EFFECTS OF TIER 1 ENHANCEMENT TRAINING ON TEACHERS' PERCENTAGE OF CORRECTLY IMPLEMENTED INSTRUCTIONAL UNITS

by

Allison Graves Kretlow

A dissertation submitted to the faculty
of the University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Special Education

Charlotte

2009

Approved by:

____
Dr. Nancy Cooke

____
Dr. Kent Curran

____
Dr. Ya-yu Lo

____
Dr. Charles Wood

ABSTRACT

ALLISON GRAVES KRETLOW. Effects of tier 1 enhancement training on teachers' instructional unit accuracy. (Under the direction of DR. NANCY L. COOKE)

Early intervention models such as Response to Intervention have shown promise in reducing risk of academic failure (Bursuck et al., 2005; Foorman, Fletcher, Francis, & Schatschneider, 1998). General education teachers assume primary responsibility for instruction in RTI; however, many report lack of preparation for this role due to lack of high quality, sustained professional development (Schumaker et al., 2002). Professional development models that include a combination of inservice and coaching have demonstrated effectiveness in promoting sustained changes in teachers' instruction (Yoon et al., 2008). This study examined the effects of inservice plus coaching on 1st grade teachers' accurate delivery of instructional units in CM. The extent to which changes in instruction generalized to an untrained math session was also examined. Teachers were trained to use a combination of whole-class instruction strategies, including model-lead-test for introducing new concepts and correcting errors, choral responding, and response cards. Results indicated that all teachers improved their delivery of instruction after the inservice, with a second level of growth achieved after coaching. Results indicated some generalization to an untrained math session as well. Teachers also reported high levels of satisfaction using the strategies, as well as high levels of acceptability with the training model.

DEDICATION

I would like to dedicate this dissertation to my family. First, I would like to thank my husband, Phillip for making space in our marriage for this professional journey. Second, I am grateful to my parents, who pushed me to go after my goals, and who always believed I could accomplish them. Third, I am grateful to my family and friends outside the doctoral program, especially my sisters, for helping me remember the most important things in life. Finally, I am thankful to Pam for her constant support, collaboration, and most of all, her friendship throughout this program.

ACKNOWLEDGEMENTS

I would like express my sincere appreciation to several individuals who have been vital to my success. First, I would like to thank my advisor, Nancy Cooke for setting high expectations for me to become a better thinker, a better writer, and a better researcher, and for raising the bar every time I met it. Second, I am grateful to Charlie Wood for giving me so many opportunities to develop critical professional skills, and for providing constant, helpful feedback along the way. Third, I am thankful to my other committee members, Ya-yu Lo, Bill Bursuck, and Kent Curran for their expertise during my dissertation. Next, I am thankful for the teachers who participated in this study, for working so hard to improve instruction. I am also grateful to Melissa Kelly and Angela Burns who spent many hours collecting data for this study. Finally, I am very thankful for the Giles Dissertation Year Fellowship, which allowed me the time and resources to focus wholly on my dissertation for the last year.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Statement of the Problem

Young students at risk for failure. Many young students in general education classrooms are at risk for academic failure and special education referral due to poverty, English Language Learner (ELL) status, disability and lack of early academic experiences (Coyne, Kame'enui & Carnine, 2007). Poverty has been repeatedly identified as the strongest predictor of risk in reading and math for students entering school, and up to one in five students in the United States live in poverty (U.S. Census Bureau, 2007). These data suggest that as many as 20 percent of students in a typical classroom may be at risk for academic failure. In high poverty schools the number of students considered at risk is even higher. Many students who fail to meet grade level expectations in early grades often continue to fail and eventually qualify for special education services (Simmons, Kame'enui, Coyne & Chard, 2002). These students often need more intensive instruction to demonstrate mastery with skills and prevent the need for special education services (Coyne et al.). Substantial empirical research has demonstrated that early academic intervention is key to preventing failure (Bursuck et al., 2004; Foorman, Francis, Fletcher & Schatschneider, 1998; Fuchs & Fuchs, 2006).

Challenges for general educators. In early intervention models, general educators are responsible for providing primary (i.e., Tier 1) instruction for students at risk for failure. These models of intervention (e.g., Response to Intervention, RTI) have substantially changed the role of the general educator (Haagar & Mahdavi, 2007).

Historically, general educators provided primary instruction tied to grade level curriculum and standards, whereas special educators often focused on designing, implementing, and monitoring intensive instruction for students with disabilities (Haagar & Mahdavi). In RTI models, general educators assume primary responsibility for providing research-based Tier 1 instruction to students with a wide range of instructional needs. Given diverse populations, general educators need not only research-based curricula, but also need instructional practices that will support students at risk for academic failure who typically participate in Tier 1 instruction.

Substantial special education research has identified that instructional interventions developed to improve achievement can be implemented in whole-class settings, and are appropriate for a wide range of student achievement levels (Baker, Gersten, & Lee, 2002; Gersten, Chard, Baker, & Lee, 2002; Swanson & Hoskyn, 1998). Recent research has demonstrated that general educators can impose many of these strategies on existing core curricula. For example, Bursuck et al. (2005) trained general educators to strengthen core reading instruction using a combination of instructional “enhancements” such as unison responding, systematic error correction and scaffolding using a Direct Instruction (DI) model-lead-test (MLT) procedure. Also, Fuchs and colleagues have trained general educators to deliver more systematic, explicit instruction to students at risk for failure during Tier 1 instruction across a variety of grade levels and content areas (i.e., reading and math; Compton, Fuchs, Fuchs, & Bryant, 2006; Fuchs et al., 2006; Fuchs et al., 2008).

Despite the success of previous investigations with general educators, empirical evidence also suggests that general educators may be unprepared to deliver instruction that is designed to improve achievement of students at risk for academic failure (Brownell, Ross, Colon, & McCallum, 2005). Previous research has unveiled several reasons for this. First, general teacher education programs lack concentration on instructional methods for students at risk (Brownell et al.; National Council on Teacher Quality, 2006). Second, general education curricula often do not include critical features of instructional design that work for students who struggle with learning (Coyne et al., 2007). Third, many in-service teachers lack access to quality professional development focused on improving achievement for students at risk. Specifically, in several surveys teachers have reported limited access to professional development opportunities for strategies targeted toward students at risk (Boardman, Arguelles, & Vaughn, 2005; Schumaker et al., 2002).

Other empirical data suggest that teachers at schools with high numbers of students at risk face particular gaps in pedagogical knowledge. For example, Ingersoll (2001) found that teachers in high-poverty schools were less skilled at instruction than teachers in low-poverty schools. In addition, principals have reported that teachers at low performing schools are less effective than teachers at high performing schools. The most frequent reason principals report for this ineffectiveness is lack of lesson implementation skills (Torff & Sessions, 2005). The combined implications of these data are discouraging for students who need highly systematic, teacher directed, effective instruction to be successful academically, and lead directly to the need for high quality professional development to improve instruction in general education classrooms.

Challenges of professional development and teacher change. A growing body of research suggests that alternative models of professional development (e.g., beyond the one-day inservice) more effectively support teachers' behavioral change. Reviews of professional development literature (Knight & Wiseman, 2005; Yoon et al., 2007) have identified important variables that reliably produced sustained teacher change and improved student achievement. These variables include: (a) at least 14 cumulative contact hours; (b) combination of inservice and follow-up support (e.g., coaching, booster workshops); and (c) content provided directly to teachers by experts (i.e., not a "train the trainer" approach). Other research conducted with general education teachers has demonstrated that in order for a specific practice to be sustained, it must be easy to implement, not require additional materials, be relatively inexpensive, not require substantial time, and help all students in the classroom (Delquadri, Greenwood, Stretton, & Hall, 1983; Maheady, 2008). In addition, a few of single subject studies have demonstrated that follow-up coaching more effectively supports sustained teacher changes after inservice professional development than no follow up or other types of follow up (i.e., additional meetings, group inservices; Hasbrouk, 1997; Morgan, Menlove, Salzberg, Hudson, 1994).

Limitations of Current Research

Current research in general education settings. Research conducted with general educators is still limited in several ways. First, very few studies have investigated whole class interventions in Tier 1. Most studies conducted in preventive models have focused on Tier 2 and Tier 3 interventions and involved small group, supplemental instruction provided by a paraeducator or special educator, not the general educator. Tier 1 interventions are a critical area of exploration given that the most common form of instruction in many general education classrooms is whole group (Schumaker et al., 2002).

Second, the majority of studies examining Tier 1 interventions have been conducted in reading (Bursuck et al, 2004; Fuchs & Deshler, 2007; Vaughn & Klingner, 2007). Only a few studies to date have been conducted in math (Fuchs, et al., 2005; Fuchs, Fuchs, et al., 2006). Math is a critical area for early intervention, as it is assessed under No Child Left Behind (NCLB, 2001) and an area of special education identification under the category Specific Learning Disability (Individuals with Disabilities Education Act, 2007). In addition, mastering foundational math skills is critical to success in later grades when concepts build upon prerequisite skills (Harniss, Carnine, Silbert, & Dixon, 2007). For example, to demonstrate mastery solving an algebraic equation, students must have mastered arithmetic operations. Similarly, to demonstrate mastery with arithmetic operations, students must understand basic numeracy concepts such as rote counting, one-to-one correspondence, and more and less, which are taught in early elementary grades.

Abundant research has demonstrated that direct and explicit instruction improves at-risk learners' mastery of basic and higher order math skills (Przychodzin, Marchand-Martella, Martella, & Azim, 2004; Harniss et al., 2007); however, the National Math Panel (2008) reported that studies are still needed to determine the impact of highly systematic instruction and additional practice with feedback for students in general education settings. In addition, the panel reported a need to "unpack the underlying variables" (Gersten et al., p. 235) associated with teacher directed, explicit instruction in mathematics. More specifically, studies are needed to identify and validate components of math instruction which improve achievement of students at risk.

Professional development research. In addition to the need for early intervention math studies, research is also limited in the area of professional development. For example, although coaching studies have demonstrated positive results, there are very few of them. Also, reviews of professional development literature have found that very few high quality studies have demonstrated the effects of sustained professional development models on teacher change. For example, Yoon et al. (2007) retrieved over 1,000 studies examining the impact of professional development, yet only nine met the Institute for Education Sciences (IES) quality indicators for research design. Yoon and colleagues noted that many studies were excluded from the review because they were qualitative, used teacher reported measures of change instead of behavioral measures, or did not demonstrate experimental control. Other researchers have also noted the lack of quality single subject studies on professional development (Maheady, Harper, Mallette, & Karnes, 2004).

Similarly, measurement of teacher and student changes as a direct result of professional development warrants further investigation. The majority of professional development studies used subjective measures such as teacher reports or qualitative observation tools to examine teacher change (i.e., over 900 out of 1,000 studies; Yoon et al., 2007). Although teacher perceptions and contextual factors are important variables in the transportability of any new instructional practice; direct, behavioral measures of teacher change are critical in validating the effects of professional development activities on effective instruction. Data collection systems are needed that allow for sensitive measurement of critical teaching behaviors that are transportable across skills, activities, and content areas. Measurement of student changes as a result of teacher training, specifically related to coaching needs further investigation as well. Although most of the empirical coaching studies investigate changes in student achievement, these studies used group performance on existing classroom measures (e.g., spelling tests, number of lessons mastered). More studies need to examine the effects of changes in teacher instruction measured by sensitive, observable tools. In addition, the achievement of students at risk for failure as a direct result of changes in teaching practice needs to be measured.

Finally, no studies to date have examined the extent to which teachers generalize skills learned in professional development to instructional areas not directly trained. The level of support needed to promote generalization is an important variable as schools consider the design of teacher training and the allocation of resources for professional development activities.

Purpose of Study and Research Questions

Given the limitations of previous research, the need exists to examine the effects of a high quality, sustained professional development model on general educators' accurate delivery and generalization of new instructional behaviors in the area of math. The purpose of this study was to investigate the effects of enhancement training on 1st grade general education teachers' percentage of instructional units implemented correctly and the extent to which teachers generalized the correct implementation of instructional units to an untrained area of math.

