction for kindergarten through twelfth grade teachers. This survey provides information on a broad range of instructional activities; however, it relies on the assumption that each student within a class in exposed to the same instruction (SEC Online, n.d.).

Herman, Klein, and Abedi (2000) used classroom artifacts to supplement teacher interviews and student surveys allowing for more in-depth analyses of learning opportunities.

Teachers were asked to provide typical assignments and tests given during the previous quarter of instruction. These materials were coded by two raters who judged the proportion of items within the set that reflected various reform practices that were queried on the surveys. The classroom artifacts scale was found to correlate well with both teacher interview results (r = .69) and student survey results (r = .64).

Ball et al.'s (1999) web-based teacher logs provide a method for integrating information on daily teacher practices that address whole class, small group, and individual instruction. The web-based logs provide data on both the content and pedagogy of teachers' instruction as well as the materials used, other adults involved in instruction, student activities, and student attendance. Data from the pilot study suggest that web-based teacher logs can accurately capture information regarding the number of reading and mathematics lessons, duration of each lesson, student attendance, materials used, and instructional grouping. The collection of additional information such as the content of instruction, the nature of instructional tasks, and teacher and student activities was limited by a lack of common language for teachers to use in describing their daily instruction.

As part of the *Reform Up Close* study, Smithson and Porter (1994) evaluated the use of daily logs, weekly surveys, observations, interviews, and comprehensive questionnaires. They found moderate to strong levels of agreement between daily logs and direct observations of teachers (average percents of agreement ranging from .59 to .78) and promising, yet inconclusive, results for the agreement of daily logs with related responses on questionnaires. These results suggest that questionnaires administered once or twice a year

may not reliably capture elements of instruction at the same level of specificity as can be captured using daily logs.

The use of teacher self-report to evaluate students' opportunities to learn has been the subject of considerable debate. Smithson and Porter (1994) believe there is ample evidence to support the use of survey instruments, whereas Snow-Renner (2001) has cautioned that large-scale surveys, such as the Third International Mathematics and Science Study, lack sufficient technology to provide data beyond general information about students' experiences. Mayer (1999) found encouraging support for the reliability and validity of teacher self-report surveys on composites of instructional practices. Self-report of the percent of time spent on a composite of 13 instructional practices demonstrated high test-retest stability over the course of four months (r = .68) and correlated strongly with direct observations (r = .85). Results were less encouraging when instructional practices were examined individually, although the rank order of time devoted to each instructional practice was generally maintained on the second survey administration. Shim, Felner, Shim, and Noonan (2001) confirmed Mayer's (1999) findings that the reliability of measuring instructional practices improves when items are grouped to reflect underlying characteristics. Other research has also supported the reliability and validity of teacher self-report (Burstein et al., 1995).

Porter and Smithson (2001) and Blank (2002) have cautioned against the use of teacher self-report measures when these measures are linked to rewards or sanctions that might cause teachers to intentionally misrepresent aspects of their instruction. Teachers may also unintentionally misrepresent aspects of their instruction because they are unable to recall

elements of their instruction or are unfamiliar with the terminology used in the measure (Porter & Smithson, 2001).

Overall the research has provided three recommendations for developing a reliable and valid teacher self-report measure. First, analyses at the item level for teacher self-reports might not provide meaningful information, but these items may be aggregated to reflect underlying characteristics that produce reliable and valid measures of teachers' instruction (Mayer, 1999; Shim, Felner, Shim & Noonan, 2001). Therefore, multiple indicators are needed for each domain the survey purports to address. Second, to reduce biased responses, teachers should remain anonymous and no rewards or sanctions should be attached to an evaluation of their responses (Blank, 2002; Porter & Smithson, 2001). Finally, analyses of validity such as the correlation of self-report with classroom artifacts should be conducted to assess both intentional and unintentional misrepresentation of instructional practices (Porter & Smithson, 2001).

Measuring opportunities to learn for an individual student. Except for portions of data collected using the web-based teacher logs, each of the classroom based methods I have reviewed focuses on the instruction delivered to an entire group of students. These methods assume that students within a class have equal opportunities to learn. Ball et al. (1999), however, identified potential variations in classroom instruction that may exist including "variation in content delivered to particular students, variation due to grouping practices, materials, or representations used to deliver content to different students" (p. 7). Dreeben and Barr (1988) conducted full-day observations, at 12 3-week intervals, for 13 first grade classes including a total of 50 instructional groups for reading. They observed that assigning

first grade students to reading groups based on ability often placed "non-English-speaking and extremely low-aptitude students in limbo, where they initially receive[d] either no grouped instruction or infrequent individualized instruction" (p. 136). Floden (2003) expanded the conception of an individual student's opportunities to learn to include out-of-school opportunities. He noted, "In acquisition of written literacy, for example, children with literate parents have out-of-school opportunities to learn that are not available to [children of] illiterate parents" (p.255). Documented variations among students in the same classroom suggest that assessments of learning opportunities must include individual indicators in addition to those shared by the classroom or instructional group.

Shapiro (1996) developed a structured teacher interview to address an individual child's academic problems that includes a section on delivery of reading instruction. This qualitative form of data collection addresses the type of reading series or program used by the teacher, the grouping of students by ability, the allocation of time to different types of instruction (e.g., independent seatwork, small group, and cooperative groups), outside opportunities for reading instruction, and methods of assessment. The teacher is also asked to compare the student's performance to other students within the same reading group on skills related to oral reading, word attack, sight words, and comprehension. This structured interview is primarily used to guide the development of effective interventions, but may also serve as an indicator of an individual student's previous opportunities to learn.

Assessing the opportunities a teacher provides to an individual student is necessary but not sufficient to determine that student's previous opportunities to learn. Access to the enacted curriculum must also be considered. At the school level, tracking may limit student

access to specific courses (e.g., Oakes, 1985). At the individual level, a student's attendance may also limit access to the enacted curriculum. Students with frequent absences and tardiness have fewer opportunities to learn than students who attend school regularly. Walker-Hill, Block-Pedego, Todis, and Severson (1991) developed the *School Archival Records System* to track eleven student variables that may limit access to the enacted curriculum including days absent, number of schools attended, within-school referrals, special education status, out-of-class placement, and Title I status. Although Walker-Hill and colleagues designed this system to identify students at risk for behavior disorders and at risk for dropping out of school, portions of this archival rating system could be used to measure access to opportunities to learn for individual students.

Visual impairments and limited English proficiency (LEP) are two additional factors that may influence an individual student's access to learning opportunities. Bradley-Johnson and Morgan (2002) suggested that teaching visually impaired students in the general classroom without sufficient resources may limit their access to quality instruction in Braille. Scribner (2002) reported that "nearly one-third of LEP students receive no tailored assistance in understanding what is being taught. This means that these students receive limited instruction in how to speak English or in understanding content area subjects" (p. 1485). Other factors that may influence an individual student's access include illness, hunger, and other forms of disability.

Summary

The first section of this review addressed the development of the essential components of reading instruction and provided definitions for their use in the present study.

Put Reading First (Armbruster, Lehr, & Osborn, 2001) provides the blueprint for an intended reading curriculum for early elementary instruction. It is well supported by previous and current research, applies to classrooms across the nation, and is specified in federal legislation. For these reasons it was selected to serve as the intended curriculum for the present study, and to then guide the development of instrument content.

The second section of this review addressed previous methods for evaluating opportunities to learn. There are a number of critical points to consider when developing measures of opportunities to learn, including the alignment between curricular areas, the unit of analysis (e.g., state, school, classroom, or individual), and the method of recording data (e.g., observations, teacher surveys, daily logs, or interviews). Measures that focus on the quality of the intended curriculum and its alignment with the assessed curriculum promote opportunities for student learning, but do not directly assess whether these opportunities actually occur. Classroom based measures that focus on the alignment of the enacted curriculum with the intended, assessed, and learned curricula provide more direct measures of students' learning opportunities. However, these measures generally assume that students within a class share the same opportunities to learn. Researchers such as Dreeben and Barr (1988) have called this assumption into question. Therefore, to adequately determine an individual student's previous learning opportunities, a measure must focus on the enacted curriculum that the individual student actually receives, evaluate the degree of alignment across curricular areas, and consider factors which may limit the student's access to the enacted curriculum.

Teacher self-report, although not an ideal tool for measuring an individual student's previous learning opportunities, appears to be a viable option especially if the instrument's content is aligned with an intended curriculum that is supported by sound research and if multiple indicators are developed for each domain to be addressed. The use of teacher self-report on a measure of reading instruction may be further enhanced by collecting assessment products to supplement the self-report and by ensuring that teacher responses remain anonymous.

Currently no tool exists that measures the opportunities provided to a single student for learning to read that uses the latest national consensus on reading research, and is based on a sound theoretical model for systematically measuring opportunities to learn. Such an instrument could provide a meaningful and potentially useful tool for measuring a student's previous opportunities to learn from explicit and systematic instruction in the five essential components of reading.

Rationale for the Present Study

The S*pecial Rule for Eligibility Determination* in the Individuals with Disabilities Education Improvement Act of 2004 provides the primary rationale for developing an instrument to assess teachers' implementation of scientifically based reading instruction for an individual student. The legislation requires educators to determine whether a child's insufficient progress in reading is due to lack of opportunity, but does not specify methodologies for doing so. The instrument developed for this study, which I have labeled Teacher's Implementation of Scientifically Based Reading Instruction (TISBRI), could

provide such a measure. This study was an effort to describe the psychometric properties of this instrument.

Predictions

The specific goals of this study include developing a scoring method for the TISBRI, establishing a final item pool, and determining the instrument's reliability and validity as a measure of teacher self-report of reading instruction. To address these goals, seven predictions were tested. The first three predictions are supported by principles of test development (Gregory, 2000), the fourth prediction is supported by *Put Reading First* (Armbruster, Lehr & Osborn, 2001) and findings of the National Reading Panel (2000a), and the final three predictions are supported by Porter and Smithson's (2001) research on the alignment of curricular areas within an educational system. Operational definitions for each of these predictions are provided at the conclusion of the Methods chapter.

Prediction 1. The final item pool for the TISBRI will have adequate psychometric characteristics.

a) Items will be positively and meaningfully correlated with other items related to the same component.

b) Items will be positively and meaningfully correlated with component scale composites and the total TISBRI composite.

c) The item pool will have a simple factor structure that corresponds to the underlying component structure of the TISBRI.

Prediction 2. The TISBRI will be a stable measure of teachers' self-reported enacted reading curricula.

Prediction 3. The rubric developed to score the assessment products will demonstrate acceptable inter-rater agreement.

Prediction 4. Emphasis on the five components of reading instruction will shift from phonemic awareness, phonics, and fluency instruction in earlier grades to emphasis on vocabulary and text comprehension instruction in later grades.

Prediction 5. Scale scores addressing the enacted curriculum for the teacher self-report portion of the TISBRI will correlate positively with scale scores of the assessed curriculum.

Prediction 6. Scale scores addressing the enacted curriculum will correlate positively with scale scores addressing the learned curriculum.

Prediction 7. TISBRI component scores addressing the enacted curriculum will correlate positively with teachers' estimates of proportion of time devoted to instruction in each of the components.

Chapter 2

Method

Respondent and Student Sample

The respondents in this study were 16 classroom reading teachers, 2 resource teachers, and 1 Title I teacher who completed a rating scale for one (n = 19) of their students in grades kindergarten through 3. The sample was drawn from schools in two counties in North Carolina. Descriptive data including ethnicity, gender, and degree obtained were collected for each teacher and are presented in Table 1. Additional information including the number of courses taken in reading instruction, years of teaching experience, and grade taught were also collected and are presented in Table 2. Teachers removed all identifiers of the students from the assessment products they submitted. The students' identities remained anonymous to the researcher throughout data collection. Descriptive data collected for the students included the amount of reading instruction they received relative to the other students in the class (*more*, *about the same*, or *less*), grade, retention, age, gender, and ethnicity (see Table 3).