The specific research questions addressed in this study were:

(a) What is the effect of enhancement training in the form of inservice and coaching, on teachers' accurate delivery of instructional units in calendar math?

(b) To what extent does enhancement training, focused on teaching skills addressed during calendar math, generalize to instructional delivery during a second untrained daily math session focused on numeracy and problem solving?

(c) How many 1st grade students who receive enhancements during Tier 1 insruction change risk level (i.e., from "some risk" to "low risk" or "low risk" to "some risk")?

(d) How many 1st grade students who do not receive enhancements during Tier 1 instruction change risk level?

(e) What additive value do teachers place on coaching in relation to enhancement training through group inservice alone?

(f) To what degree do teachers attribute enhancements in lessons to improvements in student achievement and progress through the curriculum?

Significance of the Study

This study has the potential to contribute to the research base in the following ways. First, the study may provide a model for training general educators to use efficient, low cost teaching practices that improve academic engagement and achievement in Tier 1 math instruction. Second, the study may add to the growing but limited research in which general educators serve as "change agents" in preventive instruction models. Specifically, changes in general educators' instructional methods using this model have the potential to improve achievement of large numbers of students over time. In addition, the study may provide empirical evidence for a professional development model that involves supporting teachers' acquisition of new skills through individualized side-by-side coaching. The study may also contribute more evidence to the small, growing body of RTI math research using a design that demonstrates experimental control of a generalized set of specific teaching behaviors.

Definitions

The terms defined below are critical to understanding the conceptual framework and methodology of the proposed study, are used throughout the following chapters.

Active student responding- "Active student responding can be defined as an observable response made to an instructional antecedent. ASR occurs when a student emits a detectable response to ongoing instruction" (Heward, 1994, p. 286). Examples include words read, questions answered, or problems solved.

Choral responding- "Each student in the class or group responds orally in unison to a question, problem, or item presented by the teacher" (Heward, 1994, p.286).

Coaching- Coaching is a component of professional development that involves an expert (e.g., trainer, peer) providing individualized support to teachers after an initial training has occurred. Coaching is typically provided in one of two ways. Supervisory coaching involves observation and feedback given after a lesson. Side-by-side coaching involves the coach directly intervening during the lesson by providing a model, a rationale for the change, and additional opportunities for the teacher to teach the same format again with immediate feedback from the coach (Blakely, 2001).

Enhancements-instructional procedures added to existing curriculum in order to increase student responding and task mastery (Bursuck & Damer, 2007). Some examples of enhancements include unison responses, brisk instructional pace, increased practice, model-lead-test format and systematic error correction.

Learning trial-"A learning trial consists of three major elements: antecedent (i.e., curricular) stimuli, the student's response to those stimuli, and consequent stimuli (i.e., instructional feedback) following the response" (Heward, 1994, p. 284).

Model-Lead-Test-A teaching format in which "…the teacher first demonstrates how to do the new skill so that the students have no difficulty understanding exactly what the new skill looks like" (i.e., model). Then, "the teacher practices the skill with his or her students until they are able to do it without him or her" (i.e., lead). Lastly, "the teacher monitors students as they do the skill independently" (i.e., test, Bursuck & Damer, 2007, p. 16).

Professional development- continuing education in the form of specific training designed to update teachers' knowledge of current and changing practices. Professional development formats can include any combination of the following: workshops, inservices, field study, demonstration, coaching, and meetings (Garet et al., 2001).

Response cards-"…cards, signs, or items that are held up simultaneously by all students to display their response to a question or problem presented by the teacher" (Heward, 1994, p. 299).

Response to Intervention- "RTI is a method for both preventing and helping to identify learning disabilities. An important feature is its multi-tier structure: *primary intervention* (tier 1) refers to classroom instruction; *secondary intervention* (tier 2) usually involves more intensive pullout, small-group instruction; and *tertiary intervention* (tier 3) typically denotes most intensive, often special education services. General education represents primary intervention. Students demonstrating unsatisfactory progress in the regular classroom enter a more intensive secondary intervention. In most RTI research, this involves one or more rounds of small-group tutoring in which instruction is driven by an evidence-based standard protocol. Students who also respond poorly to the secondary intervention are understood to have demonstrated 'unexpected failure,' and they become candidates for the most intensive, tertiary intervention, or special education. Prior to moving from secondary to tertiary intervention, students undergo an abbreviated evaluation to determine a likely cause for the observed academic failure and an appropriate instructional placement and plan" (Fuchs, Compton, et al., 2008, p. 415).

Scaffolding-temporary support provided by the teacher for students to learn new material. This support is eventually faded as students are able to demonstrate mastery independently (Coyne et al., 2007)

Students at risk-students who, based on poverty, English language learner status, disability or lack of early academic experiences, enter school with deficits critical to academic success, putting them at risk for academic failure. These students can be identified by below benchmark performance on predictive measures of early literacy and mathematics skills (Bursuck & Damer, 2007; Vaughn, Wanzek, Woodruff, & Linan-Thompson, 2007).

Systematic error correction-teacher corrects students immediately after they make an error using a model-lead-test, model-test, or lead-test procedure (Bursuck & Damer, 2007).

Tier 1-Tier 1 is composed of three elements: "1) a core program or curriculum based on scientific research, 2) screening and benchmark testing of students at least three times each year (i.e., fall, winter, and spring) to determine instructional needs, and 3) ongoing professional development. Tier 1 instruction is designed to address the needs of the majority of a school's students" (Vaughn et al., 2007). In addition, Tier 1 instruction is delivered solely by the general education teacher and is "the least intensive, first level of instruction that consists of the current program used in the classroom" (Bursuck & Damer, 2007, p. 10).

Tier 2-Tier 2 is designed to meet the needs of students who need core and supplementary instruction by providing them with an additional amount of small group instruction daily. Supplementary instruction is designed to provide students with additional practice on skills taught during Tier 1 instruction (Vaughn et al., 2007). Tier 2 instruction typically (Bursuck & Damer, 2007, p. 11) ranges from 10-30 min to ensure that students become accurate and fluent. Tier 2 instruction is delivered by a variety of school personnel including general educators, special education teachers, and paraprofessionals.

Tier 3-Tier 3 is designed to meet the needs of students who show difficulty in acquiring academic skills even after Tier 2 intervention. Tier 3 students need more intensified and lengthy individualized explicit instruction (e.g., 45-60 minutes per day). In some RTI models, Tier 3 involves special education services. (Vaughn et al., 2007).

Tier 3 instruction is “more intensive instruction using an alternative program conducted daily for 30-90 min in separate pull-out settings with small, skill-based groups of 2 to 5 students” (Bursuck & Damer, 2007, p. 11). Tier 3 instruction is most often delivered by a special educator.

CHAPTER 2: REVIEW OF LITERATURE

During the course of the 1997 IDEA reauthorization, concern that neither accurate nor early identification or intervention for students with learning disabilities was occurring led the U.S. Office of Special Education Programs (OSEP) to consider alternative procedures for early intervention (Bradley, Danielson, & Doolittle, 2007). The ramifications of the traditional assessment and intervention model (i.e., discrepancy model) were twofold. First, students had to fail academically for many years in order to demonstrate a large enough discrepancy to warrant any intervention. Second, the traditional model led to an over-representation of minority and male students in special education.

In response to the concerns related to traditional assessment and intervention, OSEP joined with researchers and other stakeholders to form the Learning Disabilities (LD) Initiative, whose work resulted in the investigation of responsiveness to intervention as an alternative method of early intervention that could potentially prevent and help accurately identify learning disabilities (NJCLD, 1997). RTI includes universal screening and progress monitoring in conjunction with an increasingly intensive tiered system of intervention. Tier 1 intervention occurs in the general education classroom with all students and includes instruction with a research-based core program, Tier 2 intervention involves more intensive, small group pull-out intervention, and Tier 3 intervention involves even more intensive instruction, smaller group sizes, and often special education services (Fuchs, Fuchs, Compton et al., 2008). RTI is considered a worthy alternative to the traditional assessment and intervention model because it does not require students to fail before intervention is provided, provides early intervention in the general education setting, and allows teachers to gain important information about student progress on a frequent basis (Kame'enui, 2007).

Critical Features of Effective Tier 1 Intervention

Since the work of the LD Initiative began, subsequent RTI research has demonstrated that early intervention in the general education classroom is critical for preventing academic failure and referral to special education (Vellutino, Scanlon, Small, & Fanuele, 2006). The majority of RTI research has been focused on reading, given that the majority of students with LD are identified because of reading deficits. There is now substantial evidence demonstrating that the quality of Tier 1 instruction can improve academic achievement of students at risk for failure. In particular, high levels of certain instructional characteristics including explicitness, engagement, monitoring, and planning are most effective when used in Tier 1 intervention (Cirino, Pollard-Durodola, Foorman, Carolson, & Francis, 2007; Foorman et al., 1998; Foorman et al., 2006; McIntosh, Graves, & Gersten, 2007).

For example, Foorman et al. (1998) compared three levels of explicitness of alphabetic code instruction on reading achievement for first and second graders at risk: direct instruction, less direct instruction (i.e., embedded in connected text), and indirect (i.e., incidental instruction fully embedded in connected text). Three levels of explicitness were examined during whole-class instruction delivered by general educators. Using Hierarchical Linear Modeling researchers accurately predicted phonemic awareness, word growth, and vocabulary knowledge by level of explicitness for 285 first and second grade students. The model demonstrated that the highest level of explicitness (i.e., direct instruction) predicted a higher mean rate of change (i.e., growth curve) for end of year reading achievement. The most explicit instruction had the greatest impact for students who began the year with the lowest reading skills; specifically students with the lowest achievement demonstrated the lowest overall growth across the year except for students in the most explicit condition. These students still managed to show substantial growth in reading. In addition, logistic regression revealed no difference for effects based on whole group or small group instruction in the explicit group, suggesting that explicit instruction may be effective enough for students at risk to use in whole class settings.

In a similar investigation, Foorman et al. (2006) also found that higher ratings on a seven item observational scale better predicted language arts skills for over one thousand first and second graders in high poverty schools. Researchers used the Checklist of Teacher Competencies (CTC; Foorman et al., 1998) to rate teachers on the following instructional characteristics using a five point scale ranging from “never” to “always”: (a) lesson planning (the extent to which teachers used and followed a prescribed lesson plan); (b) engagement (measured by on and off task behavior of four target students); (c) monitoring of student progress during lessons; and (d) feedback (i.e., praise, correction). Higher total scores on the five item measure accurately predicted higher end of year scores on word reading, a cloze measure, and an oral spelling measure.

Cirino et al. (2007) extended the evidence for the correlation between teaching quality and student achievement to ELL. First, researchers rated the quality of instruction provided by 141 teachers across 35 schools using the Texas Teacher Appraisal System

(TTAS; Texas Education Agency, 1984) which includes observational measures of instructional strategies, classroom management, active participation of students, and scaffolding; these measures are summed into a total score. The TTAS was completed three times across the school year. Second, researchers also used the CTC developed by Foorman et al. (1998) and used again by Foorman et al. (2006) to rate planning, monitoring, instruction, and engagement. Third, researchers combined the TTAS ratings with results from a 70-item survey of teachers' knowledge of reading content and specialized instructional practices (i.e., the Beginning of Year Survey; BOYS). Results of the teacher observations echoed those found by Foorman et al. (2006); that is, teachers with higher scores on both the TTAS and the CTC had students with higher end of year oral language and reading proficiency. However, Cirino et al. also extended the work of Foorman et al. by demonstrating that teacher knowledge of the content and practices significantly reduced the variability of students' outcomes; therefore when teacher knowledge was combined with high teaching quality ratings, the combination best predicted end of year performance for ELL. Foorman et al. (1998, 2006) and Cirino et al. (2007) clearly demonstrated how instructional quality can explain changes in student achievement.

McIntosh, Graves, and Gersten (2007) further explained the relationship between teacher quality and student outcomes by demonstrating that instructional quality could also predict referral to special education for ELL. Similar to the other three studies, McIntosh et al. measured teacher effectiveness using an observation tool with moderate inference (i.e., Likert scale). Researchers used the English Learners Classroom Observation Instrument (ELCOI; Haager, Gersten, Baker, & Graves, 2003) during Tier 1 and Tier 2 instruction to rate explicit teaching (e.g., modeling, prompting, making relationships overt), interactive teaching, and items related to content (vocabulary, phonemic awareness, etc.). Researchers tracked reading outcomes of 111 ELL at risk from first to third grades and discovered that students with higher quality Tier 1 instruction were less likely to be identified in need of special education. Specifically, 9 students qualified for special education at the end of third grade; 5 of the 9 were from classrooms with lower instructional quality ratings.

Studies by Foorman et al. (1998; 2006), Cirino (2007), and McIntosh et al. (2007), examining the relationship of Tier 1 teaching quality and academic outcomes for students at risk, suggest that particular instructional characteristics better predict and explain higher achievement. First, these studies support direct, explicit instruction over embedded explicit instruction and indirect, incidental instruction. Second, these studies support active engagement, frequent monitoring of student learning during lessons, feedback on learning, and lesson planning. The three studies described also suggest that these instructional characteristics can influence student achievement in whole class settings and may predict special education placement. Finally, these studies support specialized knowledge of content and practices within a curricular domain.

Other researchers have offered experimental support for Tier 1 interventions using these characteristics in reducing academic risk and referral to special education (Berninger et al., 2006; Fuchs, Fuchs & Burish, 2000, & McIntosh, Graves, and Gersten, 2007). However, in general, research on teacher delivered Tier 1 interventions is sparse. The majority of existing Tier 1 intervention research has been focused on Classwide Peer Tutoring (Fuchs, Fuchs, Otaiba et al., 2001; CWPT; Greenwood et al., 2001; Greenwood, Maheady, & Delquadri, 2002). RTI researchers have used this instructional strategy because it is well supported by empirical evidence and is easily implemented in whole class settings. Although CWPT is not teacher delivered instruction (i.e., it involves reciprocal peer tutoring), it is teacher directed instruction and to be implemented successfully depends heavily on factors found critical by Foorman and colleagues (1998, 2006, 2007) such as explicit instruction, planning, engagement, monitoring of learning, and feedback.

Greenwood, Maheady, and Delquadri (2002) and Greenwood et al. (2001) demonstrated that general educators effectively implemented CWPT procedures to improve engagement during instruction, reading, language, and math performance for students at risk including ELL when compared to low, average, and high performing students. Researchers demonstrated that progressive implementation of CWPT (i.e., students engaged in peer tutoring across first through fourth grade) resulted in continued academic growth across 4 school years.

Similarly, Fuchs, Fuchs, Otaiba et al. (2001) demonstrated the efficacy of Tier 1 interventions in two studies using another model of CWPT (i.e., Peer Assisted Learning Strategies; PALS). In these two studies, general educators implemented PALS and K-PALS in reading. After PALS implementation, students with low achievement showed higher phonological awareness and word reading skills compared to students in control groups and the skills were maintained 5 months later. Lane et al. (2007) extended the results of PALS to show improvements in reading fluency and a decrease in variability of academic engagement in general education classrooms. Fuchs, Fuchs, Otaiba et al. and

Lane et al. attributed academic growth to frequent, structured opportunities for academic engagement during reciprocal tutoring and corrective feedback provided by peers.

Although much of the Tier 1 research has been focused on CWPT, one study has experimentally examined the impact of teacher delivered Tier 1 interventions (Bursuck, et al., 2004). Bursuck et al. trained kindergarten, first and second grade teachers at high poverty schools to deliver whole class systematic, explicit reading instruction. Tier 1 instruction included a core reading program with added enhancements (e.g., advance organizers, unison responding, model-lead-test). Teachers in project PRIDE were provided with scripted teaching formats to accompany each lesson from the core reading program that prompted them to use enhancements. Teachers received group workshops as well as individual demonstration and coaching when workshop and simulated practice was not sufficient.

Results indicated that over half (53%) of the students identified at risk achieved benchmark scores on the Nonsense Word Fluency and Oral Reading Fluency subtests of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS, Good & Kaminski, 2002) from Tier 1 instruction alone. Overall, more students reached DIBELS benchmark scores in the PRIDE group than did in the control (60.8% and 32.9% respectively), although Tier 1 instruction was only one component of these results (i.e., overall benchmark scores were a composite result of Tier 1, 2, and 3 interventions). Also, the effect sizes for overall reading growth across all three tiers of PRIDE group were large (i.e., 1.91, 2.11, and 1.78). In addition, teachers' acceptability of Tier 1 enhancements was high (i.e., range of ratings across three years was 3.3-3.4 on a 4-point scale).

In addition to the support for Tier 1 reading intervention, Fuchs and colleagues have begun to provide additional research support for the effectiveness of Tier 1 intervention in math. For example, Fuchs, Fuchs, and Karns (2001) and Fuchs, Fuchs, and Yazdin (2002) compared teacher implemented math Peer Assisted Learning Strategy (PALS) in kindergarten and 1st grade with traditional basal math instruction on the math achievement of students at risk. To use PALS, teachers organized their entire classrooms into tutoring dyads. Students tutored one another daily for 20-30 min using the PALS math curriculum. The PALS curriculum was designed to improve students' number sense and was composed of activities across four skills: (a) number concepts; (b) number comparisons; (c) adding and subtracting concepts; and (d) adding and subtracting with numerals. Students in PALS classrooms exceeded growth of students in control classrooms (i.e., basal math instruction).

In another investigation of Tier 1 math interventions, Fuchs, Fuchs, and Hollenbeck (2007) trained general educators to use "Hot Math" instruction in problem solving with the whole class. Hot Math instruction consisted of teacher delivered explicit rule and strategy based instruction combined with self-regulation strategies in basic math operations and problem solving specifically designed to promote skill acquisition and generalization. General educators implemented Hot Math two to three times per week for 25 to 40 min per lesson. Preliminary results indicated that after Hot Math instruction the rate of non-responsiveness (i.e., lower than 16^{th} percentile on three researcher created mathematical problem solving measures) for students at risk decreased from 86% to 29%. Achievement results for students in this longitudinal study have not yet been reported.

In summary, studies have indicated that particular teaching variables can predict and explain improved student achievement in Tier 1 intervention including the level of explicitness, active engagement, lesson planning, monitoring of student learning, feedback, and teachers' specialized knowledge. These variables have been found effective as individual practices in previous experimental investigation (Cirino et al., 2007; Foorman et al., 1998; Foorman et al., 2006; McIntosh et al., 2007). In addition, studies have shown that systematic implementations of these combined components (e.g., in the form of core instruction plus enhancements, CWPT, KPALS, Hot Math) as Tier 1 interventions can improve achievement, reduce risk and special education referral for students at-risk, and are acceptable to teachers.

General Education Preparation for Tier 1 Instruction

Despite the success and acceptability of Tier 1 interventions, evidence suggests that classroom practice is often not aligned with current research (Greenwood & Abbott, 2001; Greenwood & Maheady, 2001). It is well documented that general educators often do not use research-based interventions. Schumaker et al. (2002) highlighted the research to practice gap in time sampling observations of 34 general educators across nine schools. Across all observations, researchers recorded no intervals in which teachers used research-based instructional programs or practices such as the ones summarized above. Similarly, in two surveys (Gagnon & Maccini, 2006; Maccini & Gagnon, 2007), less than half of general educators reported implementing various research-based instructional methods in mathematics instruction (e.g., direct instruction, self-monitoring).

There is also evidence to suggest why general educators do not use research-based strategies more widely. Specifically, data suggest general educators do not receive adequate, effective professional development which supports implementation of research-based strategies for students at risk. It is well documented that identifying research based practices is much easier than implementing and sustaining them (Maheady, 2008). One of the most salient barriers to effective implementation is lack of dissemination to teachers through ongoing, high-quality professional development (Greenwood & Abbott, 2001). Previous research has demonstrated the positive relationship between high quality, ongoing professional development and sustainability of practices, and improved student achievement (Joyce & Showers, 1995; Yoon et al., 2007).

Obtaining effective professional development focused on research-based teaching methods to practicing teachers can be difficult for two main reasons, lack of quality and lack of access. First, across several studies teachers have reported that professional development they received did not prepare them to successfully and independently implement research-based practices (Schumaker et al., 2002). Teachers also reported that they often did not sustain use of a research-based practice because of “not having an in-depth understanding of the practice” and “forgetting” to implement or sustain and “needing a refresher” due to many other classroom responsibilities (Klingner, Vaughn, Hughes, & Arguelles, 1999, p. 271). Secondly, teachers reported that few opportunities for professional development targeting students at risk are offered to them (Boardman, Arguelles, Vaughn, Hughes, & Klinger, 2005). Another widespread professional development issue is that the one session inservice is the most popular form of training provided to teachers, despite its documented inability to produce significant changes in teacher behavior or student achievement (Garet et al., 2001).

Similarly, in another survey of general educators, Williams and Coles (2007) found that general education teachers had positive attitudes toward research but did not have access to sources that would show them how to implement specific research based practices. Bursuck, Munk, Nelson, and Curran (2002) also found that general education kindergarten teachers had not received professional development related to teaching reading to students at risk, but were receptive to changing their methods to use a research based, explicit, systematic approach. Other researchers have demonstrated the importance of more opportunities for sustained professional development by demonstrating that more contact hours in training better predict effective implementation of new practices (Banilower, Heck, & Weiss, 2007).

Another substantial barrier in providing general educators with effective inservice training is the lack of quality research on professional development. Existing research allows for little causal inference about what variables are effective in promoting teachers' behavior change. Yoon et al. (2007) conducted a meta-analysis to identify professional development strategies that led to marked improvements in student achievement using only studies that met the IES quality indicators (U.S. Department of Education, 2003). Over one thousand studies examining the impact of professional development were retrieved, but only nine satisfied the quality indicators for research design.

Similar results were found by Knight and Wiseman (2005) in a review of studies of professional development targeting students at risk and by one of the National Math Panel (2007) task groups. The task group's review of the available research and the rigor of this research highlighted the critical need for more and better studies empirically examining the relationship between specific approaches to professional development and teachers' capacity for teaching and their students' learning. In summary, there are limited venues for disseminating to teachers, low quality professional development, and a lack of research on effective professional development. These problems with professional development may inhibit teachers' knowledge and implementation of effective teaching practices.

Effective Professional Development

Although the experimental research base for professional development is limited, the few quality studies that exist have pointed toward certain variables that led to teachers' behavior change. From the nine studies Yoon et al. (2007) found that met IES quality indicators, statistically significant effects on achievement were found from professional development that (a) lasted more than 14 hours, (b) included follow-up support for teachers after the initial training, and (c) was provided directly to teachers by experts (i.e., not a "train the trainer" approach). Similarly, the review conducted by Knight and Wiseman (2005) found high effects on student achievement for professional development models that involved a combination of inservice and follow-up support for teachers.

Federal policy related to professional development supports research as well. According to the No Child Left Behind Act (2001), high quality professional development is (a) sustained, intensive and content focused, (b) aligned with academic standards and assessments, (c) improves teacher content knowledge, (d) improves teachers' use of evidence-based instructional methods, and (e) is evaluated for student and teacher effects. Other research (Maheady, 2008) has revealed that teachers implement and sustain practices that (a) are user friendly (i.e., clear and easy to implement, inexpensive, take little time), (b) include systematic training and ongoing monitoring and (c) are tied to socially important organizational outcomes (e.g., meeting Adequate Yearly Progress).