Instrumentation

Teachers' Implementation of Scientifically Based Reading Instruction (TISBRI) is the instrument developed for use in this study. The initial development of the TISBRI is described in Appendix A. The TISBRI is designed to measure teachers' implementation of the five essential components of scientifically based reading instruction as described in *Put Reading First* (Armbruster, Lehr & Osborn, 2001). This instrument is focused on the delivery of instruction to a single student during the course of a school year. The background

Table 1

Descriptive Data for Teachers Sampled Within Each Grade Level

Grade Level	K	1[a]	2	3	K-3[a]
			% (*n*)		
Demographic Characteristics					
Female	37 (7)	32 (6)	21 (4)	5 (1)	95 (18)
Caucasian	32 (6)	26 (5)	21 (4)	5 (1)	84 (16)
African American	5 (1)	5 (1)	0 (0)	0 (0)	11 (2)
Hispanic	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Asian	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Educational Background					
Bachelors	21 (4)	16 (3)	5 (1)	5 (1)	47 (9)
Masters	16 (3)	16 (3)	16 (3)	0 (0)	47 (9)
Total *N*	7	7	4	1	19

[a]One teacher within this group did not report gender, ethnicity, or educational background.

Table 2

Mean Number of Reading Courses and Years of Teaching Experience for Teachers Sampled Within Each Grade Level

Grade Level	K	1	2	3	K-3
Reading Courses	3.50[a] (1.52)	4.29 (1.70)	3.67[a] (0.58)	3.00 (0.00)	3.82 (1.42)
K-3 Experience	6.29 (6.73)	11.00 (12.62)	16.75 (5.25)	5.00 (0.00)	10.16 (9.68)
Total Years Teaching	9.17 (6.46)	16.43 (14.35)	21.00 (10.10)	5.00 (0.00)	14.39 (11.39)
Total *N*	7	7	4	1	19

Note. Values in parentheses represent standard deviations.

[a]One teacher within this group did not report the number of reading course completed.

Table 3

Descriptive Data for Students Sampled Within Each Grade Level

Grade Level	K	1	2	3	K-3
			% (*n*)		
Instructional Setting					
General Education	32 (6)	32 (6)	16 (3)	5 (1)	84 (16)
Resource	0 (0)	5 (1)	5 (1)	0 (0)	11 (2)
Title 1	5 (1)	0 (0)	0 (0)	0 (0)	5 (1)
Demographic Characteristics					
Female	11 (2)	26 (5)	0 (0)	0 (0)	37 (7)
Caucasian	5 (1)	26 (5)	5 (1)	5 (1)	42 (8)
African American	16 (3)	11 (2)	11 (2)	0 (0)	37 (7)
Hispanic	11 (2)	0 (0)	5 (1)	0 (0)	16 (3)
Asian	5 (1)	0 (0)	0 (0)	0 (0)	5 (1)
Other	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Retained	0 (0)	5 (1)	5 (1)	0 (0)	11 (2)

Table 3. (cont.)

Grade Level	K	1	2	3	K-3
			M (*SD*)		
Months of Instruction	4.57 (3.10)	6.57 (2.51)	7.75 (2.22)	6.00 (0.00)	6.05 (2.78)
Age	5.29 (0.49)	6.33 (0.52)	8.00 (0.82)	8.00 (0.00)	6.38 (1.24)
Total *N*	7	7	4	1	19

information portion of this instrument includes data on factors that may influence the enacted curriculum a student receives. These factors include level of teacher training and experience, opportunities for assistance outside of the classroom, and the kind of reading program used by the teacher.

The rating scale portion of the instrument is focused on the enacted and learned curricula. Teachers were asked to determine the frequency and quality of their instruction for items reflecting each of the five components. To measure the degree of student learning for each item, teachers were asked to determine the level of progress their students had made in response to instruction. The assessment portion of this instrument required teachers to provide information on the type and frequency of assessments they used to address each of the five components. In addition, teachers provided three different examples of assessments they had conducted during the course of the school year with the student they selected. These assessments were evaluated based on the degree to which they reflected each of the essential components of reading instruction and their alignment with teacher self-reports.

Description. The TISBRI includes four parts: *Background Information*, *Implementation*, *Components of Instruction*, and *Assessment* (see Appendix B for the field trial version used in this study). In part one, Background Information, respondents selected the middle student from their "middle" reading groups to focus their responses or selected the middle student from their class distribution based on reading level or current grade in reading. The respondents provided demographic and other relevant background information, described the reading programs they used, and identified whether the program was considered "scientifically-based."

Part two, Implementation, is divided into three sections: *Learning Objectives*, *Student Strategies*, and *Teacher Strategies*. Learning Objectives refers to skills the student has demonstrated. Student Strategies refers to specific strategies that the student has used while engaged in reading activities, and Teacher Strategies refers to instructional techniques or processes used by the teacher to help a student achieve an objective and/or employ a strategy. For each section, respondents rated the frequency with which each item was addressed with this student (*Never*, *Sometimes*, *Often, Almost Always*). For those items they had addressed at least once, respondents rated the quality of their delivery (*Poor*, *Fair*, *Good*, *Excellent*) and indicated if the student demonstrated progress due to the teacher's instruction on this item (*No Progress*, *A Little Progress*, *Some Progress*, *A Lot of Progress*).

Part three, Components of Instruction, requires the respondents to read a list of descriptors for the five different components of reading instruction and estimate the percent of instructional time they have spent on each component with the student they selected. Respondents could also select "Other Instruction Related to Reading" and describe any aspects of their instruction not reflected in the descriptors. The sum of percentages for each type of reading instruction was intended to be equal to 100.

In part four, Assessment, respondents identified which types of assessments they conducted for each component of reading and how frequently these assessments occurred. Then teachers collected three different products completed by the student that were representative of their assessment practices. Respondents were guided to choose the products that represented their common assessment practices in reading and were discouraged from selecting assessments based on student performance. They described each

assessment and its purpose, then estimated the number of times they had conducted each assessment with the student since the beginning of the school year.

Scale development. Items were assigned to groups based on the essential component of reading instruction that each item reflected. To evaluate my original assignment of items to groups, seven graduate students completing a course in the theory and practice of reading instruction reviewed the total item pool and assigned each item to a component of reading instruction based on their knowledge of *Put Reading First* (Armbruster, Lehr & Osborn, 2001) and other reading research. Percent of agreement was calculated for the original item pool, and items that resulted in disagreement by at least two evaluators were considered for removal.

Of the original pool of 73 items, 49 items (67%) were assigned to the same component by at least six of the seven raters resulting in 24 items considered for removal. Four of these items were removed, and nine items were revised. The remaining 11 items were assigned to the same component by five of the seven raters and were retained in their original format to preserve the breadth of the item pool.

Three repeat items and two veracity items were placed within the rating scale portion of the instrument to monitor the consistency of teacher responses. The percent of agreement between original and repeat items was calculated for each respondent, and average frequency ratings for the veracity items were also calculated. These results are presented in Chapter 3.

Scoring. Scoring procedures developed for the TISBRI include methods for evaluating rating scale responses and assessment products. Rating scale responses were evaluated across four dimensions: *Items Addressed*, *Frequency*, *Quality of Delivery*, and

Student Progress. Items Addressed was calculated by dividing the number of items addressed by the total number of items within each component to create a percentage score. A score for each component (Phonemic Awareness, Phonics, Fluency, Vocabulary, and Text Comprehension) and one composite score were derived from this procedure.

Frequency scores of 0, 1, 2, and 3 were respectively assigned to Frequency ratings of *Never*, *Sometimes*, *Often* and *Almost Always*. Frequency ratings were averaged for all items within each component of reading instruction (Phonemic Awareness, Phonics, Fluency, Vocabulary, and Text Comprehension) and for the total item set yielding five component scores and one composite score.

Quality of Delivery scores were calculated by assigning scores of 0, 1, 2, and 3 for *Poor*, *Fair*, *Good*, and *Excellent* ratings, respectively, and averaging these scores. Five component scores and one composite score were derived.

Student Progress was calculated by assigning scores of 0, 1, 2, and 3 for respective ratings of *No Progress*, *A Little Progress*, *Some Progress*, and *A Lot of Progress*. These scores were averaged across items resulting in five component scores and one composite score.

The scoring of the Assessment section involved several procedures. First, the Assessments Conducted portion was divided into two parts: breadth of coverage and depth of coverage. Teacher responses to which types of assessments they conducted for each component of instruction (breadth) were scored on a 0 to 3 scale with 0 indicating that no assessments were conducted and three indicating at least three different types of assessments were conducted. This resulted in five component scores and one composite score for breadth

of coverage. Depth of coverage was calculated by first assigning point values from 0 to 5 for ratings of *Never*, *Once*, *Quarterly*, *Monthly*, *Weekly*, and *Daily* respectively. Points were totaled within each component resulting in five component scores and one composite score. The scoring of the Assessments Products involved three steps. First, a rubric was developed to evaluate the products submitted by the teachers (see Appendix C). Each product received five ratings (one for each component of instruction). The rater used a set of descriptors to decide which assessment components were present in the product. The product received a 0 for each component that was not present. For components that were present, the product received scores of 1, 2, or 3 depending on how many measures of each component were illustrated.

Second, two raters independently scored each product based on the rubric described above resulting in five ratings per product, per rater. When differences in ratings occurred that were within one point, the average of the two ratings was used as the final score. When two or more points separated the raters' scores, a third rater reviewed the product. If this rater's scores did not coincide with either of the two previous raters' scores, then the two closest scores were averaged. Finally, a composite score for all three products was calculated by averaging ratings within components.

Additional information was also summarized that did not involve specific scoring methods. *Percent of Instructional Time Devoted to Each Component* presents the respondents' estimates of how their instructional time was divided for the target student. In addition, *Potential Targets for Intervention* were identified from respondents' self-reports. Items that were not addressed or were only addressed once, items that were delivered poorly,

and items that were addressed but the student demonstrated little or no progress were identified for this list.

Procedure

Consultants for the North Carolina Department of Public Instruction, speech/ language pathologists, district superintendents, directors of Exceptional Children's services, curriculum specialists and school principals were contacted to obtain permission to distribute the rating scale. Distribution of the instrument and consent forms occurred in one of three ways: by mail, in person at staff development meetings, or via a school psychologist working within the schools. In all conditions, teachers were provided with a stamped, self-addressed envelope to return their materials to NCSU. In April 2005, 180 packets were distributed via mail to kindergarten through third grade teachers in 13 elementary schools within Rockingham County yielding eight respondents. In November and December of 2005, an additional 33 packets were distributed to teachers who were either enrolled in graduate level courses or attending workshops related to reading instruction. Additionally, 105 packets were given to Wake County school psychologists to distribute to teachers within their schools. This second round of data collection yielded eleven additional respondents. Upon returning the materials to the research staff, teachers' names were entered into a drawing to win a $100.00 gift certificate to a teaching store (round one of data collection) or a $40.00 gift certificate to a local restaurant (round two of data collection). All teachers were given the opportunity to complete the scale a second time to have their names submitted twice into the drawing. Five respondents submitted additional scales.

Teachers' names did not appear on their responses nor were their names included in the database. Teachers were assigned identification numbers to record on their instrument and to use as a signature on their consent forms. Teachers were provided with four options for consent: (a) They could remain completely anonymous and provide only their identification number on all forms submitted to the research staff; (b) They could give consent to participate in the stability study and provide contact information so they could receive a second mailing of the instrument; (c) They could give consent to be entered into the drawing for a chance to win a gift certificate and provide contact information on a separate sheet of paper to be placed in the drawing box; or (d) They could give consent to participate in the stability study and consent to have their names entered into the drawing twice (once for each set of materials they submitted). A list of names and contact information for those teachers who provided this information for the stability study was kept separate from the responses and used for mailing purposes only. Contact information for the drawing was kept in a secure box, and only the winners were contacted. See Appendix D for the consent form and Appendix E for the IRB approval letter.