Among several models of ongoing professional development, experimental research supports the combination of inservice and coaching as one of the most effective methods of promoting and sustaining changes in teachers' instructional practices (Jackson et al., 2006; Jager, Reezigt, & Creemers, 2002; Joyce & Showers, 1995; Yoon et al., 2007).

Coaching is a component of professional development that involves an expert (e.g., trainer, peer) providing individualized support to teachers after an initial training has occurred. The purpose of coaching is to encourage accurate and sustained implementation of new teaching behaviors and to prevent the isolation that often occurs after teachers attend an initial training (Joyce & Showers, 1995). In addition, coaching is intended to provide teachers with "a means of examining and reflecting on what they do in a psychologically safe environment where it is all right to experiment, fail, revise, and try again" (Raney & Robbins, 1989, p. 37).

There are two primary models of coaching identified in the professional development literature: the supervisory follow-up method (Joyce & Showers, 1995) and the in-class feedback method or side-by-side coaching method (Gleason & Hall, 1991). Using the supervisory follow-up method, the coach conducts an observation of a teacher implementing a procedure he or she has recently learned to use in a prior training. During the observation the coach records the presence or absence of particular instructional characteristics the teacher was instructed to use in the initial training. After the lesson, the coach provides descriptive, non-evaluative feedback to the teacher regarding the strengths and opportunities for improvement noticed during the observation. Researchers have found that when supervisory follow-up coaching is provided in combination with an initial professional development opportunity, teaching accuracy improves (Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Kohler, Ezell, & Paluselli, 1999).

The second model of coaching, the side-by-side coaching method, allows teachers to receive in vivo feedback specific to the accuracy of their implementation of new teaching behaviors. In addition, side-by-side coaching allows teachers an opportunity to observe specific teaching procedures demonstrated by an expert (i.e., the coach) with their own students in the context of a real classroom lesson. During a side-by-side coaching session, the coach directly intervenes during the lesson, provides a model and a rationale for the change and then provides additional opportunities for the teacher to teach the same format again with immediate feedback from the coach. Experimental investigations have shown that side-by-side coaching can improve the rate of acquisition of new teaching behaviors (Kohler et al., 2001), the accuracy of teaching behaviors (Kretlow et al., in preparation) and result in longer maintenance of accurate teaching behaviors than the supervisory follow-up method (O'Reilly et al. (1992).

Also, in a large survey comparing professional development follow-up techniques, Blakely (2001) found that the majority of teachers (61%) rated side-by-side coaching most effective in helping them acquire new DI teaching procedures. Teachers rated side-by-side coaching higher than a single demonstration lesson, an after school meeting with a coach, verbal feedback provided by a coach after a lesson, and group training sessions. Teachers preferred the in-class coaching method because it was "hands on," "experiential," and "allowed them to practice techniques to mastery" (Blakely, p. 79). Teachers also reported that the second most helpful form of follow up was in class demonstration (i.e., 22%). In particular, 96% of teachers reported that observing a model provided by an expert helped them understand teaching techniques better. In contrast, less than 10% percent of teachers reported that an after-school training session was sufficient to aid implementation of a new DI procedure.

In a series of single subject studies, Kohler and colleagues examined the effects of inservice and coaching on general education teachers' accurate implementation of reading instruction. All studies used a combination of group inservice and individual coaching sessions. In addition, all researchers in the following studies used either the supervisory follow-up method or the side-by-side coaching method.

First, Kohler et al. (1999) provided a combination of inservice and four coaching sessions to three kindergarten teachers to improve teachers' facilitation of student engagement, quality of teacher-student interaction, and successful completion of academic tasks. Professional development included a half-day inservice to introduce teachers to Rosenshine's (1983) direct instruction model (i.e., review, model, guided practice, corrective feedback, independent practice, weekly and monthly assessment). Following the inservice each teacher received four individual peer coaching sessions using the supervisory follow-up method. During the individual sessions the coach used a Likert scale observation checklist to rate the teacher's implementation of the teaching components, then conducted a post conference with the teacher to give oral feedback related to the lesson. Researchers used a multiple-baseline-across-teachers design to examine the effects of inservice plus coaching on two dependent variables: the number of academic statements to target students and their classmates during partner activities and changes in the type of corrections teachers gave to students (i.e., skills, materials, or interaction/roles/processes). Results indicated an improvement in level and trend for the first two dependent variables across two participants during the coaching phase.

In a second study, Kohler et al. (2001) investigated the effects of a two level inservice plus demonstration and coaching on teachers' accurate implementation of partner reading, collaborative strategic reading, and making words. The first level of inservice was content instruction, the second level was teacher oriented discussion of the content learned in level one. Follow-up demonstration and coaching was provided to each teacher once a week for the entire school year. Researchers used a multiple-baseline-across-teachers design to evaluate the effects of inservice plus extensive coaching on two dependent variables. First, researchers recorded the number of steps completed accurately on a checklist of organization and procedures during lessons. Second, researchers used a 10-s time sampling procedure to code academic subject matter, group structure, instructional mode and function, teacher behaviors, student talk, and student engagement. Results indicated that more improvements in teaching behaviors occurred during coaching than in the initial independent phase. Areas of the curriculum that were not routinely discussed with a coach showed little or no change.

Two additional single subject studies examined the efficacy of coaching on accuracy of teacher behaviors. Morgan, Menlove, Salzberg, and Hudson (1994) examined the effects of coaching on the quality of reading instruction delivered by five low performing general education teachers. During baseline, teachers implemented and video recorded Reading Mastery (Engelmann & Bruner, 1995) instruction with no feedback or training. During the coaching condition, a coach met with each teacher for 30-45 min twice per week to review videotapes of reading sessions. After evaluating the videotapes, coaches assisted teachers in evaluating their own performance on the tapes, compared coach and teacher evaluation of performance, established objectives with teacher for improved performance, and modeled specific behaviors. Coaching occurred privately with each teacher and did not take place during regular instruction (i.e., supervisory follow-up method). Researchers used a multiple-baseline-across-teachers design to evaluate the effects of coaching on effective instructional trials defined by six teaching behaviors: (a) effective cues (e.g., what word, together, get ready); (b) effective pauses (e.g., 1 to 2 s pause between question and signal); (c) effective response to errors in unison responding (e.g., staggered entry of student response); and (d) effective response to student errors (e.g., model-lead-test, model-test). Two other dependent variables were measured including the ratio of specific praise statements to total praise statements and the rate of student responses. All four teachers' percentage of correct instructional trials increased as a direct result of coaching as indicated by a change in trend during intervention.

One other study also examined the effects of coaching on improvements in teaching behaviors and extended the evaluation to student outcomes. Maheady, Harper, Mallette, and Karnes (2004) evaluated the effects of inservice plus coaching on the accuracy of general educators' CWPT implementation. Teachers attended a 2 hr workshop conducted by researchers and were given a procedural implementation manual. The workshop included relevant theory, explicit instruction in spelling CWPT components using 15 min video clips on CWPT, and role play with positive and corrective feedback. Following the workshop, individual support was provided to each teacher by researchers and included feedback about teacher performance and occasional modeling of tutoring procedures. In contrast to the other coaching studies, Maheady et al. did not implement a prescribed number of coaching sessions. Teachers received coaching sessions until they completed 85% of teaching procedures correctly during CWPT lessons (i.e., measures through direct observation). The dependent variable was measured by a three category 36 item procedural checklist including measures of teachers' use of CWPT materials, teaching procedures, and students' tutoring behavior. In addition, researchers assessed student academic outcomes including the percentage of words spelled correctly and the amount of normalized gain made above pretest performance, both administered orally by teachers. Results of a quasi-experimental pretest-posttest examination of teacher and student data indicated that teachers with coaching learned to use CWPT with high degrees of accuracy (mean 88%) and eight of ten teachers reached pre-established training criterion in just over 2 hr of total training. Mean pretest and posttest spelling grades and normalized gain scores increased from averages of 69% in posttest averages of 94%, which constituted a 25% increase in students' accrual spelling averages and two to three letter grade improvements.

Experimental studies cited support the findings of both educational policy (NCLB, 2001) and reviews of literature (Knight & Wiseman, 2005; Yoon et al., 2007) related to the efficacy of an inservice plus coaching model of professional development. Kohler and colleagues (1999; 2001), Morgan et al., (1994) and Maheady et al., (2004) all found that coaching improved the accuracy of specific teaching behaviors in reading and spelling instruction. In addition, one study extended the positive effects of coaching to student achievement (Maheady et al.). All of the studies demonstrated that coaching was effective in helping teachers develop skills to implement explicit instruction methods while following a highly structured program (e.g., Direct Instruction, CWPT).

Explicit Instruction for Preventive Intervention in Math

Explicit instruction is critical to the success of preventive instructional efforts. In math, explicit instruction involves teacher demonstrated step-by-step plans for answering a question or solving a problem. The teacher demonstrates a specific plan for a set of problems then directs students to use the same procedures independently to solve the problem (Gersten, 2007). The Math Panel task group on instructional practices (Gersten et al., 2007) concluded that students with learning difficulties should receive explicit instruction on a regular basis. Explicit instruction was deemed essential for building proficient computation and subsequent translation of word problems into mathematical equations and solutions. The Math Panel also recommended that explicit instruction should also be used to ensure that students gain foundational skills and conceptual knowledge necessary for understanding the math they are learning at their assigned grade level.

Gersten et al. (2007) found nine studies that looked at the effect of explicit instruction for students who were at risk or had learning disabilities, which also met high quality methodological criteria (e.g., randomized control trials, strategically matched quasi-experimental studies). Studies were examined across three instructional areas: problem solving, computation, and generalization. Combined results revealed significant effects for explicit instruction across all three instructional areas, with large effect sizes ranging from .8 to 1.3, indicating explicit instruction is an effective means for improving achievement in math.

In another review of evidence-based practices in math, Miller and Hudson (2007) also recommended teachers use explicit instruction as a framework for teaching math across strands (i.e., numbers and operations, algebra, geometry, measurement, data analysis and probability) and the three knowledge areas (i.e., conceptual, procedural, and declarative). Miller and Hudson's recommendations were based on the positive results of explicit math instruction on student achievement derived from 30 years of experimental research with students with mild disabilities. Miller and Hudson defined explicit instruction as step-by-step instruction that includes direct teacher demonstration followed by a gradual shift of responsibility from teacher to student for solving problems (i.e., scaffolding). Miller and Hudson further defined explicit math instruction as direct teaching sequences that include new material presented in small pieces using examples and non-examples followed by continuous practice until the skill is mastered.

Features of Direct Instruction Presentation

Although step-by-step instruction is one of the critical components of effective math instruction for students at risk, there is also evidence to suggest it is not effective enough in isolation to produce substantial student achievement gains in math (Dixon, Carnine, Lee, Walin, & Chard, 1998). Specifically, in a review of high quality experimental math research, Dixon et al. found that two-step explicit instruction in which teachers demonstrated how to compute or solve a problem in math while students watched passively then provided an opportunity for independent practice did not produce substantial gains in student achievement. A three-step explicit instruction model including an intermediate step in which teachers actively engaged students by asking many questions to check for understanding was more effective. During the intermediate step teachers also provided feedback and correction (e.g., additional explanations when students had difficulty). In the third step of the three step explicit model teachers assess students' independent skill in applying knowledge practiced in step two to untaught problems.

Dixon et al. (1998) demonstrated that other components of instruction including active engagement and scaffolded opportunities to practice new skills when combined with explicit instruction are more effective than explicit instruction alone. One specific model of explicit instruction, DI, also combines active engagement and scaffolding with explicit instruction (Watkins & Slocum, 2002). DI has been deemed highly effective by a sizeable body of experimental research (Iver & Kemper, 2002; Przychodin-Havis et al., 2005; Simonsen & Gunter, 2001).

DI programs are highly effective teaching curricula which include: (a) systematic program design in which the critical concepts, rules and strategies of a content domain (i.e., "big ideas") are taught through careful selection, sequencing, and implementation of instructional trials; (b) instructional organization (e.g., small, skill-based groups, curriculum based assessment, and (c) highly engaged teacher-student interactions (e.g., choral responding, brisk pace; Watkins & Slocum, 2004). A substantial body of research supports the use of DI programs to improve the achievement of at-risk learners across content domains, as indicated by high effect sizes found in multiple research reviews (Adams & Engelmann, 1996; Przychodzin-Havis, 2004; Przychodzin-Havis et al., 2005). In addition, many studies have demonstrated that DI programs are effective when delivered by elementary general education teachers to students who have disabilities or are at risk (Przychodzin-Havis, et al.)

Although DI programs are highly successful, teachers, especially general educators who are primarily responsible for instruction based on specific academic content standards, may not have the autonomy to select DI programs to deliver core curricula. It may be possible, however, for teachers in general education settings to embed effective components of DI in lessons as strategies to improve student engagement and achievement. Component analyses of DI programs have helped identify the specific practices that contribute to its effects on student achievement (Watkins & Slocum, 2002). Three salient components of DI are model-lead-test (MLT), systematic error correction, and unison responding. These components have been found effective instructional procedures when embedded in DI programs and also in individual interventions.

One component, MLT is a teaching procedure in which "…the teacher first demonstrates how to do the skill so that the students have no difficulty understanding exactly what the new skill looks like" (i.e., model). Then, "the teacher practices the skill with his or her students until they are able to do it without him or her" (i.e., lead). Lastly, "the teacher monitors the students as they do the skill independently" (i.e., test, Bursuck & Damer, 2007, p. 16). Studies have shown that using a MLT format to introduce new concepts to students, especially students at risk, promotes mastery of new content (Hollingsworth & Woodward, 1993; Idol-Maestas, 1995; Parks, Weber, & McGlaughlin, 2007).

Hollingsworth and Woodward (1993) taught 37 secondary students with learning disabilities health facts and concepts which they then applied to problem solving exercises presented through computer simulation games. Students who were taught through MLT an explicit strategy for solving problems performed better on two generalization measures than students given only supportive feedback and encouraged to use their own strategies.

Idol-Maestas (1995) found similar results using a MLT format to teach daily story comprehension to five students with LD or low reading achievement. Results of a multiple-baseline-across-participants design showed that the MLT procedure led to students answering more story questions correctly and completing post-reading story maps more correctly. In a similar multiple-baseline-across-participants design, Parks, Weber, and McGlaughlin (2007) replicated the effectiveness of MLT on the name writing skills of preschool students with developmental delays. During baseline students had difficulty with letter identification and formation. Intervention included a MLT procedure to teach students to write their names that was eventually faded. When MLT was faded, both participants' handwriting remained at an improved level. Results of all three studies suggest that MLT is an effective procedure for supporting independent mastery of new content.

MLT is also used in a second DI component, systematic error correction. Using systematic error correction, a teacher immediately provides a brief re-teaching using MLT or variation (i.e., model-test, lead-test) after students make an error. Studies have indicated the efficacy of systematic error correction in promoting accuracy of academic responses (Barbetta et al., 1993; Barbetta et al., 1994; Gettinger, 1993; Nelson et al., 2004; Van Houton, 1993). For example, Barbetta et al. (1993) used an alternating treatments design to compare the effects of error correction (i.e., model-test) with and without a required student response (i.e., model only) on sight word acquisition. In both conditions, error correction procedures were followed by specific praise, either for attending to the word or providing the correct response. Results indicated a functional relationship between the response condition and the number of words read correctly on same day and next day tests. In addition, the number of correct student responses during sight word instruction was higher in the response condition than in the no response condition.

In a subsequent study, Barbetta et al. (1994) extended the investigation of error correction to compare immediate and delayed correction. In one condition, researchers corrected student errors on sight words immediately after they occurred, in the other condition researchers corrected all errors made in a set of sight words after the student had read the entire set. Results indicated a functional relationship between immediate error correction and improved sight word acquisition and maintenance. Other research also supports the combination of immediate, systematic error correction in the areas of reading (Nelson et al., 2004), math (Van Houten, 1993), and spelling (Gettinger, 1993).

MLT and systematic error correction are often combined with a third component of DI, unison responding, to maximize students' active responding during instruction. Unison responding occurs when all students in a group respond together at the same time to a curricular-related antecedent (Bursuck & Damer, 2007; Heward, 1994). Unison responding can occur verbally (i.e., choral responding) or physically (e.g., all students hold up response cards or signs, hold up number of fingers, thumbs up) after a teacher presents a question, problem, or item.

Research supports the use of unison responding alone and in combination with MLT and systematic error correction. In the first examination of unison responding, Sindelar et al. (1986) compared the effects of ordered and choral responses during small group reading instruction on the sight word acquisition of eleven 8-10 year old participants with mild disabilities. Using an alternating treatments design, 10 sight words were taught in two conditions: an ordered response condition and a choral response condition. During the ordered response condition, researchers called on individual students in a prescribed order to read a sight word. During the choral response condition, researchers prompted all students to respond in unison when each sight word was presented. Results indicated a functional relationship between choral response and a higher number of words learned across all three groups. Kamps (1994) and Godfrey, Grisham-Brown, Schuster, and Hemmeter (2003) found similar results with children with autism and preschoolers with attending difficulties. Choral responding was more effective than individual turns (hand raising) in improving participation and on-task behavior.

Research with response cards produced similar results. Gardner et al. (1984), Narayan et al. (1990), Sweeney et al. (1999), Christle and Schuster (2003), Davis and O'Neil (2004) and Wood, Mabry, Kretlow, Lo, and Galloway (in press) all found response cards more effective than individual turns (e.g., handraising, ordered individual turns) in improving a variety of dependent variables including participation, on-task behavior, and achievement in various academic areas across grade levels. Similarly, in a review of research on response cards, Randolph (2007) found large, statistically significant effect sizes for response cards compared to hand raising for test achievement, quiz achievement, participation, and reduction in intervals of disruptive behavior. Randolph also found that response cards were most effective for students with learning difficulties and that students preferred response cards to hand raising.

Combining DI Presentation Features

DI programs have combined MLT, systematic error correction, and unison responding in published programs and individual components have been taught to teachers as a way of strengthening core instruction. Bursuck et al. (2004) combined MLT, systematic error correction, and choral responding in the form of “enhancements” and embedded them in existing reading curricula for Tier 1 instruction. Bursuck et al. trained kindergarten, 1st grade and second grade teachers to add enhancements to two reading programs (i.e., Harcourt Trophies or Open Court) using structured teaching formats for each lesson. Although enhancements were not the only support provided (additional enhancements, supplemental instruction and data-based decision making were also used), enhancements were one component of instruction that led to improved reading for students at risk for failure across three grades. In addition, teachers had high fidelity with the enhancements and reported that enhancements helped students at risk and were easy to use.

Given the positive effects of DI programs, investigations of individual components of DI and experimental effects of enhancing core instruction with a combination of specific DI components, Kretlow, Wood, and Cooke (submitted) conducted a study to train teachers to effectively select and use the components. Kretlow et al. extended the Bursuck et al. (2005) study by training kindergarten teachers to use three enhancements (i.e., MLT, systematic error correction, and unison responding) during whole class math instruction. The study examined the effects of two levels of training on kindergarten teachers' accurate delivery of enhancements. Teachers were trained to use choral responding and response cards in combination with a MLT procedure for introducing new skills and correcting errors. A multiple-baseline-across-teachers design was used to examine the effects of a half-day workshop and coaching on teachers' accurate delivery of instructional units within math calendar lessons. Results indicated that the workshop followed by coaching was more effective than a workshop alone in increasing teachers' instructional accuracy.

Measuring Change in Teacher Behavior

With the exception of a few studies (Kohler et al. 1999; Kohler et al. 2001, Kretlow et al., in preparation, Maheady et al., 2004, Morgan et al., 2004), changes in teaching behaviors after professional development are rarely measured behaviorally in a way that allows causal inference about the effectiveness of training (Yoon et al., 2007). Researchers have suggested that instructional effectiveness should be gleaned from objective observation, not subjective evaluation (Ross, Greer, & Singer-Dudek, 2005). One challenge in evaluating effects of professional development on teachers' instructional behaviors is identifying a measure that is both practical for an applied study and sensitive to observable change. It is equally difficult to identify a singular measure that encompasses critical features of effective group instruction in natural teaching contexts (i.e., using existing curricula). The majority of measures developed for observing teacher change were designed for use with highly structured programs (e.g., DI, CWPT). For example, Morgan et al. (1994) used an observable measure to determine the extent to which teachers implemented the components of DI correctly during a scripted DI program, Reading Mastery lessons; so the outcome may not be easily transferrable to implementation of non-scripted programs or non-DI programs.

The instructional trial is one objective unit of analysis for empirical measurement of teaching behaviors (Heward, 1994). An instructional trial is essentially a three-term contingency applied in a teaching and learning context. First, an antecedent (i.e., question or direction given by the teacher) is presented to student(s). Second, the antecedent is followed by students' behavior (i.e., students' response or answer to the question posed). Third, the students' response is followed with feedback from the teacher (i.e., correction or praise which serves as the consequence).

The instructional trial has been applied in previous research to create an empirical way of measuring the quality of teachers' instruction. Greer and colleagues (1985; 1991; 1992; 1996) created a system of measurement (i.e., the Teacher Performance Rate and Accuracy Scale; TPRA) based on interlocking instructional trials termed "learn units" to record the occurrence and quality of antecedents, behaviors, and consequences during instructional events. The learn unit was developed using critical behavioral components for each contingency identified in previous research. A correct learn unit is composed of: (a) an unambiguous antecedent; (b) an active student response; and (c) a correction or reinforcement for student responses (Ross, Greer, & Singer-Dudek, 2005). Using the TPRA, each component of the learn unit that occurs during instruction is scored (i.e., teacher antecedent, response, and correction or reinforcement); then the total learn unit is given a discrete score (i.e., correct or incorrect). The total number of correct learn units per minute is the final outcome score on the TPRA. The rate of student responses is also scored. Research using the TPRA in applied settings indicated a positive correlation between higher TPRA scores and student learning, functional relationships between higher TPRA scores and correct students responses, and higher numbers of correct student responses with the consequence component than with the opportunity to respond alone (Ross et al).

For example, Ingham and Greer (1992) compared the effects of a general observation procedure and a TPRA observation on teacher performance and student responding. Researchers conducted 10 to 20 min observations of teachers using task-analysis, discrete trials or incidental trials with students who had significant cognitive disabilities. During baseline (i.e., general observation condition), a supervisor repeatedly observed the teacher using a descriptive measure of teaching effectiveness. During the sessions the supervisor calculated the number of correct learn units per min and the total number of correct student responses per min during each session. After each observation, the supervisor met with the teacher to provide general, nonspecific verbal feedback about the session including praise, comments on instruction and comments on student behavior. During intervention, the supervisor used the TPRA to observe teachers. Supervisors also calculated the number of correct student responses per session. After each observation, the supervisor gave specific, verbal and written feedback to the teacher, including showing them data on their instructional accuracy and the accuracy of student responses. A multiple-baseline-across-teachers showed that teachers had higher accurate learn units per min during the TPRA condition than during the descriptive observation condition. Similarly, a multiple-baseline-across-students design showed that the rate of correct student responses was also higher in the TPRA condition.

In a second study Ingham and Greer (1992) examined the extent to which teachers maintained teaching accuracy across weeks and the extent to which correct student responses generalized to the remainder of the day after an observation. During baseline researchers conducted general observations with non-specific feedback once weekly. During intervention researchers conducted observations using the TPRA and gave specific feedback on teacher and student performance once weekly. Researchers collected data on the rate of accurate learn units and the rate of correct student responses during weekly observations in both conditions. In addition, teachers recorded the rate of correct and incorrect student responses during the remainder of instruction on observation days in both conditions. A multiple-baseline-across-teachers design showed a functional relationship between weekly TPRA observations and higher rates of correct learn units. A multiple-baseline-across-students design also showed a higher rate of correct student responses during the TPRA condition. Results of TPRA studies suggest that giving teachers specific feedback related to each contingency within a learn unit can improve instructional accuracy and student performance in classroom lessons.