Teachers were asked to complete the following steps:

1. Select the middle student from their "middle" reading group. (If instructional groups were not used, the teacher selected the student whose reading level or average test score was in the middle of the class distribution.)
2. Complete the TISBRI on the reading instruction delivered to that student since the beginning of the school year.

3. Collect products completed by the student that represent common assessment practices in reading.
4. Remove all identifiers of the student from the products to be submitted.
5. Mail the products and the instrument to me at NC State.

Chapter 3

Analyses and Results

Using the scoring methods described in the previous chapter, mean response values were generated for each respondent on parts two through four of the TISBRI (*Implementation, Components of Instruction,* and *Assessment*). Component and composite scores for each part of the TISBRI will be summarized and descriptive characteristics of the instrument will be highlighted prior to evaluating each of the seven predictions.

Component and Composite Scores

Implementation. Table 4 displays average scores within each grade level for the *Implementation* portion of the instrument. The average percent of items the respondents indicated they had addressed in their reading instruction ranged from a mean of 75% (*SD* = 0.25) in kindergarten to a mean of 96% (*SD* = 0.05) in first and second grade. On average, composites for Frequency ratings ranged from a mean of 1.62 (*SD* = 0.56) in kindergarten to a mean of 2.04 (*SD* = 0.38) in second grade indicating that respondents typically addressed each item *Sometimes* to *Often* in their reading instruction. Quality of Delivery ratings ranged from 1.91[1] in third grade to a mean of 2.34 (*SD* = 0.30) in first grade suggesting that respondents felt their delivery of instruction ranged from *Fair* to *Good* for those items they had addressed. Respondents also felt that, on average, the student made *A Little Progress* to *Some Progress* on those items the respondents had addressed as reflected in Student Progress composite scores that ranged from 1.58 in third grade to a mean of 2.45 (SD = 0.15) in second grade.

[1] Because there was only one respondent for third grade, values for this group have no variance. Therefore, a standard deviation was not reported.

Table 4

Mean Response Values for Teachers Sampled Within Each Grade Level on Part Two of the TISBRI (Implementation)

Scale	K	1	2	3	K-1	2-3
Percent of Items Addressed[a]						
PA	78 (22)	99 (2)	100 (0)	100 __	89 (19)	100 (0)
P	78 (25)	92 (11)	100 (0)	77 __	85 (20)	95 (10)
F	81 (24)	95 (9)	90 (13)	100 __	88 (19)	92 (12)
V	66 (32)	96 (4)	91 (18)	100 __	81 (27)	93 (16)
TC	76 (26)	97 (8)	99 (2)	100 __	87 (21)	99 (2)
C	75 (25)	96 (4)	96 (5)	95 __	85 (20)	96 (5)
Frequency[b]						
PA	1.54 (.56)	2.12 (.34)	2.07 (.13)	1.53 __	1.83 (.54)	1.96 (.27)
P	1.59 (.45)	2.16 (.39)	2.31 (.35)	.92 __	1.88 (.50)	2.03 (.69)
F	1.99 (.75)	2.08 (.41)	1.90 (.55)	2.08 __	2.04 (.58)	1.93 (.49)
V	1.31 (.84)	1.84 (.25)	2.04 (.62)	1.86 __	1.57 (.65)	2.00 (.54)

Table 4 (cont.)

Scale	K	1	2	3	K-1	2-3
Frequency						
TC	1.63 (.70)	1.95 (.39)	2.06 (.60)	2.05 __	1.79 (.57)	2.05 (.52)
C	1.62 (.56)	1.98 (.21)	2.04 (.38)	1.72 __	1.80 (.44)	1.97 (.36)
Quality of Delivery[b]						
PA	2.03 (.25)	2.43 (.41)	2.06 (.25)	1.94 __	2.23 (.38)	2.04 (.22)
P	2.04 (.26)	2.54 (.41)	2.10 (.19)	1.6 __	2.29 (.42)	2.00 (.28)
F	2.46 (.42)	2.41 (.28)	2.23 (.46)	2.42 __	2.43 (.34)	2.27 (.40)
V	2.11 (.48)	2.17 (.37)	2.20 (.49)	1.79 __	2.14 (.41)	2.12 (.46)
TC	2.19 (.46)	2.15 (.38)	2.09 (.60)	2.14 __	2.17 (.40)	2.10 (.52)
C	2.15 (.27)	2.34 (.30)	2.20 (.31)	1.91 __	2.25 (.29)	2.14 (.30)
Student Progress[b]						
PA	2.00 (.53)	2.53 (.29)	2.43 (.17)	1.59 __	2.26 (.50)	2.26 (.40)
P	2.06 (.45)	2.48 (.37)	2.62 (.20)	1.20 __	2.27 (.45)	2.33 (.66)
F	2.24 (.69)	2.53 (.35)	2.42 (.34)	1.50 __	2.38 (.55)	2.24 (.51)

Table 4 (cont.)

Scale	K	1	2	3	K-1	2-3
Student Progress						
V	1.81 (.55)	2.14 (.42)	2.47 (.31)	1.64 __	1.97 (.50)	2.31 (.46)
TC	1.89 (.59)	2.18 (.37)	2.43 (.39)	1.82 __	2.04 (.50)	2.30 (.43)
C	2.00 (.48)	2.35 (.33)	2.45 (.15)	1.58 __	2.18 (.44)	2.28 (.41)

Note. PA = Phonemic Awareness, P = Phonics, F = Fluency, V = Vocabulary, TC = Text Comprehension, and C = Composite. Values in parentheses represent standard deviations. Dashes indicate the standard deviations could not be estimated because no variance occurred in responses (n = 1).

[a]Values represent percentages. [b]Values represent average ratings on scales ranging from 0 to 3.

Components of Instruction. Table 5 displays average scores within each grade level for the percent of instructional time teachers estimated devoting to each component of reading. Total hours of reading instruction per week ranged from 8.00 hours in third grade to a mean of 9.75 (*SD* = 2.87) hours in second grade.

Assessment. Mean response values for the *Assessment* portion of the TISBRI are displayed in Table 6. On average, teachers reported using two to three or more different kinds of assessments for each component of reading instruction as reflected in Assessments Conducted – Breadth composite scores that ranged from 2.77 (*SD* =.60) in kindergarten to 3.00 in first through third grades. The Assessments Conducted – Depth scores are not easily interpretable because ratings from 0 to 5 were summed across nine types of assessment yielding a potential score range of 0 to 45. Composite values ranged from 17.80 (*SD* = 6.46) in kindergarten to 25.70 (*SD* = 5.55) in second grade. Component scores for Assessment Products (based on the scoring rubric designed for this instrument) indicate that on average, none to two measures were used to evaluate student progress in each component.

Descriptive Characteristics of the TISBRI

Veracity items. For the first veracity item, respondents indicated that, on average, they *Never* to *Sometimes* taught students to "outline textbook chapters using formal outlining strategies" (*M* = 0.42, *SD* = 0.84). A much wider range of responses from *Never* to *Often* was found for the second fake item: "provide a full hour of intensive phonics instruction each day" (*M* = 1.42, *SD* = 1.12).

Repeat items. The three repeat items added to the pool provided a measure of consistency across responses. Evidence suggests respondents were moderately consistent

Table 5

Mean Response Values for Teachers Sampled Within Each Grade Level on Part Three of the TISBRI (Components of Instruction)

Scale	K	1	2	3	K-1	2-3
Percent of Instructional Time						
PA	23 (10)	19 (9)	21 (13)	15 __	21 (10)	20 (12)
P	23 (8)	19 (9)	19 (12)	5 __	21 (8)	16 (12)
F	13 (7)	16 (6)	14 (5)	25 __	15 (6)	16 (7)
V	15 (7)	17 (5)	16 (5)	10 __	16 (6)	15 (5)
TC	17 (4)	19 (9)	30 (16)	45 __	18 (7)	33 (16)
Total Hours	9.64 (6.87)	8.64 (3.38)	9.75 (2.87)	8.00 __	9.14 (5.23)	9.40 (2.61)

Note. PA = Phonemic Awareness, P = Phonics, F = Fluency, V = Vocabulary, and TC = Text Comprehension. Values in parentheses represent standard deviations. Dashes indicate the standard deviations could not be estimated because there was no variance in response (n = 1).

Table 6

Mean Response Values for Teachers Sampled Within Each Grade Level on Part Four of the TISBRI (Assessment)

Scale	K	1	2	3	K-1	2-3
Assessments Conducted – Breadth[a]						
PA	3.00 (.00)	3.00 (.00)	3.00 (.00)	3.00 __	3.00 (.00)	3.00 (.00)
P	2.86 (.38)	3.00 (.00)	3.00 (.00)	3.00 __	2.93 (.27)	3.00 (.00)
F	2.57 (1.13)	3.00 (.00)	3.00 (.00)	3.00 __	2.79 (.80)	3.00 (.00)
V	2.57 (1.13)	3.00 (.00)	3.00 (.00)	3.00 __	2.79 (.80)	3.00 (.00)
TC	2.86 (.38)	3.00 (.00)	3.00 (.00)	3.00 __	2.93 (.27)	3.00 (.00)
C	2.77 (.60)	3.00 (.00)	3.00 (.00)	3.00 __	2.89 (.43)	3.00 (.00)
Assessments Conducted – Depth[b]						
PA	19.14 (5.87)	23.14 (7.34)	24.00 (8.04)	12.00 __	21.14 (6.71)	21.60 (8.79)
P	18.86 (5.87)	23.43 (7.55)	24.75 (7.09)	13.00 __	21.14 (6.92)	22.40 (8.08)
F	16.86 (8.28)	21.14 (9.49)	24.00 (2.94)	23.00 __	19.00 (8.84)	23.80 (2.59)
V	16.00 (7.26)	23.57 (8.83)	26.75 (5.32)	28.00 __	19.79 (8.70)	27.00 (4.64)

Table 6 (cont.)

Scale	K	1	2	3	K-1	2-3
Assessments Conducted – Depth						
TC	18.14 (6.23)	22.71 (8.08)	29.00 (7.79)	30.00 __	20.43 (7.32)	29.20 (6.76)
C	17.80 (6.46)	22.80 (8.14)	25.70 (5.55)	21.20 __	20.30 (7.52)	24.80 (5.21)
Assessment Products[a]						
PA	.77 (.27)	.18 (.31)	.04 (.09)	.00 __	.45 (.42)	.03 (.08)
P	1.51 (.64)	.71 (.68)	.71 (.58)	.50 __	1.08 (.76)	.67 (.51)
F	.19 (.22)	.32 (.27)	.31 (.24)	.00 __	.26 (.24)	.25 (.25)
V	.28 (.44)	.26 (.29)	.33 (.47)	1.00 __	.27 (.35)	.47 (.51)
TC	.50 (.76)	1.25 (.60)	1.29 (.34)	2.00 __	.90 (.76)	1.43 (.44)

Note. PA = Phonemic Awareness, P = Phonics, F = Fluency, V = Vocabulary, and TC = Text Comprehension. Values in parentheses represent standard deviations. Dashes indicate the standard deviations could not be estimated because there was no variance in response (n = 1).

[a]Values represent average ratings on scales ranging from 0 to 3. [b]Potential value range is from 0 to 45.

when completing the TISBRI. Seven of the nineteen respondents (37%) had perfect agreement between their first and second responses across all three item sets. An additional seven respondents (37%) had perfect-agreement for two of the items and were within one point on the third item.

Potential targets for intervention. Items that at least 6 respondents (32%) indicated had not been addressed in their instruction are displayed in Table 7. Also displayed are items that received Student Progress ratings of zero or one by at least 6 respondents indicating that the respondents had addressed the item in their instruction but the student had made little or no progress. No Quality of Delivery items are displayed in Table 7 because respondents rarely indicated that their delivery of instruction was poor.

Psychometric Properties of the TISBRI

To evaluate the psychometric properties of the TISBRI, seven predictions were tested. Each prediction and the conditions necessary to support the prediction will be reviewed before describing the results. The first prediction was tested using the original item values for each respondent. The remaining predictions were tested using the component and composite scores described at the beginning of this chapter.