Although the learn unit is sensitive to the discrete teaching and learning events that occur within an instructional trial, it may not be sensitive to other critical features of effective instruction. More specifically, the learn unit does not include a measure of the use of unison responding or scaffolding. Using the learn unit as a measure, the three term interaction between a teacher and one student would be rated the same as an interaction between a teacher and 10 students using choral responding. No additional credit is given to a teacher who engages all students. In the same way, no additional credit is given to a teacher who uses appropriate scaffolding for the acquisition stage of student learning. A teacher who provides a model for the same skill in 10 trials would receive the same score as a teacher who systematically prompted students toward independent mastery of the skill.

Kretlow et al. (submitted) expanded the concept of the learn unit to the instructional unit, which is a measure of the quality of interlocking three term contingencies that includes quality ratings of unison responding and scaffolding, found to be critical measures of teacher quality in other previous research (Morgan et al., 1994). Similar to the learn unit, each contingency within a teaching event is rated for occurrence and quality: (a) teacher antecedent; (b) active student response; and (c) correction or reinforcement. However, in order for an instructional unit to be counted accurate, the teacher must not only accurately implement each part of the three term contingency, he or she must do so using unison responding and an appropriate level of scaffolding (e.g., provide a model for new skills, no model for review skills). In addition, instructional units may be composed of more than one interlocking three term contingency. An instructional unit begins with a teacher antecedent related to an identified skill and does not end until a correct unison response is achieved for that skill or when the teacher provides an antecedent related to another skill. For example, an instructional unit could include three interlocking three term contingencies all following a prompt by the teacher to count by 2s to 100. If students made errors during the first trial, the subsequent three term contingencies related to the errors for counting by 2s would be rated within the original instructional unit. Figures 1 through 5 depict correct and incorrect instructional units with one three-term contingency and multiple interlocking three term contingencies. Although Kretlow et al. examined important additional instructional variables within a behavioral measure of teaching effectiveness, no effects on student achievement were examined. Future research with the instructional unit needs to investigate the effects of observable changes in teachers' behaviors on the academic outcomes of students, particularly students at risk.

Figure 1. Correct instructional unit with one three-term contingency.

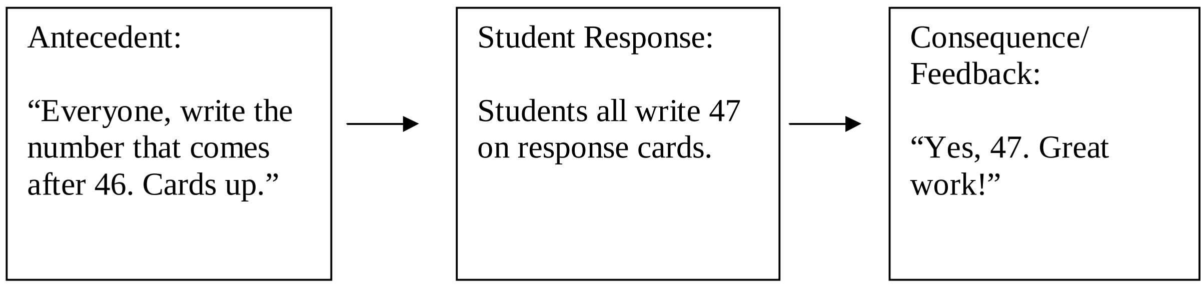


Figure 2. Incorrect instructional unit with one three-term contingency.

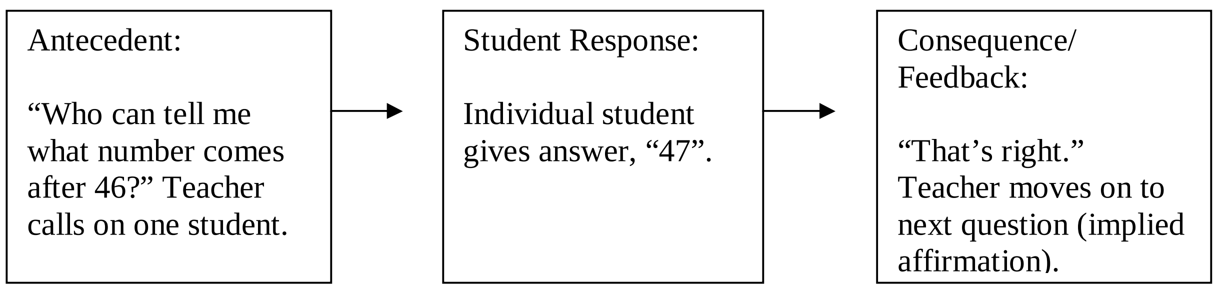


Figure 3. Incorrect instructional unit with one three-term contingency.

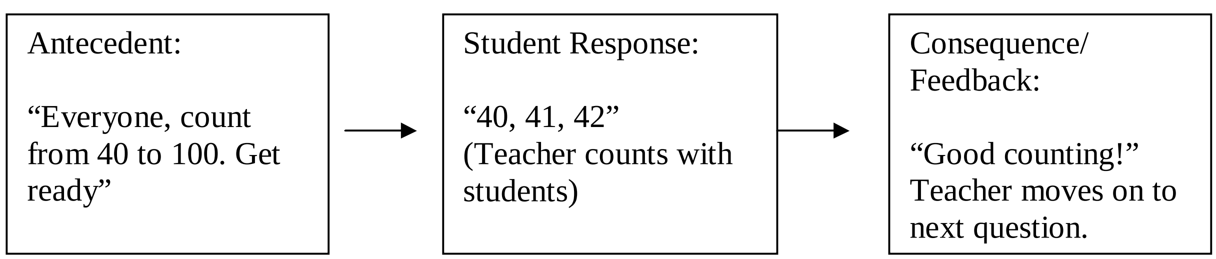


Figure 4. Correct instructional unit with a series of three-term contingencies.

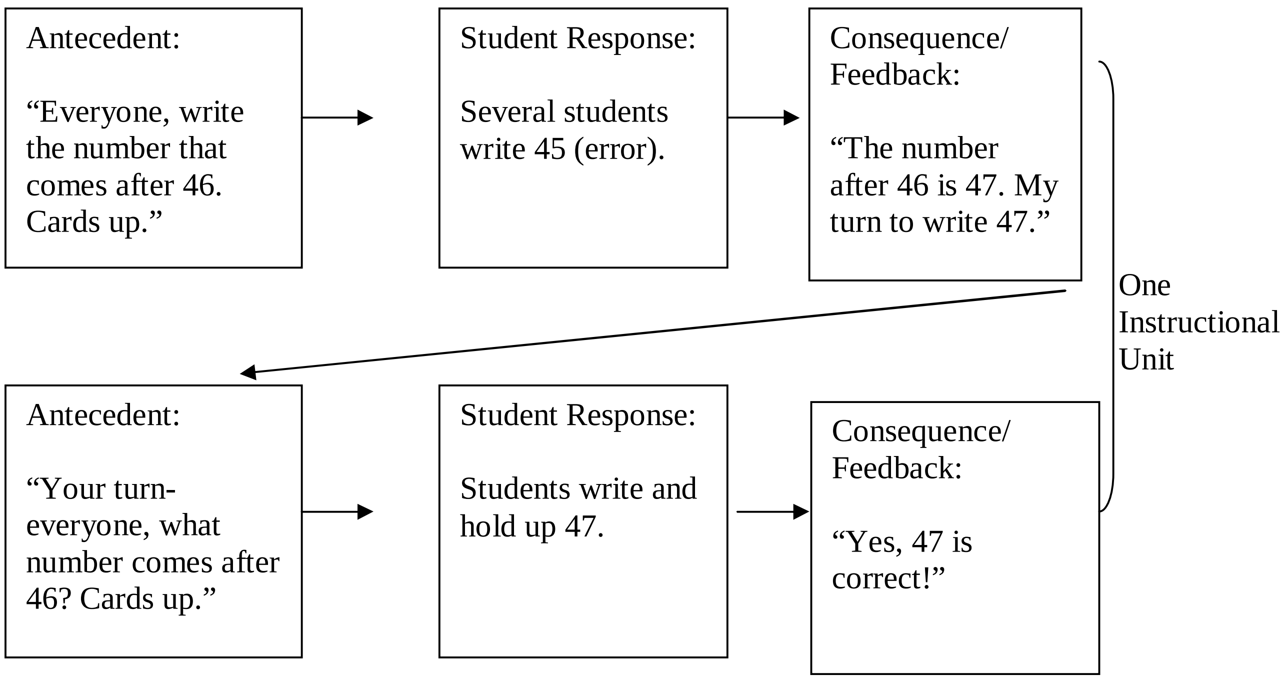
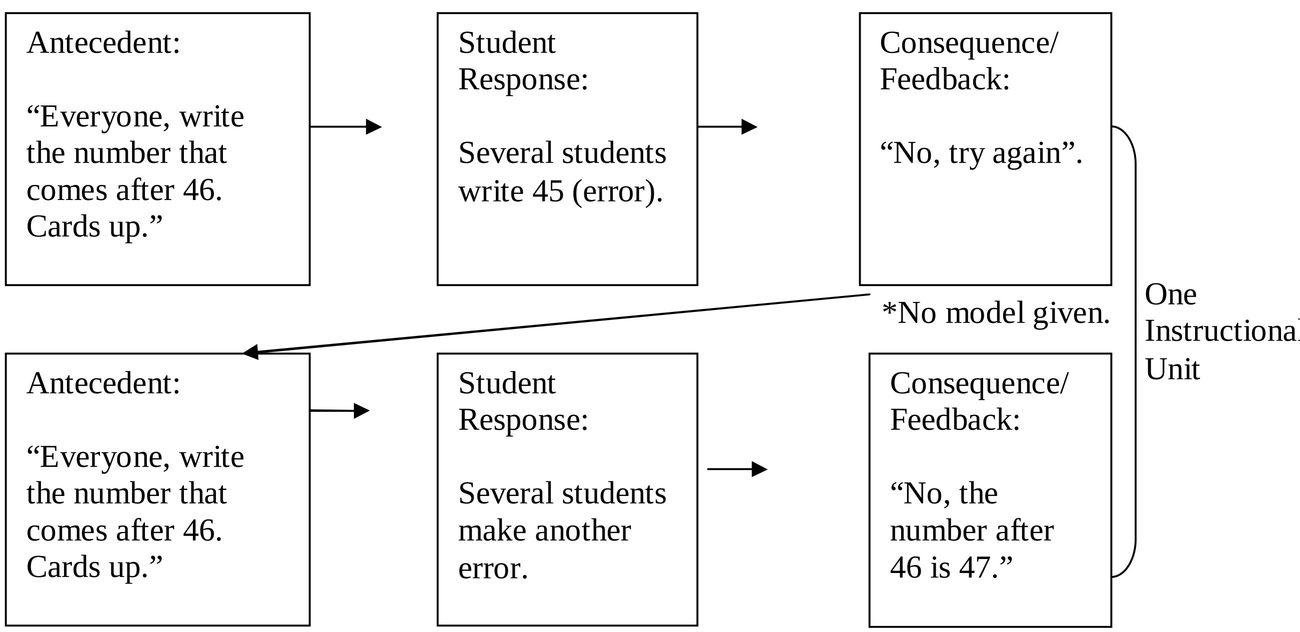


Figure 5. Incorrect instructional unit with a series of three-term contingencies.



Summary of Research

Early intervention efforts such as RTI predicate the need for high quality Tier 1 instruction in general education classrooms. Previous research has established that certain teacher variables significantly predict student outcomes in early intervention. These variables include explicit instruction in combination with active student engagement, scaffolding, frequent monitoring during instruction, and feedback for student responses. The literature on Tier 1 reading instruction which includes the components of effective instruction mentioned above is well established; however, the research on Tier 1 math instruction is only emerging. More research needs to be conducted on the efficacy of Tier 1 math interventions.

Although high quality Tier 1 instruction is critical, the lack of high quality professional development opportunities for general educators prevents the widespread implementation of effective teaching strategies. Teachers lack access to professional development related to effective instruction for students at risk and when they do receive professional development of this type it rarely aligns with the components of effective training indicated in the literature (Garet et al., 2001).

One model of professional development that holds promise for training teachers to implement explicit instruction is inservice combined with coaching. A few studies have shown that this combination can lead to improved teaching accuracy across content areas (Kohler et al.,1999; Kohler et al., 2001; Morgan et al., 1994) and improve student achievement (Maheady et al., 2004). More studies are needed to replicate and extend the effects of inservice and coaching to math; specifically examining changes in instructional accuracy and student outcomes. One challenge for researchers is selecting a sensitive, observable measure, as opposed to a subjective, self-reported, or general measure, to examine changes in instruction directly related to training. Using the instructional trial as the basic unit of analysis has shown promise. More investigations of changes in teacher behavior using quality ratings of interlocking three term contingencies (e.g., learn unit, instructional unit) may lend additional support for critical variables related to training in-service general educators to improve instruction for students at risk.

CHAPTER 3: METHOD

The purpose of this study was to investigate the effects of enhancement training on 1st grade general education teachers' percentage of instructional units implemented correctly in Calendar Math (CM) and the extent to which teachers generalized the correct implementation of instructional units to an untrained area of math, Numeracy/Problem Solving (N/PS). This chapter will present the methods that were used to investigate the research questions. Specifically, the chapter will present information about participants, research design, data collection, and intervention procedures.

Participants

Teachers. Four 1st grade teachers were purposefully selected to participate in a single-subject design examining the impact of enhancement training on the percentage of instructional units implemented correctly. Teachers were invited to receive enhancement training if they (a) were nominated by the principal, (b) taught a DI program for at least one entire school year, (c) were the primary classroom instructor for the duration of the study, (d) responded positively to questions about teaching DI programs, and (e) provided written consent to participate, required by the UNC Charlotte Institutional Review Board. It was important that the teachers selected for the study structured math lessons similarly to one another, both in content (e.g., skip counting, addition, telling time) and structure (e.g., daily use of whole group vs. rotating centers). The structure of lessons was important for consistent comparison across tiers in a single subject design. All students in the classrooms of the four teachers selected for enhancement training received daily instruction (i.e., approximately 28-30 students per class).

Kristy held a bachelor's degree in Elementary Education, and had four years of teaching experience, 2 years in 1st grade and 2 years in kindergarten. She had previously taught DI reading and language programs for two years (i.e., Reading Mastery, Language for Learning). Megan also held a bachelor's degree in Elementary education. She had 2 years teaching experience in 1st grade and taught Reading Mastery for one year. Beth held a bachelor's degree in Elementary Education and a master's degree in Literacy. She had 9 years experience teaching 1st grade and 5 years experience teaching DI reading and language programs. Jade also had a bachelor's degree in Elementary Education and 9 years teaching experience (2 years in 1st grade). Jade had 6 years experience teaching Reading Mastery.

Setting

Teacher training. All study activities took place at an elementary school in a suburban North Carolina school district. The study occurred during the second year of the school's operation. The school housed approximately 800 students, of whom 63% were Caucasian, 18.6% were Hispanic, 15.7% were African American, 2% were Asian and less than 1% were other. The group inservice occurred in the special education teacher's classroom. Individual preconferences, coaching and feedback sessions occurred in teachers' respective classrooms. No students or other staff were present during the preconferences or feedback sessions. The coaching sessions occurred during regularly scheduled math lessons, therefore students were present and on occasion other school personnel were present (e.g., paraprofessionals).

Daily instruction. Daily math instruction was provided in two separate sessions using the Everyday Math program (EDM, Bell & Bell, 2004) and the North Carolina math content standards (North Carolina Department of Public Instruction, 2008). In the first session, teachers provided math instruction in the context of the calendar as prescribed by EDM. Daily instruction included a routine of math activities related to the number of days students had attended school. For example, each day students orally counted days in school by 1s, 2s, or 5s, represented the number of days in school in ones, tens, and hundreds by placing straws in the correct pocket to demonstrate place value, count coins a variety of ways for the number of days in school (e.g., 1 nickel and 4 pennies for 9 days in school), and represented the number of days in various ways (e.g., tally marks, write the number, addition and subtraction sentences). Students also completed patterning activities during CM. Pictures of objects were imposed on dates on a calendar posted on the board. The teacher asked the students to complete the pattern by stating what object will come next today, tomorrow, or in several days. Students also stated the date, the months of the year, days of the week, and graphed how many days they have experienced a particular weather pattern (e.g., rainy, sunny).

During the second daily session, teachers provided instruction related to numeracy and problem solving concepts using portions of EDM in a whole class setting. Teachers typically presented this instruction by projecting a workbook page on the Smart Board as students sat at their desks in groups of five. First, the teacher demonstrated a procedure (e.g., rolling dice, counting the dots, tallying the number of times she rolls a particular number). Next, the teacher demonstrated the procedure again while asking the students guiding questions about the task (e.g., “how many tallies will I make now?”).

Finally the teacher instructed students to complete the remaining problems in their individual workbooks independently, or with the help of the classmates at their table.

Researcher

The researcher was a special education doctoral student who served as the primary interventionist and data collector for this study. She held elementary education and special education teaching licenses and taught special education at the elementary level for 3 years. In addition, she was a trained DI coach and had one year experience using side-by-side coaching with teachers who were implementing various DI reading programs. The researcher also co-taught a graduate course on instructional design for students with diverse learning needs that covered several content areas, including math which is the focus area for this study.

Second Observers

Two graduate students who were also special education teachers at the school served as the second observers for the primary dependent variable and the generalization measure across all teachers and phases of the study. In addition, the graduate students collected procedural fidelity data by observing the researcher during inservice and listening to recordings of preconferences and coaching sessions. Both data collectors had bachelor's degrees in special education were currently enrolled as master's degree students in UNC Charlotte's Special Education Program.

Dependent Variables

Teacher measures. Two dependent variables were included in this study. The first dependent variable measure was the percentage of instructional units implemented correctly during 10 min trained segments of daily CM lessons. The second dependent variable measure (i.e., generalization) was the percentage of instructional units implemented correctly during 10 min untrained segments of math lessons (e.g. numeracy and problem solving instruction not delivered during CM) collected three times per week. Percentage of correct units was used because the number of opportunities for teachers to use correct unit varied across sessions and teachers. For example, if teacher A had 15 correct units and teacher B had 8 correct units, it would appear that teacher A had more correctly implemented instructional units. However, using number correct could have been deceiving if teacher A had 30 opportunities (i.e., 50% correct) and teacher B had 8 (100% correct). Therefore, percentage allowed for a more sensitive comparison across teachers. The percentage of correct instructional units was calculated by tallying the total number of correct units and dividing that number by the total number of instructional units that occurred during the 10 min segment. The mean number of opportunities to respond (given to the whole class) presented in 10 min across all conditions and phases of the study were also collected as a descriptive measure of active student responding provided during lessons.

A correct instructional unit was defined as a single three-term contingency or a series of three-term contingencies that began with a correct teacher provided antecedent and ended with a correct independent unison response. Teacher provided antecedents were counted correct if (a) the teacher provided a model for a new skill then provided an opportunity for independent practice using a test (e.g., My turn to count by 2s. 2, 4, 6, 8, 10. Your turn to count by 2s.") or (b) the teacher provided a model for a new skill, lead the students in a response, then provided an opportunity for independent practice with a test (e.g., "My turn to count by 2s. 2, 4, 6, 8, 10. Count by 2s with me, 2, 4, 6, 8, 10. Your turn to count by 2s."), or (c) if the teacher provides an opportunity for practice of a review skill using a test with no model (e.g., "Get ready to count by 2s to 10"). If a teacher used a MT procedure or a MLT procedure, either would be counted correct, even if one was preferred over another.

If students made errors when practicing new or review skills, an instructional unit was only counted correct if the teacher provided a model-test, model-lead-test, or lead-test error correction procedure that resulted in a correct independent unison response. An instructional unit was counted correct if the teacher provided any of the three types of error correction immediately after students made a mistake. In a model-test correction, the teacher immediately stopped the students after detecting an error by presenting a model, then an opportunity for a group response using a test. For example, if the teacher presented the question "What shape is this?" and students responded with an incorrect answer, the teacher would then say "This shape is a triangle. Everyone, what shape is this?" In a model-lead-test correction, the teacher also immediately stopped the students after detecting an error by presenting a model, then an opportunity for students to repeat the correct answer with her, followed by a group response using a test. For example, if the teacher asked students to count by 5s to 20 and students made an error, the teacher might say "My turn to count by 5s. 5, 10, 15, 20. Do it with me 5, 10, 15, 20. Now by yourselves." Finally, in a lead-test error correction, the teacher corrected student errors by using a series of steps or questions appropriate to the task then provided an opportunity for a group response using a test. For example, if the teacher asked students "What is 2 plus 2?" and students responded with an incorrect response, the teacher could have said "Let's count up together starting at 2. 2, 3, 4. What is 2 plus 2?" (i.e., teacher and students together). Then the teacher provided another opportunity for an independent group response (i.e., “Your turn again. What is 2 plus 2?”).

Only instructional units in which a group response was prompted were counted as correct. Instructional units that occurred with individual students (i.e., individual turns) were counted incorrect if the teacher did not first present the question or direction to the whole group. Individual turns related to a skill that were presented immediately prior using a unison response were not counted correct or incorrect (i.e., teachers were encouraged to use individual turns after first presenting the item or question to the entire group).

The researcher and the second data collector used the form in Appendix A to score the 10 min audio-recorded segments of CM and N/PS. When scoring a recording, the data collectors used the following procedure: (a) wrote a short descriptor of the question or direction given by the teacher in the activity column; (b) marked whether a prompt for a group response occurred, and which type of group response was used (i.e., choral responding or response cards); (c) marked the level of scaffolding used by the teacher (e.g., MLT, MT, T); (d) marked the occurrence of any student errors; (e) marked the occurrence and type of error correction used by the teacher (i.e., MLT, MT, LT). Error correction was counted as part of the original instructional unit, as were individual turns and repetitions of the same question or direction (e.g., teacher asks students to count by 2s twice in a row). Next, the data collectors determined if the entire instructional unit was correct or incorrect, based on the rules listed in the preceding paragraph. Lastly, the data collectors calculated the percentage of correct instructional units that occurred during the 10 min recording. Figure 6 below depicts a sample scored data collection sheet.

Figure 6. Sample data collection form.

Data Collection Form

Date 1/13 **Data Collector:** Allison **Teacher:** June Math Calendar or N/PS **Data Collection Duration:** 10 min. **Lesson Description:** Reviewing Time

Teacher Direction/Question	Opportunity for Group Response	Model, Lead, Test	Errors: Task Error, Unison Error, No Error	Controlled Response	Notes
What does the short hand tell you?	CR RC	M L T	T U No	Y N	C
What does the long hand tell you?	CR RC	M L T	T U No	Y N	C
What time is on the clock?	CR RC	M L T	T U No	Y N	C
2 hours later	CR RC	M L T	T U No	Y N	Students confused about hour/min hand C
Error correction	CR RC	M L T	T U No	Y N	Demonstrated moving hands of the clock
What time is it now?	CR RC	M L T	T U No	Y N	Only a few students responded, no signal I
Error Correction	~~CR RC~~	M L T	T U No	Y N	
Write the time that is 3 hours from now	CR RC	M L T	T U No	Y N	White boards/erasers
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	

4 (total number of correct instructional units) divided by 5 (total number of total instructional units) x 100 = 80%

Interobserver reliability. Interobserver reliability was collected across all phases of the study. A second observer listened to 25% of all recordings across teachers and recorded instructional unit data on the primary dependent variable (i.e., CM lessons) and the generalization variable (i.e., 10 min of untrained numeracy and problem solving lessons). The second observer's results were compared with the researcher's results using an item-by-item method (i.e., compare the number of correct and incorrect instructional units). The experimenter calculated the reliability coefficient by dividing the number of agreements by the sum of agreements and disagreements. The researcher trained the second observer to use the data collection system by modeling the accurate data collection method with one 10 min recording, then provided opportunities to score recordings together until 90% agreement was reached across three separate teachers.