Prediction 1. The final item pool for the TISBRI will have appropriate psychometric characteristics.

(a) Items will be highly correlated with other items related to the same component. Evidence to support this condition would include coefficient alphas of at least .80 within each component.

Table 7

Potential Targets for Intervention

Item		% of Respondents
	Items Not Addressed[a]	
SS01F.	Listen to recorded readings of a text while following along.	32
SS04V.	Use a dictionary and other reference aids to define new words.	32
SS20F.	Build fluency by reading from a script to rehearse and perform a play.	37
SS24TC.	Work with a partner or in a small group to understand content-area texts.	32
TS30V.	Teach root words (words from other languages that are the origin of . . .	42
	Limited Student Progress[b]	
SS16TC.	Look forward in the text to see if new information may resolve the difficulty.	32
SS19TC.	Look back through the text to find breakdowns in comprehension.	32

Note. SS = Student Strategy, TS = Teacher Strategy, F = Fluency, V = Vocabulary, and TC = Text Comprehension.

[a]List includes items that respondents indicated not addressing in their instruction. [b]List includes items that were addressed by the respondents and received Student Progress ratings of zero (No Progress) or one (A Little Progress).

(b) Items will correlate with composites for each component and the total TISBRI composite. Evidence to support this condition would include item-total correlations greater than .30 for component scale composites and for the total TISBRI composite.

(c) The item pool will have a simple factor structure that corresponds to the underlying component structure of the TISBRI. Meeting a majority of the following decision rules would provide support for this condition: exploratory factor analysis would yield only five factors with Eigenvalues > 1.0, the "bend" in a scree plot of Eigenvalues would occur at or near the fifth factor, and each item in the final pool would have loadings of at least .40 on the intended factor and less than .30 on unintended factors. (Note: this prediction was not tested due to insufficient sample size.)

Due to small sample size and insufficient distribution of respondents across grade levels, item-analyses were used only to identify those items which might be considered for removal from the item pool if similar results were found upon further data collection and analysis. The item pool consisted of 76 items each assigned to one of the five components of reading instruction, 3 repeat items, and 2 veracity items. Frequency ratings were used for all item-analyses. Three missing values (less than 1% of the total number of values) across two respondents were replaced with the mean frequency rating for each respondent to preserve the sample size when calculating coefficient alphas. Item-analyses for Quality of Delivery and Student Progress ratings could not be determined because seventeen respondents omitted at least one item from each of the scales. Three of these respondents omitted 28 or more items (35% of the items in the scale). (If the respondent indicated that an item had not been

addressed, the respondent did not have to complete the Quality of Delivery and Student Progress scales for that item.)

The first condition for Prediction 1 was supported by coefficient alphas of at least .80 for all five components indicating that items were highly correlated with other items related to the same component (see Table 8). The second condition was fulfilled by item-total correlations greater than .30 for component scale composites and for the total frequency composite for 65 of the items (80%). See Table 9 for a list of items that did not meet this condition and have been considered for removal.

Prediction 2. The TISBRI will be a relatively stable measure of teachers' self-reports. Correlation coefficients for test-retest stability on Items Addressed, Frequency, Quality of Delivery, and Student Progress that are greater than .80 would provide strong evidence to support this prediction.

Correlations were calculated between scores for a subset of teachers (n = 5) who completed the scale again one to three weeks later. (They were not required to resubmit assessment product data on the second time.) Due to the small sample size, Spearman's rho was used instead of the Pearson product moment correlation to determine test-retest stability. Five component scores and one composite score were evaluated for each scale (Items Addressed, Frequency, Quality of Delivery, and Student Progress) yielding twenty-four measures of test-retest stability.

Partial support was found for the stability of the TISBRI over a one to three week period. See Table 10 for all test-retest correlations. Six of the twenty-four scores evaluated

Table 8

Internal Consistency of the TISBRI

Scale	Cronbach's Alpha	*N* of Items
Frequency		
Phonemic Awareness	.86	17
Phonics	.83	13
Fluency	.82	12
Vocabulary	.93	14
Text Comprehension	.93	23
Composite	.97	81[a]

Note. Values based on 19 cases.

[a]Includes two veracity items not assigned to a component.

Table 9

Items Considered for Removal from the TISBRI

Item	Item-Total Correlation	Cronbach's Alpha if Item Deleted[a]
LO19PA. Identify and make oral rhymes.[b]	.30	.97
SS17F. Read along as part of a group of students with you or . . .	.18	.97
TS01PA. Focus instruction on one or two types of phoneme . . .[b]	-.36	.97
TS02PA. Make explicit connections between phonemic . . .[b]	.25	.97
TS04P. Use a phonics program that has carefully selected . . .	.09	.97
TS07TC. Model how to apply a particular comprehension . . .	.20	.97
TS12P. Use a phonics program that provides precise directions . . .	.16	.97
TS17PA. Deliver phonemic awareness instruction in a small . . .[b]	.04	.97
TS26P. Help student understand why he/she is learning the . . .	.30	.97
TS28P. Help student apply what he/she learns about sounds	.29	.97
TS36P. Teach letter shapes and names.[b]	-.06	.97
TS37F. Read aloud from a big book showing students where . . .[b]	-.08	.97

Note. The values represent internal consistency measures for the Frequency scale. LO = Learning Objective, SS = Student Strategy, TS = Teacher Strategy, PA = Phonemic Awareness, F = Fluency, P = Phonics, TC = Text Comprehension.

[a]Cronbach's Alpha with all items included = .98. [b]Item also met removal criteria for within component values.

Table 10

Test-retest Stability of the TISBRI

Component	Items Addressed	Frequency	Quality of Delivery	Student Progress
Phonemic Awareness	.50	.67	.21	.90* (.09-.99)[a]
Phonics	.79	.15	-.24	1.00**
Fluency	.76	.90* (.09-.99)[a]	.21	.67
Vocabulary	.95* (.42-.1.00)[a]	.30	.29	.70
Text Comprehension	1.00**	.05	.15	.70
Composite	.90* (.09-.99)[a]	.00	.80	.70

Note. $n = 5$. Values represent Spearman's *rho* correlations.

[a]Values represent 95% confidence intervals for significant Spearman's *rho* correlations.

*$p < .05$. **$p < .01$

(25%) met criteria for test-retest stability (Spearman's *rho* greater than .80). The rank order of respondents' scores remained consistent across administrations for Items Addressed in Text Comprehension and Student Progress in Phonics resulting in perfect Spearman's *rho* correlations. Confidence intervals for the remaining four significant correlations are displayed in parentheses in Table 10.

Prediction 3. The rubric developed to score the assessment products will demonstrate acceptable inter-rater agreement. This prediction would be adequately supported by at least 80% agreement between the two raters and a correlation coefficient of .80 or greater for inter-rater agreement. In addition, a weighted coefficient kappa of .60 or greater would provide strong evidence for the reliability of the rubric. (Note: coefficient kappa was not calculated due to the insufficient distribution of values across raters on the rubric scores.)

Two raters scored 42 assessment products from 17 respondents using the rubric developed for this instrument. One respondent did not submit any assessment products, and a second respondent submitted assessment products that were unrelated to reading instruction. Each product received five scores (one for each component of reading). The rubric developed to score the assessment products demonstrated acceptable inter-rater agreement for two of the five reading components, Fluency and Text Comprehension, as demonstrated by correlation coefficients of .81 and .92 respectively ($p < .01$). The 95% confidence interval for Fluency ranged from .67 to .89, which included the prediction of $r \geq .80$. The 95% confidence interval for Text Comprehension remained above this level, ranging from .86 to .96. Correlation coefficients for the remaining components were moderate (Phonemic awareness, .74; Phonics, .67; Vocabulary, .57) and statistically

significant ($p < .01$)). Only 15 of the 42 products that were scored (36%) had perfect agreement between the two raters. Six additional products had near perfect agreement 14%), which meant that 50% of the products elicited scores with less than near agreement between raters.

Prediction 4. Emphasis on the five components of reading instruction will shift from phonemic awareness, phonics, and fluency instruction in earlier grades to emphasis on vocabulary and text comprehension instruction in later grades. Evidence to support this prediction would manifest in the form of significant interaction effects between grade level (K-1 and 2-3) and the five components of reading on multiple dependent measures (i.e., percent of Items Addressed, Frequency of instruction, Quality of Delivery, and Percent of Instructional Time Devoted to Each Component) based on a MANCOVA at alpha = .05. [Should the interaction be significant, the eta squared (proportion of explained variance) will be calculated. Strong evidence would be indicated by a medium or higher effect size (i.e., > .20). Post hoc analyses using Bonferroni's method will be calculated to identify the nature of the interaction between grade level and reading components.] Due to the small sample size, kindergarten and first grade teacher responses were combined into a single group, as were second and third grade teacher responses.

Limited support was found for the prediction that emphasis on instruction would shift from phonemic awareness, phonics, and fluency instruction in the earlier grades to emphasis on vocabulary and text comprehension instruction in the later grades. A significant interaction effect between grade level and the five components of reading was found on only one of the four dependent measures: teachers' estimates of Percent of Instructional Time

Devoted to Each Component ($F(4, 95) = 3.03$, $p = .02$)[2]. Although significant, the interaction explained only 13% of the variability in the model (partial $\eta^2 = .13$). An insufficient number of groups for grade level prevented post hoc analyses to identify the nature of this interaction. However, upon visual inspection of the mean response values for each group (See Table 5), K-1 teachers estimated spending more time on phonics instruction than 2-3 teachers estimated, and 2-3 teachers estimated spending more time on text comprehension instruction than K-1 teachers estimated suggesting the nature of the interaction may be in line with my prediction for these two components of instruction.

Prediction 5. Component scores addressing the enacted curriculum for the teacher self-report portion of the TISBRI will correlate positively with component scores of the assessed curriculum. This prediction would be supported by positive correlations of the Frequency component scores with Assessments Conducted scores and Assessment Products scores. These correlations should be significant at alpha = .05.

Evidence partially supports the positive and significant correlations of one of the scales representing the assessed curriculum (Assessments Conducted but not Assessment Products) with a scale representing the enacted curriculum (Frequency). See Table 11 for all correlations. Assessments Conducted - Breadth scores for Fluency, Vocabulary, and Text Comprehension correlated moderately with their corresponding component scores within the Frequency scale (.65, .61, and .59 respectively) ($p < .05$). However, Breadth scores correlated just as well with Frequency ratings across components as they did with Frequency ratings within components. These correlations were likely due to a lack of variability in responses for the Breadth subscale. (Eighteen of the nineteen respondents indicated using at

[2] $N = 95$ for this analysis because each respondent was treated as five separate cases (one for each component).

Table 11

Correlations Among Measures of the Enacted Curriculum (Frequency Subscales) and Measures of the Assessed Curriculum (Assessment Subscales)

	Frequency Subscales				
	Phonemic Awareness	Phonics	Fluency	Vocab.	Text Comp.
Assessments Conducted[a]					
Phonemic Awareness	—[b] (.65**)	—[b] (.54*)	—[b] (.44)	—[b] (.46*)	—[b] (.53*)
Phonics	.68** (.68)	.34 (.55*)	.65** (.40)	.61** (.47*)	.59** (.53*)
Fluency	.68** (.70**)	.34 (.45)	.65** (.59**)	.61** (.65**)	.59** (.69**)
Vocabulary	.68** (.69**)	.34 (.46)	.65** (.53*)	.61** (.63**)	.59** (.63**)
Text Comprehension	.68** (.67**)	.34 (.40)	.65** (.54*)	.61** (.66**)	.59** (.70**)
Assessment Products					
Phonemic Awareness	-.56*	-.56*	-.24	-.66**	-.43
Phonics	-.30	-.14	-.02	-.26	-.24
Fluency	.18	.17	-.09	.06	.02
Vocabulary	.17	-.01	.06	-.01	.03
Text Comprehension	.31	-.04	.29	.45	.28

[a]Top value represents Breadth; bottom value in parentheses represents Depth. [b]Value could not be computed because no variation occurred for this subscale. *$p < .05$. **$p < .01$

least three different assessment methods for each component of instruction resulting in ratings of 3.00 for each component across respondents.) Assessments Conducted – Depth scores for all five components correlated moderately with their corresponding component scores within the Frequency scale (.65 for Phonemic Awareness, .55 for Phonics, .59 for Fluency, .63 for Vocabulary, and .70 for Text Comprehension) ($p < .05$). Similar to Breadth scores, Depth scores often correlated just as well with Frequency ratings across components as they did with Frequency ratings within components suggesting limited variability in responses.

Component scores within the Frequency scale did not correlate positively and significantly with component scores within the Assessment Products scale. Contrary to expectation, the Phonemic Awareness scores for the Assessment Products were negatively correlated with their corresponding scores within the Frequency scale (-.56, $p < .05$). (Very few of the assessment products that were scored reflected examples of phonemic awareness assessments.) Table 12 displays the 95% confidence intervals for the significant correlations found between Frequency and Assessment subscales (i.e., within components). Given the small sample size, these intervals provide more cautious estimates of the true relationships between these scales.

Prediction 6. Scale scores addressing the enacted curriculum will correlate positively with scale scores addressing the learned curriculum. This prediction would be supported by positive correlations of the Frequency and Quality of Delivery scale scores with the Student Progress scale scores. These correlations should be significant at alpha = .05.

Table 12

Confidence Intervals for Significant Correlations Among the Frequency and Assessment Subscales

	Frequency Subscales				
	Phonemic Awareness	Phonics	Fluency	Vocab.	Text Comp.
Assessments Conducted - Breadth					
Fluency			.28-.85		
Vocabulary				.22 - .83	
Text Comprehension					.19 - .82
Assessments Conducted – Depth					
Phonemic Awareness	.28 - .85				
Phonics		.13 - .80			
Fluency			.19 - .82		
Vocabulary				.25 - .84	
Text Comprehension					.36 - .88
Assessment Products					
Phonemic Awareness	-.81- -.14				

Note. Values represent 95% confidence intervals.

Limited support was found for correlations between scales reflecting the enacted curriculum (Frequency and Quality of Delivery) and the scale reflecting the learned curriculum (Student Progress). See Table 13 for all correlations. The Fluency and Text Comprehension component scores for Frequency correlated moderately and significantly with their corresponding Student Progress component scores (.46 and .51 respectively; $p < .05$). The true correlation of these scales is likely to range from .01 to .76 for Fluency and from .07 to .78 for Text Comprehension (95% confidence interval). No positive and significant correlations were found between Quality of Delivery component scores and Student Progress component scores.

Prediction 7. TISBRI component scores addressing the enacted curriculum will correlate positively with teachers' estimates of proportion of time devoted to instruction in each of the components. Evidence to support this prediction would manifest in the form of significant positive correlations between percent of Items Addressed and Percent of Instructional Time Devoted to Each Component. These correlations should be significant at alpha = .05.

Partial support was found for the correlation of scores addressing the enacted curriculum (Percent of Items Addressed) with teachers' estimates of proportion of time devoted to instruction. See Table 14 for all correlations. Fluency and Vocabulary component scores for Percent of Items Addressed correlated moderately and significantly with teachers' estimates (.46 and .53 respectively, $p < .05$). The true correlation of these scales is likely to range from .01 to .76 for Fluency and from .10 to .79 for Vocabulary (95% confidence interval). A negative correlation was found for Phonemic Awareness component scores. As

Table 13

Correlations Among Measures of the Enacted Curriculum (Frequency and Quality of Delivery Subscales) and Measures of the Learned Curriculum (Student Progress Subscales)

	Student Progress Subscales				
	Phonemic Awareness	Phonics	Fluency	Vocab.	Text Comp.
Frequency Subscales					
Phonemic Awareness	.23	.21	.45	.36	.44
Phonics	.27	.38	.29	.23	.21
Fluency	-.18	-.15	.46*	.28	.35
Vocabulary	-.08	-.01	.33	.43	.44
Text Comprehension	-.13	-.12	.37	.44	.51*
Quality of Delivery Subscales					
Phonemic Awareness	.32	.31	.39	.33	.28
Phonics	.34	.39	.38	.30	.21
Fluency	-.23	-.17	.33	.24	.22
Vocabulary	-.13	-.01	.23	.36	.25
Text Comprehension	-.22	-.12	.31	.41	.38

*$p < .05$. **$p < .01$

Table 14

Correlations Among Percent of Items Addressed and Teachers' Estimates of Their Distribution of Instructional Time

	Percent of Instructional Time				
	Phonemic Awareness	Phonics	Fluency	Vocab.	Text Comp.
Percent of Items Addressed					
Phonemic Awareness	-.52*	-.34	.49*	.47*	.31
Phonics	-.28	-.36	.16	.39	.15
Fluency	-.36	-.28	.46*	.44	.21
Vocabulary	-.61**	-.38	.59**	.53*	.36
Text Comprehension	-.42	-.31	.37	.49*	.26

*$p < .05$. **$p < .01$

the percent of Phonemic Awareness items that were addressed increased, the percent of time teachers estimated devoting to Phonemic Awareness instruction decreased (-.52, $p < .05$). The 95% confidence interval for this estimate is also quite wide, ranging from -.79 to -.09.

Evidence provided limited to partial support for each of the seven predictions regarding the psychometric properties of the TISBRI. Interpretations of these results and a discussion of limitations and directions for future use of the TISBRI will be presented in the following chapter.

Chapter 4

Discussion

This study was an attempt to design a meaningful and useful instrument for measuring a student's previous opportunities to learn from explicit and systematic instruction in the five essential components of reading. The *Special Rule for Eligibility Determination* in the Individuals with Disabilities Education Act of 2004 provided the primary rationale for developing such an instrument. The specific goal of this study was to determine the instrument's psychometric properties. In the following sections, I will evaluate the progress made toward this goal, discuss the limitations of the present study, and address next steps for measuring opportunities to learn.

Psychometric Properties of the TISBRI

Due to a small sample size and an insufficient distribution of respondents across grade levels, limited insight was gained into the psychometric properties of the TISBRI. Each area of analysis will be discussed in light of these limitations.

Item analyses and reliability. Item-analyses suggest strong internal consistency estimates; however, these estimates are based on a respondent sample that included only four second grade teachers and one third grade teacher. A greater distribution of respondents across grade levels would likely increase the diversity of items addressed as well as the variability of frequency ratings which may, in turn, influence the internal consistency estimates for the TISBRI. Additionally, items currently considered for removal may become essential to the TISBRI if they reflect learning objectives or strategies that are more commonly addressed by second and third grade teachers.

Stability. With only five respondents completing the instrument a second time, Spearman's *rho* correlations were used in place of Pearson's *r* correlations to evaluate test-retest stability. Spearman's *rho* correlations examine the relationship between the rank order of values on a given variable and not the values themselves, and, unlike Pearson's *r,* this method does not assume random error distribution. However, given the small sample size, wide confidence intervals (see Table 10), and limited statistical power for these estimates, I recommend these results be interpreted in light of the inherent lack of precision within the estimates (i.e., they should be largely ignored).

Inter-rater Agreement. Based on an evaluation of scores for 42 student work samples (i.e., two or three samples per student), evidence did not support acceptable inter-rater agreement. This lack of support may be due to the design of the rubric itself, or to insufficient training of the raters. Raters found it challenging to decipher between examples of phonemic awareness assessments and phonics assessments given the definitions provided for them on their scoring rubric and their knowledge of differences between these two areas of instruction. Running records were the primary examples provided for fluency assessments, making this component easy to identify. Examples of text comprehension assessments also appeared more clear-cut than examples of other kinds of assessments. These reasons may explain why inter-rater agreement for fluency and text comprehension met acceptable levels while inter-rater agreement for the remaining components remained in the moderate range. Given that scores for 50% of the products did not meet an acceptable level of inter-rater agreement, revisions should be made to the rubric and to methods for training the raters.

Discriminant Validity. To perform a MANCOVA, each respondent was represented as five separate cases (one case for each component of reading) in the data set boosting the sample size from 19 to 95. Although this inflation of sample size allowed for the evaluation of interaction effects between grade level and each component of instruction across the four dependent measures, it violated statistical principles of independence between observations. It is unlikely that teachers' responses for a given component are completely independent from their responses to other components. Therefore, estimates of significance for the interaction of grade level and component on teachers' estimates of instructional time are likely to be inflated given the actual number of respondents.

Concurrent Validity. Due to small sample size, correlational studies of scales within the TISBRI provided limited insight into the relationship of curricular areas represented within the instrument. Although significant correlations fell within the moderate range, in general, 95% confidence intervals for these estimates stretched from low (.20) to high (.80) levels of correlation. Although results suggest that some relationships do exist between scales, such a wide band of confidence prevents meaningful interpretation of these relationships.

Limitations of the Present Study

Recruitment of teachers was the single greatest challenge to this study and led to its most significant limitation: insufficient sample size. A return rate of 6% suggests that there was not enough incentive to complete the instrument given the time and effort required. When asked why they did not complete the instrument, teachers in the first round of data collection (April and May, 2005) had two common responses: "It is too long" and "It is just

the wrong time of year; I'm too busy." The second round of data collection, which took place in December, 2005 and January, 2006, had a similar return rate to the first round suggesting the instrument's length may be of greater concern than time of year for data collection.

A related limitation is the lack of random selection of teachers for participation in the study. After initial efforts to solicit help with recruitment from North Carolina's Department of Public Instruction proved unfruitful, I began contacting school districts directly and recruiting teachers via mail, via the school psychologists working at their schools, and at various workshops and graduate courses. Those who chose to participate may stand out from a random sample of teachers in terms of experience, knowledge of reading instruction, interest in scientifically based practices, or some other distinguishing variable. For example, four of the respondents were recruited from workshops that focused on improving reading instruction which may have given them a greater knowledge base to draw from in their instruction than would be typical of teachers who did not attend the workshops. Although the student sample was ethnically diverse and 63% male, the respondent sample was predominantly Caucasian and female. Improving the diversity of the respondent sample across all demographics, would strengthen the conclusions that could be drawn regarding the TISBRI's psychometric properties. In addition, to investigate grade-level differences reliably, participants should be recruited via stratified random sampling to ensure that respondents in each grade level have comparable demographic characteristics.

This study was also limited by its methods for evaluating the concurrent validity of the TISBRI. The collection of assessment products provided the primary method for

determining the validity of teacher self-report. The results of this study indicate that the rubric developed to evaluate these products will need to be revised for future data collection, and raters may need to receive additional training before scoring the products. In addition, procedures for collecting assessment products may also need to be changed to limit the influence of teacher judgment. For the present study, teachers chose which products to submit and self-reported the breadth and depth of their assessment practices. To eliminate the influence of teacher judgment in the evaluation of assessment products on future data collection efforts, teachers could submit an entire student assessment portfolio (assuming there is one) for evaluation by independent raters. Contingent upon the development of a reliable rubric for evaluating these portfolios, this method may provide a more sound measure of concurrent validity between the Frequency and Assessment scales on the TISBRI.

Next Steps for Measuring Opportunities to Learn

The results of this study raise more questions than answers. Two questions will be addressed with respect to future directions for research: Is the TISBRI just too long? And, is teacher self-report a feasible method for evaluating an individual student's opportunities to learn?

Is the TISBRI just too long? Judging from feedback and poor response rate, the answer to this question is "Yes." However, having a large initial item pool allows for the removal of items that do not significantly impact the internal consistency of the instrument. Removing such items can only be done if a sufficient sample size is obtained to run the appropriate analyses. Therefore increased incentives may be necessary to recruit participants to complete the TISBRI in its current version. The TISBRI could also be revised to focus

exclusively on the frequency of instructional practices and the breadth and depth of assessment practices by omitting the Quality of Delivery and Student Progress scales. Additionally, research examining indicators of scientifically based reading instruction may prove beneficial to streamlining the TISBRI. For example, research may find that teachers who use phonics programs that provide precise directions for the teaching of letter-sound relationships also address several other objectives and strategies related to phonics instruction that teachers who do not use these programs consequently omit (e.g. segmenting written words into onsets and rimes, helping students understand why they are learning the relationships between letters and sounds). Identifying key indicators may reduce the number of objectives and strategies that would need to be assessed on the TISBRI, yielding greater parsimony of measurement.

Is teacher self-report a feasible method for evaluating an individual student's opportunities to learn? Maybe. If key indicators of scientifically based reading instruction can be identified through research and a parsimonious scale developed, teacher self-report is the most efficient method for measuring an individual student's opportunities to learn. However, this method should be compared to other methods of assessment for an individual student including teacher interviews, direct observations, and permanent product reviews to determine which methods produce the most reliable and valid estimates of opportunities to learn from scientifically based reading instruction.

Two problems are likely to remain problematic for self-report instruments, no matter how well constructed. The first is that teachers may simply be unable to evaluate the quality of instruction they provide. That is, delivering and managing instruction may exhaust

available cognitive monitoring resources, leaving teachers unable to accurately self-assess the quality of their activities. The second problem is that even accurate reports cannot reliably indicate how a student might respond to other instruction, or even the same instruction delivered by a different person. Response to intervention approaches to establishing lack of opportunity to learn cannot directly demonstrate whether a student actually received appropriate instruction prior to intervention, but they can demonstrate whether a student responds to a particular intervention delivered by a particular person.

Conclusions

Limited insight was gained into the psychometric properties of the TISBRI. The instrument demonstrated strong internal consistency, but no conclusions could be drawn for the stability of the instrument, the relationships between scales in the instrument, or its ability to detect grade level differences in response patterns. Small sample size, lack of random selection, and an unreliable method for evaluating the assessment products all contributed to the lack of meaningful findings.

Greater incentives may be needed to recruit participants to complete the TISBRI in its current version. If a sufficient sample size could be obtained, the TISBRI may be stream-lined to create a more parsimonious measure of an individual student's prior opportunities to learn. This has utility not only for addressing the *Special Rule for Eligibility Determination* but also for expediting more intense services to students who have received adequate opportunities to learn, tailoring interventions to meet the specific needs of an individual student, and enabling teachers to evaluate and improve upon their instruction. When determining appropriate uses for the TISBRI, consideration must be given to the potential

rewards and sanctions associated with its use. Using the TISBRI to help determine eligibility for special education services might cause teachers to intentionally misrepresent aspects of their instruction. Therefore, the collection of assessment products and other data may become critical to preserving the validity of teacher self-report.

Upon further development, the TISBRI could have implications for use by school psychologists. This instrument could provide insight into the breadth and depth of learning opportunities a student receives in the general education environment. Such information could help school psychologists determine if deficits in academic achievement result from lack of exposure to appropriate instruction or an innate disability. Identifying the root of the problem may lead to better recommendations for services. School psychologists could also use the TISBRI as a tool to consult with teachers regarding their instructional practices in reading. Promoting the TISBRI as a self-evaluation tool for teachers would allow school psychologists to influence the learning opportunities of all children within a classroom not just the opportunities of a single child referred for services.

References

Armbruster, B. B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read.* Washington, DC: US Department of Education.

Ball, D. L., Camburn, E., Correnti, R., Phelps, G., & Wallace, R. (1999). *New tools for research on instruction and instructional policy: A web-based teacher log* (Document W-99-2). Seattle: University of Washington, Center for the Study of Teaching and Policy.

Blank, R. K. (2002). Using surveys of enacted curriculum to advance evaluation of instruction in relation to standards. *Peabody Journal of Education*, *77*, 86-121.

Bradley-Johnson, S., & Morgan, S. K. (2002). Best practices in planning effective instruction for students who are visually impaired. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 1405-1418). Bethesda, MD: National Association of School Psychologists.

Burstein, L., McDonnel, L. M., Van Winkle, J., Ormseth, T., Mirocha, J., & Guitton, G. (1995). *Validating National Curriculum Indicators. Draft* (Report No. DRU-1086-NSF). Santa Monica, CA: National Science Foundation. (ERIC Document Reproduction Service No. ED387869)

Cunningham, J. W. (2001). The National Reading Panel report [A review.] In R. L. Allington (Ed.), *Big brother and the national reading curriculum* (pp. 49-74).

Dreeben, R., & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, *61*, 129-142.

Elementary and Secondary Education Act of 1965, Pub. L. 107-110, Sec. 1208 (3,6). Retrieved on February 11, 2005 from http://thomas.loc.gov/

Engelmann, S., & Bruner, E. C. (1995). *Reading Mastery II: Teacher's guide*. Worthington, OH: SRA Macmillan/McGraw-Hill.

Floden, R. E. (2003). The measurement of opportunity to learn. In A.C. Porter & A. Gamoran (Eds.), *Methodological Advances in Cross-National Surveys of Educational Achievement* (pp. 231-266). Washington, DC: National Academy Press.

Foorman, B. R., Chen, D., Carlson, C., Moats, L., Francis, D. J., & Fletcher, J. (2003). The necessity of the alphabetic principle to phonemic awareness instruction. *Reading and Writing: An Interdisciplinary Journal*, *16*, 289-324.

Garan, E. M. (2001). Beyond the smoke and mirrors: A critique of the National Reading Panel report on phonics. In R. L. Allington (Ed.), *Big brother and the national reading curriculum* (pp. 90-111).

Gregory, R. J. (2000). *Psychological testing: History, principles, and applications*. Boston, MA: Allyn & Bacon.

Herman, J. L., & Klein, D. C. D. (1996). Evaluating equity in alternative assessment: An illustration of opportunity-to-learn issues. *The Journal of Educational Research*, *89*, 246-257.

Herman, J. L., Klein, D. C. D., & Abedi, J. (2000). Assessing students' opportunity to learn: Teacher and student perspectives. *Educational Measurement: Issues and Practice*, *19*, 16-24.

Houghton Mifflin Company (2002). *Houghton Mifflin Reading: The Nation's Choice*. Retrieved February 11, 2005, from http://www.eduplace.com/marketing/nc/

Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No: 108-446, Sec. 614(b)(5)(A). Retrieved on February 11, 2005 from http://thomas.loc.gov/

Krashen (2001). More smoke and mirrors: A critique of the National Reading Panel report on fluency. In R. L. Allington (Ed.), *Big brother and the national reading curriculum* (pp. 112-124).

Mayer, D.P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*, 29-45.

Mickelson, R. A. (2001). Subverting Swann: First- and second-generation segregation in the Charlotte-Mecklenburg schools. *American Educational Research Journal*, *38*, 215-252.

National Reading Panel (2000a). *Report of the National Reading Panel: Teaching children to read—An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (Reports of the Subgroups).* Washington, DC: US Department of Health and Human Services.

National Reading Panel (2000b). *Report of the National Reading Panel: Teaching children to read—An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (Summary).* Washington, DC: US Department of Health and Human Services.

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven: Yale University Press. Abstract obtained from ERIC, 1986, Abstract No. ED274749

Oregon Reading First Curriculum Review Panel (2004, March). *Oregon Reading First Center: Review of Comprehensive Reading Programs*. Retrieved February 11, 2005, from http://reading.uoregon.edu/curricula/or_rfc_review_n.php

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, *31*, 3-14.

Porter, A. C., & Smithson, J. L. (2001). *Defining, developing, and using curriculum indicators* (Research Report Series RR-048). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

RAND Reading Study Group (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. Santa Monica, CA: RAND Corporation.

SEC Online (n.d.). *Surveys of the enacted curriculum: State collaborative on assessment and student standards.* Retrieved February 16, 2005, from http://facstaff.wcer.wisc.edu/jsmithson/secWebReference.htm

Scribner, A. P. (2002). Best assessment and intervention practices with second language learners. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 1485-1500). Bethesda, MD: National Association of School Psychologists.

Shanahan, T. (2003). Research-based reading instruction: Myths about the National Reading Panel report. *The Reading Teacher*, *56*, 646-655.

Shapiro, E.S. (1996). *Academic skills problems: Direct assessment and intervention*. New York, NY: Guilford Press.

Shim, M. K., Felner, R. D., Shim, E., & Noonan, N. (2001, April). *Multi-dimensional assessment of classroom instructional practice: A validity study of the Classroom*

Instructional Practice Scale (CIPS). Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Simmons, D. C., & Kame'enui, E. J. (2003). *A consumer's guide to evaluating a core reading program grades K-3: A critical analysis*. Eugene: University of Oregon, Institute for the Development of Educational Achievement.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum – The Reform Up Close study* (Research Report Series RR-31). Brunswick, NJ: Consortium for Policy Research in Education.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Snow-Renner, R. (2001, April). *What is the promise of large-scale classroom practice measures for informing opportunities-to-learn? An example using the Colorado TIMSS.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Stevens, F. I., & Grymes, J. (1993). *Opportunity to learn: Issues of equity for poor and minority students* (Report No. 93-232). Washington, DC: National Center for Education Statistics. (ERIC Document Reproduction Service No. ED356306)

Walker-Hill, M., Block-Pedego, A., Todis, B., & Severson, H. H. (1992). *School Archival Records Search*. Longmont, CA: Sopris West.

Wang, J., & Goldschmidt, P. (1999). Opportunity to learn, language proficiency, and immigrant status effects on mathematics achievement. *Journal of Educational Research*, *93*, 101-111.

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research monograph No.18). Madison, WI: University of Wisconsin-Madison, National Institute for Science Education.

Zimmerman, J., & Brown, C. (2003). Let them eat more than phonics. *Phi Delta Kappan*, *84*, 603-605.

Appendix A

Development of the TISBRI

Development of the TISBRI

Development of the TISBRI involved several steps. First, the content domain was intended to reflect the essential components of reading instruction (phonemic awareness, phonics, fluency, vocabulary, and text comprehension). *Put Reading First* (Armbruster, Lehr & Osborn, 2001) was used to identify specific learning objectives, student strategies, and teacher strategies for each component. Items were randomly placed within each section so that items reflecting the same component (i.e., phonemic awareness, phonics, fluency, vocabulary, or text comprehension) were unlikely to be clumped together. At least ten items were developed for each component. Then measures of frequency, duration, and delivery were developed to assess implementation of these objectives and strategies. Finally, a rubric was designed to measure the alignment of permanent product data with teachers' self-report.

The TISBRI was reviewed for clarity, content, style, and deliverability in a pilot study conducted prior to the current research. The initial version was submitted to four university professors, and two elementary school teachers. The professors included two from School Psychology and two from Education. The teachers included one kindergarten teacher and one first grade teacher. This feedback was used to organize, streamline, and improve the readability of the instrument. After revisions were made, a focus group was used to make further improvements to the instrument. The group included one kindergarten teacher, two first grade teachers, one second grade teacher, and one Title One teacher. Feedback from this discussion was used to modify the measurement scales and clarify the directions for the instrument.

Appendix B

Teacher's Implementation of Scientifically Based Reading Instruction (TISBRI)

Teacher's Implementation of Scientifically Based Reading Instruction

Goal:

This self-report measure has been developed to evaluate the reading instruction you have delivered to a particular student this year. Its content stems from *Put Reading First: The Research Building Blocks for Teaching Children to Read* (Armbruster, Lehr, & Osborn, 2001) and focuses on kindergarten through third grade reading instruction. Eventually this instrument may be used to measure a child's opportunities to learn from scientifically based reading instruction. Currently, data collection is focused on measuring the reliability and validity of the instrument. In addition, your responses will help to streamline future versions of this instrument. All information you provide will remain confidential.

Directions:

Please select one child who is in the middle of your "middle" reading group. If you do not group students for instruction, please select the middle student from your class distribution for reading level or current reading grade. Once you have selected a student, consider the reading instruction you have delivered to this child since the beginning of the year. If this child has not been in your class since the beginning of the year, please select another child of similar ability. Since not all students within a classroom receive exactly the same instruction, your responses should reflect only the reading instruction you have delivered to the specific student selected.

Format:

In the sections to follow, reading instruction is divided into three categories. These categories are **Learning Objectives**, **Student Strategies**, and **Teacher Strategies**. For each category you will be asked to evaluate the amount of time you have devoted to specific elements of instruction, the quality of your delivery, and degree of student progress. Definitions for each category are given below.

Definitions:

Learning Objectives refers to skills the student demonstrates. *(Ex. Break or segment a word into its specific sounds.)*

Student Strategies refers to specific strategies that the student uses while engaged in reading activities. *(Ex. Practice orally rereading text that is at his/her independent reading level.)*

Teacher Strategies refers to instructional techniques or processes used by the teacher to help a student achieve an objective and/or employ a strategy. *(Ex. Model fluent reading by reading aloud to student.)*

Response Codes:

Use the **Response Codes** sheet to guide your responses starting on page 3.

Part 1 of 4 **Background Information**

Student Age: _____ **Grade:** K 1 2 3 **Gender:** M F **Ethnicity:** _____

District: _____ **Retained in any grade?** Yes No Don't know

In general, has the child you selected received **more**, **about the same**, or **less** reading instruction than the majority of the class during this school year? _____

What factors have limited or enhanced the instruction this child has received this year? Check all that apply.

Limitations	Enhancements
☐ Poor attendance	☐ Title 1 reading assistance
☐ Behavior problems	☐ After/before school tutoring in reading
☐	☐ Parent Instruction
☐	☐ Occupational/Speech/Other therapy
☐	☐

Please comment on any of the limitations or enhancements to the reading instruction this student receives.

_____

Respondent Gender: M F **Respondent Ethnicity:** _____ **Degree(s)/Major(s):** _____

How many college or graduate level courses have you taken in reading instruction? _____

How many months have you provided reading instruction for this student? _____

How many years have you been teaching reading at this grade level? _____

How many years have you been teaching reading at **any** grade K through 3? _____

What is the name of the current reading program you use? _____

What is the focus of this program? _____

Is this program considered "scientifically based"? Yes No Don't know

Remember: On the remainder of this scale, only consider the instruction you have delivered to the selected student. Please find the Response Codes sheet to guide your responses on the following pages.

Part 2 of 4 **Implementation: Learning Objectives**

Directions: Answer question A for each item listed below. If you have not addressed the item thus far this year, you do not need to answer questions B and C for this item. See the **Response Codes** sheet to guide your selections.

Question A: How often have you covered this learning objective with this student?	Learning Objectives	Question B: How would you rate your delivery of instruction on this objective for this student?	Question C: How much progress has the student made?
N S O A	Recognize which words in a set of words begin with the same sound. *Which words begin with the /t/ sound: **tot, toy, rat, tub**?*	P F G E	0 1 2 3
N S O A	Convert letter or letter combinations into sounds and then recognizable words. *The letters **s-h** make what sound?*	P F G E	0 1 2 3
N S O A	Analyze letter-sound relationships in previously learned words. *In the word **toy**, which letters make the **/oi/** sound?*	P F G E	0 1 2 3
N S O A	Identify the same sound in different words. ***Boy, ball,** and **bat** all begin with what sound?*	P F G E	0 1 2 3
N S O A	Use parts of word families to identify unknown words in text. *Use words from the **–ay** family to identify the new word – **stay**.*	P F G E	0 1 2 3
N S O A	Isolate and say the first or last sound in a word. *What is the ending sound in **mitt**?*	P F G E	0 1 2 3
N S O A	Make words by writing letters for phonemes. *What letters do you need to make the **/st/** sounds in **stop**?*	P F G E	0 1 2 3
N S O A	Break or segment a spoken word into its separate sounds. *How many sounds are in **small**?*	P F G E	0 1 2 3
N S O A	Identify the word in a set of words that has the "odd" sound. *Which word does not belong: **bus, bark**, or **cap**?*	P F G E	0 1 2 3
N S O A	Recognize letter-sound relationships during the reading of connected text. *Find a word in the paragraph that begins with **t-h**? What sound do those letters make?*	P F G E	0 1 2 3
N S O A	Combine or blend the separate sounds in a word to say the word. *What word is **/c/ /a/ /t/**?*	P F G E	0 1 2 3
N S O A	Segment written words into onsets and rimes. *Break the word **string** into **str** and **ing**. What are the two sound parts?*	P F G E	0 1 2 3
N S O A	Produce the new word that remains when a phoneme is removed from another word. *What is **stale** without the **/s/**?*	P F G E	0 1 2 3

Part 2 of 4 **Implementation: Learning Objectives**

Question A: How often have you covered this learning objective with this student?	**Learning Objectives**	Question B: How would you rate your delivery of instruction on this objective for this student?	Question C: How much progress has the student made?
N S O A	Make a new word by adding a phoneme to an existing word. *What word do you have if you add /s/ to the beginning of* ***mile****?*	P F G E	0 1 2 3
N S O A	Substitute one phoneme for another to make a new word. *The word is* ***can****. Change /n/ to /t/. What's the new word?*	P F G E	0 1 2 3
N S O A	Identify and work with syllables in spoken words.	P F G E	0 1 2 3
N S O A	Break a word into its separate sounds saying each word as he/she taps out or counts it.	P F G E	0 1 2 3
N S O A	Identify and work with onsets and rimes in spoken syllables or one-syllable words.	P F G E	0 1 2 3
N S O A	Identify and make oral rhymes.	P F G E	0 1 2 3

Part 2 of 4

Implementation: Student Strategies

Question A: How often have you modeled and/or taught this strategy?	Student Strategies	Question B: How would you rate your modeling and/or instructing of this strategy?	Question C: How much progress has the student made?
N S O A	Listen to recorded readings of a text while following along (tape-assisted reading).	P F G E	0 1 2 3
N S O A	Read the same passage orally several times.	P F G E	0 1 2 3
N S O A	Work actively with vocabulary over an extended period of time. *(Use 3 vocabulary words from last week in a written paragraph or conversation with a friend.)*	P F G E	0 1 2 3
N S O A	Use a dictionary and other reference aids to define new words.	P F G E	0 1 2 3
N S O A	Identify where in the text the breakdown in comprehension occurs.	P F G E	0 1 2 3
N S O A	Use word parts (prefixes, suffixes, base words, root words) to figure out the meanings of words.	P F G E	0 1 2 3
N S O A	Practice orally rereading text that is at his/her independent reading level (95% accuracy).	P F G E	0 1 2 3
N S O A	Read one-on-one with an adult who models fluent reading and assists the student (student-adult reading).	P F G E	0 1 2 3
N S O A	Look back through the text to find breakdowns in comprehension.	P F G E	0 1 2 3
N S O A	Use context clues to determine word meanings.	P F G E	0 1 2 3
N S O A	Restate the difficult sentence or passage in new words.	P F G E	0 1 2 3
N S O A	Generate comprehension questions and try to answer them.	P F G E	0 1 2 3
N S O A	Use comprehension strategies flexibly as they are needed to assist their comprehension.	P F G E	0 1 2 3
N S O A	Practice rereading a variety of reading materials, including stories, nonfiction, and poetry.	P F G E	0 1 2 3
N S O A	Use context clues to help define new vocabulary.	P F G E	0 1 2 3

Part 2 of 4

Implementation: Student Strategies

Question A: How often have you modeled and/or taught this strategy?	Student Strategies	Question B: How would you rate your modeling and/or instructing of this strategy?	Question C: How much progress has the student made?
N S O A	Look forward in the text to see if new information may resolve the difficulty.	P F G E	0 1 2 3
N S O A	Read along as part of a group of students with you or another fluent adult reader (choral reading).	P F G E	0 1 2 3
N S O A	Take turns reading aloud to another student (partner reading).	P F G E	0 1 2 3
N S O A	Look back through the text to find breakdowns in comprehension.	P F G E	0 1 2 3
N S O A	Build fluency by reading from a script to rehearse and perform a play (reader's theatre).	P F G E	0 1 2 3
N S O A	Identify the elements within a story (setting, character development, problem, solution, climax, etc.).	P F G E	0 1 2 3
N S O A	Outline textbook chapters using formal outlining strategies.	P F G E	0 1 2 3
N S O A	Determine what is important in what he/she is reading by condensing the essential information and putting it into his/her own words.	P F G E	0 1 2 3
N S O A	Work with a partner or in a small group to understand content-area texts.	P F G E	0 1 2 3
N S O A	Form visual images of what he/she is reading.	P F G E	0 1 2 3

Part 2 of 4 **Implementation: Teacher Strategies**

Question A: How often have you used this strategy with this student?	Teacher Strategies	Question B: How would you rate your ability to use this strategy?	Question C: How much progress has the student made?
N S O A	Focus instruction on one or two specific types of phoneme manipulation rather than trying to cover all types of manipulation.	P F G E	0 1 2 3
N S O A	Make explicit connections between phonemic awareness and reading. *(Say the sounds, write the sounds, read the word.)*	P F G E	0 1 2 3
N S O A	Ask comprehension questions where the answers are explicitly stated in the text.	P F G E	0 1 2 3
N S O A	Use a phonics program that has carefully selected and organized letter-sound relationships in a logical instructional sequence.	P F G E	0 1 2 3
N S O A	Provide the student with repeated exposure to vocabulary in many contexts. *(Conversations, social studies and science instruction, writing samples)*	P F G E	0 1 2 3
N S O A	Ask comprehension questions about the reader's prior knowledge or experience as it relates to the text.	P F G E	0 1 2 3
N S O A	Model how to apply a particular comprehension strategy usually by "thinking aloud" while reading a text.	P F G E	0 1 2 3
N S O A	Read aloud to students and discuss the selection before, during, and after reading.	P F G E	0 1 2 3
N S O A	Ask comprehension questions where the answers are implied by information presented in multiple sentences.	P F G E	0 1 2 3
N S O A	Model fluent reading by reading aloud to student.	P F G E	0 1 2 3
N S O A	Provide a full hour of intensive phonics instruction each day.	P F G E	0 1 2 3
N S O A	Use a phonics program that provides precise directions for the teaching of letter-sound relationships.	P F G E	0 1 2 3
N S O A	Provide guidance and assistance to students as they learn how and when to apply a particular comprehension strategy.	P F G E	0 1 2 3
N S O A	Use a graphic organizer to focus student's attention to the structure of the text.	P F G E	0 1 2 3

Part 2 of 4

Implementation: Teacher Strategies

Question A: How often have you used this strategy with this student?	Teacher Strategies	Question B: How would you rate your ability to use this strategy?	Question C: How much progress has the student made?
N S O A	Help students relate new words and concepts to their prior knowledge and experiences.	P F G E	0 1 2 3
N S O A	Model fluent reading by reading aloud to student.	P F G E	0 1 2 3
N S O A	Deliver phonemic awareness instruction in a small-group setting.	P F G E	0 1 2 3
N S O A	Encourage student to develop word consciousness, for example, by noticing an author's choice of words in a text or engaging in word play (puns, palindromes).	P F G E	0 1 2 3
N S O A	Explain to student why the use of a particular strategy helps comprehension.	P F G E	0 1 2 3
N S O A	Use a graphic organizer to help students write a well-organized summary of a text.	P F G E	0 1 2 3
N S O A	Discuss specific words found in an unfamiliar text prior to reading.	P F G E	0 1 2 3
N S O A	Encourage student to read familiar texts outside of class to build fluency.	P F G E	0 1 2 3
N S O A	Use a graphic organizer to visually represent and examine relationships in a text or across texts.	P F G E	0 1 2 3
N S O A	Use a phonics program that has carefully selected and organized letter-sound relationships in a logical instructional sequence.	P F G E	0 1 2 3
N S O A	Ask students what they already know about the content of the selection prior to reading.	P F G E	0 1 2 3
N S O A	Help student understand why he/she is learning the relationships between letters and sounds.	P F G E	0 1 2 3
N S O A	Directly teach difficult words such as those with multiple meanings.	P F G E	0 1 2 3
N S O A	Help student apply what he/she learns about sounds and letters to own writing.	P F G E	0 1 2 3
N S O A	Directly teach words that are important for understanding a concept or the text.	P F G E	0 1 2 3

Part 2 of 4

Implementation: Teacher Strategies

Question A: How often have you used this strategy with this student?	Teacher Strategies	Question B: How would you rate your ability to use this strategy?	Question C: How much progress has the student made?
N S O A	Teach root words (words from other languages that are the origin of many English words) that students are likely to see often.	P F G E	0 1 2 3
N S O A	Teach students to manipulate phonemes along with letters.	P F G E	0 1 2 3
N S O A	Show students pictures or diagrams to prepare them for what they are about to read.	P F G E	0 1 2 3
N S O A	Directly teach words that students are likely to see and use again and again.	P F G E	0 1 2 3
N S O A	Help student apply knowledge of phonics as he/she read words, sentences, and text.	P F G E	0 1 2 3
N S O A	Directly teach idiomatic expressions such as "drawing a blank".	P F G E	0 1 2 3
N S O A	Teach letter shapes and names.	P F G E	0 1 2 3
N S O A	Read aloud from a big book showing students where and how you are pausing and how the text shows you when to raise or lower your voice.	P F G E	0 1 2 3

Part 3 of 4 **Components of Instruction**

Consider the total amount of reading instruction you have delivered to this child thus far this year. Please estimate the **percent** of time you have devoted to each component of reading instruction described below. Your total percent should equal 100.

_____ **Phonemic Awareness Instruction:** Any instructional activities that involve hearing, identifying, or manipulating sounds (i.e., segmenting, blending) in spoken words.

_____ **Phonics Instruction:** Any instructional activities that involve learning the relationships between letters of written language and the sounds of spoken language.

_____ **Fluency Instruction:** Any instructional activities that involve improving the accuracy, speed, and expressiveness of oral reading.

_____ **Vocabulary Instruction:** Any instructional activities that involve direct or indirect development of oral vocabulary or reading vocabulary.

_____ **Text Comprehension Instruction:** Any instructional activities that involve direct instruction of comprehension strategies or the application of comprehension strategies to text.

_____ **Other Instruction Related to Reading:** (Please describe) _____

_____ **Other Instruction Related to Reading:** (Please describe) _____

_____ **Total** (Should equal 100%)

On average, how many hours or reading instruction do you deliver to this student each week? _____

Part 4 of 4 **Assessment**

Assessments Conducted

How have you measured this student's progress in reading, and how often have you done it? Using the frequency codes provided below, please complete each box within the table indicating how often you have used each type of assessment to measure this student's progress in each component of reading. Refer back to the following page for descriptions of the components.

Frequency Codes:

N = Never
O = Once since beginning of year
Q = Quarterly
M = Monthly
W = Weekly
D = Daily

Components of Reading

Type of Assessment	Phonemic Awareness	Phonics	Fluency	Vocabulary	Text Comp.
Standardized reading assessments not linked to a particular reading text					
Tests provided by the publisher of a reading text					
Teacher made tests					
Structured oral reading (e.g., running records)					
Informal oral reading					
Structured observation (e.g., checklist of print concept skills)					
Informal observations					
Skills worksheets					
Other:					

Please provide the names of the standardized reading assessments you have used with this student:

_____

Please describe any of the assessments you listed as "Other":

_____

Part 4 of 4 **Assessment**

Assessment Products

Step 1. Collect at least three different examples of assessments you have conducted for this student.

- Choose examples that are representative of your assessment practices.
- Please do not consider the student's performance when making your selection. (We are interested in the *kinds* of assessments you use, not the *outcomes* of these assessments.)
- Assessments may be as informal as anecdotal notes or as formal as state or district measures.

Step 2. Complete the chart below for each example you have chosen.

Step 3. Copy these products, removing all identifiers of the student, and place the copies in the return packet, with this instrument.

Student Work	Please name and describe the assessment measure in as much detail as possible.	What is/are the goal(s) of this assessment?	How many times have you done this assessment with this child?
1			
2			
3			

Thank you for your time and efforts with this survey.

- **Please remember to include at least three examples of this student's work in the return packet.**
- **If you would like a copy of the results from this study, please remember to check the appropriate box on the consent form.**

Response Codes

Place this sheet at the top of each page starting with page 3 and use it to guide your responses.

Question A:	Question B:	Question C:
N = Never S = Sometimes O = Often A = Almost Always	P = Poor F = Fair G = Good E = Excellent	0 = No progress 1 = A little progress 2 = Some progress 3 = A lot of progress

Appendix C

Rubric for Scoring Assessment Products

Rubric for Scoring Assessment Products

Step 1: Decide which assessment components are reflected in the measure. To do this, consider the goal(s) of the assessment and how it was evaluated by the teacher. A measure may include more than one component.

Description of Assessment Components

Phonemic Awareness Assessment: Any measure of the student's ability to hear, identify, and/or manipulate sounds (e.g., segmenting, blending) in spoken words.
Examples: DIBELS – Initial Sound Fluency, Phoneme Segmentation Fluency

Phonics Assessment: Any measure of the student's ability to make connections between letters of written language and the sounds of spoken language.
Examples: DIBELS – Letter Naming Fluency, Nonsense Word Fluency

Fluency Assessment: Any measure of the student's accuracy, speed, and expressiveness of oral language.
Examples: Running records

Vocabulary Assessment: Any measure of the student's understanding of oral vocabulary or reading vocabulary.
Examples: Pre/Post-tests of reading vocabulary

Text Comprehension Assessment: Any measures of comprehension of written language.
Examples: CBM – maze passages, Reading passage comprehension tests

Other Reading Assessment: Any measure not reflected in the components described above.

Step 2: For those components not present in the assessment product, score a **0**. For components that are present, use the following criteria:

- Score a **1** if the product illustrates only one measure of this component.
- Score a **2** if the product illustrates two different measures within this component.
- Score a **3** if the product illustrates at least three different measures within this component.

Example 1. Running Record with Retell

Goals: Analysis of errors (meaning, language, visual/graphophonic), accuracy of reading, ability to describe the setting, characters, plot, problem/solution, and make connections to other texts or to self

Component	Score	Evidence
PA	0	
P	1	Error analysis
F	1	% words correct
V	0	
TC	3	Setting, plot, characters, etc.

Example 2. DIBELS

Goals: accuracy and speed of initial sound identification (pointing response), accuracy and speed of letter naming (oral response), accuracy and speed of phoneme segmentation (oral response), accuracy and speed of nonsense word identification (oral response).

Component	Score	Evidence
PA	2	Initial sound, phoneme segment.
P	2	Letter naming, nonsense word
F	2	Speed, accuracy
V	0	
TC	0	

Appendix D

Informed Consent Form for Research

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study Teachers' Understanding and Implementation of Scientifically Based Reading Programs

Principal Investigator Kelly Laugle Faculty Sponsor Jeff Braden, PhD.

We are asking you to participate in a research study. The purpose of this study is to assess teachers' implementation of the five components of sound reading instruction as identified by the National Reading Panel. This includes instruction in the areas of phonemic awareness, phonics, fluency, vocabulary, and text comprehension.

INFORMATION
If you agree to participate in this study, you will be asked to complete the following steps: 1) Select the middle student from your "middle" reading group to focus your responses or select a student from the middle of your class distribution for reading levels or current grades in reading. 2) Complete a self-report instrument on the reading instruction you have delivered to that child since the beginning of the school year. 3) Collect products completed by the student that represent your assessment practices in reading. 4) Remove all identifiers of the student from the products to be submitted. 5) Mail the products and completed instrument to the research staff at NC State.

RISKS
You may feel uncomfortable rating your own level of instruction. However, your name will not appear on the responses forms, nor will your name be stored in the database. You may also feel uncomfortable submitting a student's work to research staff. You will have free choice of what materials to submit and will be asked to remove all identifiers of the student before submitting.

BENEFITS
Direct benefits to the participant include exposure to potentially new methods and topics of reading instruction and an opportunity for self-evaluation. In addition, completion of this rating instrument will allow for its stream-lining in future development to reduce the amount of time required for completion by other teachers. Students may also benefit indirectly from this study. Providing a measure of teachers' reading instruction may facilitate informed decision making in the identification of students with disabilities. This measure may allow a team to rule-in or rule-out insufficient learning opportunities as a reason for inadequate academic progress.

CONFIDENTIALITY
The information in the study records will be kept strictly confidential. Data will be stored securely in an Excel database on the principal investigator's personal computer. Contact information will remain separate from response data. You will be provided with an identification number to place on the materials you submit. This number will be matched to your name only if we need to contact you to send additional materials. If you do not want us to match your number with your contact information, you may check the appropriate box at the end of this form. No reference will be made in oral or written reports which could link you to the study.

COMPENSATION
For completing this study you will have the opportunity to be entered into a drawing to receive a $100 gift certificate to a teacher supply store. Your name will be entered in the drawing (if you provide consent) when all materials have been received by research staff. If you would be willing to complete the instrument a second time, your name will be entered into the drawing again after the second set of materials has been received to increase your chances of winning the gift certificate.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the researcher, Kelly Laugle at kmlaugle@ncsu.edu, or 919-673-3222. If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Dr. Matthew Zingraff, Chair of the NCSU IRB for the Use of Human Subjects in Research Committee, Box 7514, NCSU Campus (919/513-1834) or Mr. Matthew Ronning, Assistant Vice Chancellor, Research Administration, Box 7514, NCSU Campus (919/513-2148)

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed at your request.

CONSENT

If you understand the above information and agree to participate, please complete the instrument and return it, along with the bottom of this page. Please keep this letter for your records.

Investigator's signature___________________________________ Date _________________

Your Research Identification number is: ___________________________________

_ _

CONSENT

"I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may withdraw at any time."

Please check one of the following:

☐ I would like to remain completely anonymous.

☐ I will provide my contact information below so that I may receive a second mailing of the instrument, but I do not wish to be entered in the drawing to win a $100.00 gift certificate to a teacher supply store.

☐ I will provide my contact information on a separate sheet of paper to be entered into the drawing to win a $100.00 gift certificate to a teacher supply store, but I do not wish to receive a second mailing of the instrument.

☐ I will provide my contact information below so that I may receive a second mailing of the instrument, and I will provide my contact information on a separate sheet of paper to be entered TWICE into the drawing to win a $100.00 gift certificate to a teacher supply store.

Please use your research identification number as your signature.

Subjects research ID #: ___________________________________ Date: ________________

Contact Information to receive a second mailing:

Name: _____ Email: _____

Phone number:_____

Address:

_____

If you would like to receive a copy of the results from this study, please check the box below and provide your email address.

☐ Yes, I would like to receive a copy of the results. My email address is: _____

Appendix E

IRB Approval Letter

Sponsored Programs and
Regulatory Compliance
Campus Box 7514
1 Leazar Hall
Raleigh, NC 27695-7514

919.515.7200
919.515.7721 (fax)

From: Debra A. Paxton, Regulatory Compliance Administrator
North Carolina State University
Institutional Review Board

Date: October 22, 2004

Project Title: Evaluating Teachers' Understanding and Implementation of the Essential Components of Reading Instruction

IRB#: 229-04-10

Dear Ms. Laugle:

The research proposal named above has received administrative review and has been approved as exempt from the policy as outlined in the Code of Federal Regulations (Exemption: 46.101.b.2). Provided that the only participation of the subjects is as described in the proposal narrative, this project is exempt from further review.

NOTE:

1. This committee complies with requirements found in Title 45 part 46 of The Code of Federal Regulations.
   For NCSU projects, the Assurance Number is: FWA00003429; the IRB Number is: IRB00000330
2. Review de novo of this proposal is necessary if any significant alterations/additions are made.

Please provide your faculty sponsor with a copy of this letter. Thank you.

Sincerely,
Debra Paxton
NCSU IRB