Social validity. Social validity data were collected to assess the impact and feasibility of the enhancement training reported by teachers. First, teachers will be given a written questionnaire which included open-ended and closed (i.e., Likert) items evaluating each individual enhancement and training component (see Appendix B for Social Validity Questionnaire). The second data collector distributed the questionnaire to the four teachers and directed them to complete it anonymously.

In addition to the questionnaire, the researcher intended to collect student data to evaluate the extent to which students' risk levels change during the study by reporting the number of students per class who changed risk level on AIMSWeb TEN and Computation assessment (Edformation, 2008) after enhancement training ended. However, AIMSWeb math assessments were given in early January, when only one of four teachers had received coaching.

Experimental Design

The impact of inservice, coaching, and feedback sessions on teachers' accurate implementation of instructional units was evaluated using a multiple-baseline-across-teachers design. Data were collected on the percentage of correctly implemented instructional units during 10 min CM segments for all teachers across all conditions and phases. CM data were collected daily, while N/PS solving data were collected 3 days per week. N/PS data were collected to determine the extent to which correct instructional unit implementation during CM generalized; however, CM data were used to make decisions through baseline and the first two intervention phases.

The initial baseline included a minimum of five data points or until a stable data path or decrease in teachers' behaviors was established for all teachers. Once a stable data path across teachers was established, the group inservice occurred. Data collection continued for all teachers in the post-inservice phase until one teacher demonstrated the lowest and most stable trend (i.e., of at least five data points), at which point that teacher received the preconference, one individual coaching session, and one feedback meeting. While the first teacher received individual training, data collection continued for all remaining teachers in the post-inservice phase. When the first teacher's data indicated a clear change in level, trend, or variability of at least five data points, the second teacher having the next lowest and stable data received individual training. The same procedure was used to introduce the intervention to the third and fourth teachers.

Procedures

General study procedures. During all conditions teachers provided and audio-recorded daily math instruction. Teachers provided instruction using the state curriculum standards and selected portions of the EDM (Bell & Bell, 2004) program. The EDM curriculum includes instruction across the following math strands: (a) algebra and the use of variables; (b) data and chance; (c) geometry and spatial sense; (d) measures and measurement; (e) numeration and order; (f) patterns, functions, and sequences; (g) operations; and (h) reference frames. EDM lessons emphasize the application of mathematics to real world situations. Therefore, teachers presented numbers, skills, and mathematical concepts in the context of real-world situations and daily routines. Typical lessons included teacher demonstration, guided practice, independent practice using a student workbook (including review skills across several strands), and games. EDM lessons did not include prompts for teachers to use any enhancements teachers were trained to use in this study. EDM has demonstrated success with low achieving elementary students in a few experimental studies (ARC Center, 2003; Baxter, Woodward, & Olson, 2001) but does not currently meet the IES What Works Clearinghouse standards for evidence-based practices (IES, 2007).

During all conditions and phases, data were collected daily during two recurring lessons (i.e., CM and numeracy/problem solving). The researcher listened to 10 min of audio-recorded CM segments daily and listened to 10 min of audio-recorded numeracy/problem solving segments every third day to code the percentage of instructional units implemented correctly.

Baseline . During baseline, no training was provided to teachers. Teachers were asked to audio-record entire daily lessons in CM and numeracy/problem solving. In addition, the researcher informed teachers that she will be listening to student responses during math instruction. Teachers audio-recorded their lessons without any suggestions or prompts to make changes from their typical lesson delivery.

Post inservice. All four teachers attended an inservice (approximately 3 hours) conducted by the researcher. The inservice provided an overview of four instructional enhancements including: (a) appropriate use of model-lead-test (i.e., teacher models the correct response, teacher and students say correct response together, students say correct response independently); (b) systematic error correction (i.e., teacher provides model-lead-test, model-test, or lead-test immediately after students make an error); (c) unison responding (i.e., choral responding and write-on response cards); and (d) use of mastery to move forward in the curriculum (i.e., using student responses during instruction to determine when to introduce new skills).

The researcher provided teachers with the following instruction during the inservice: (a) a rationale for increasing active student responding; (b) an explanation of the critical features of each enhancement; (c) live and video demonstrations of the teaching procedures for each enhancement, including specific math examples using skills taught across the state curriculum standards; (d) opportunities for teachers to identify places in mock lessons for where and which enhancements could be used; (e) rules for reviewing and introducing new skills based on student mastery during lessons; and (f) opportunities for teachers to practice using enhancements in pairs with feedback from the primary investigator.

The researcher also provided teachers with two visual prompts to help teachers self-manage daily enhancement use. First, each teacher received a binder of blank daily enhancement checklists. The checklist was a short form to be completed by the teacher before each CM lesson, and included a series of closed-ended questions to help teachers make decisions about which enhancements to use during a given lesson (e.g., Is the skill new or review? Will you use model-lead-test, model-test or test? Which materials will you need?). See Appendix D for the daily enhancement checklist.

Second, each teacher received a laminated checklist (i.e., Visible Curriculum, Appendix G) depicting a list of the skills covered on each of the four quarterly math assessments. The checklist also included blank spaces next to each skill. The researcher prompted teachers to publicly post the checklist in the classroom and check off each skill when their students mastered it. During the inservice, teachers were prompted to mark skills as mastered when students in Tier 2 were independently firm for a few days of instruction on a particular skill (i.e., assessed during lessons using choral responding, response cards, or individual turns). In addition, teachers were given guidelines for reviewing mastered skills (e.g., review skills that are prerequisites for new skills, distribute review of skills over time).

At the conclusion of the inservice the researcher provided the teachers with all materials needed to implement the enhancements then prompted them to begin applying enhancements to all daily math lessons. Teachers continued to audio-record daily math calendar and numeracy/problem solving lessons. Data continued to be collected on the percentage of correctly implemented instructional units.

Coaching. Intervention consisted of three components (a) one preconference and planning meeting, (b) one coaching session, and (c) one feedback meeting. First, the researcher conducted a preconference with the teacher selected to enter intervention. The preconference lasted approximately 15-20 min. During the preconference, the researcher provided the teacher with specific feedback regarding strengths and opportunities for improvement using enhancements derived from post-inservice audio-recorded data. In addition, the researcher and the teacher planned the coaching lesson using the daily enhancement checklist. The researcher modeled selecting enhancements for targeted skills, then supported the teacher in selecting enhancements by asking guiding questions, providing feedback and error correction. Only skills within math calendar time were discussed, demonstrated, or coached so that untrained skills within the other math lessons (i.e., numeracy/problem solving) could be evaluated for generalization. Teachers were not be prompted to use the daily enhancement checklist during the math lessons used for generalization.

Second, the researcher attended the next CM lesson to demonstrate and coach the skills targeted in the preconference. The researcher used a side-by-side coaching model during this session. Specifically, the researcher: (a) modeled at least one instructional unit within each skill area taught during the CM lesson (e.g., one unit for time, one unit for money, one unit for counting by 2s); (b) modeled each target enhancement identified in the preconference and when possible, across skills (e.g., model error correction across money and rote counting); (c) immediately after modeling each skill, prompted the teacher to try at least two instructional units within the same skill; (d) gave specific praise to the teacher at least once per skill; (e) provide corrective feedback; and (f) provided another opportunity for the teacher to implement the instructional unit independently after error correction. All feedback was oral and provided to the teacher in a non-evaluative manner. The coaching sessions lasted approximately 30 to 45 min.

After the coaching session, the researcher instructed the teacher to implement the strategies during all math lessons. The teacher continued to audio-record daily lessons. The researcher continued to collect data on the percentage of correctly implemented instructional units in CM daily and the untrained numeracy/problem solving lessons 3 days per week. Then, the researcher conducted a feedback meeting with the teacher at least five sessions after the initial coaching session to follow up on the skills coached in the first session. At the feedback meeting, the researcher provided verbal feedback from audio-recordings, answered questions, and provided corrections. If at the time of the feedback meeting the teacher had not generalized the accurate use of enhancements to the numeracy/problem solving lessons, the researcher prompted the teacher to begin using the daily enhancement checklist during the second daily math period.

Materials. The researcher developed a PowerPoint® presentation to present an overview of the enhancements to teachers during the inservice. In addition, the researcher used video clips (Bursuck & Damer, 2007; Heward & Wood, 2006) to demonstrate the procedures for using each enhancement. The researcher also provided each teacher with a folder at the beginning of the group inservice containing a copy of the PowerPoint slides and sample teaching formats for each enhancement (i.e., scripted examples of model-lead-test, systematic error correction, and unison responding) to use during paired practice. In addition, each teacher received a binder containing blank daily enhancement checklists (i.e., one for each day for the duration of the study). See Appendix D for an example teaching format. Teachers also used write-on response cards. The response cards were be 10 inch by 10 inch square white, erasable boards on which students use markers to write their answers during instruction. The school already provided each teacher with a sufficient number of white boards and markers for the class. Finally, teachers audio-recorded daily lessons using Olympus® battery-operated digital recorders (Model # WS-110) provided by the researcher. The recorders were 3.8 x 2 x 9.8 inches, weighed less than 1 pound, and recorded up to 69 hours with 256 MB of internal memory. The recorders also contained a USB which allowed for transfer of the digital recording to a computer file.

Procedural fidelity. A second observer attended the group inservice and used a checklist to mark the occurrence of each of the prescribed steps included by the researcher during the inservice. In addition, one of the second observers marked the occurrence of each of the prescribed steps the researcher includes during preconference, coaching and feedback sessions. These sessions were audio-recorded so teachers can receive feedback privately with the researcher. The second observer used the audio-recording of the sessions to score procedural fidelity. Procedural fidelity was calculated by dividing the number of steps the researcher performs by the total number of steps to obtain two separate scores (i.e., inservice fidelity and coaching fidelity). In addition, another observer scored fidelity on the inservice and one of the preconference, coaching, and feedback sessions to establish interobserver reliability. See Appendix E and F for procedural fidelity checklist.

Data Analysis

To empirically evaluate the impact of enhancement training the researcher graphed the percent of correctly implemented instructional units for every session across all four teachers using Microsoft Excel®. Visual analysis of the graphs was used to determine changes in level, trend, or variability across all conditions and phases of the study for both the primary dependent variable and generalization. Using the multiple-baseline design, experimental control was demonstrated if improvements in teachers' level, trend, or variability of accurate instructional unit implementation were replicated across tiers as the coaching intervention was individually applied.

CHAPTER 4: RESULTS

Interobserver Reliability

Second observers scored 25% of the Calendar Math (CM) and Numeracy/Problem Solving (N/PS) sessions across teachers. Overall, interobserver reliability ranged from 60% to 100% with a mean of 91%. Overall CM reliability ranged from 60% to 100%, with a mean of 90.4%. Overall N/PS reliability ranged from 79% to 100%, with a mean of 93%.

CM. Across teachers, second observers evaluated 21% of the baseline recordings, 39.2% of the post-inservice recordings, and 28.2% of the post-coaching recordings. During baseline for CM, mean interobserver reliability was 89.3%, with a range of 71% to 100% across sessions. The mean post-inservice reliability for calendar was 90.1%, with a range of 60% to 100%, and a mean of 93% for post-coaching with a range of 86% to 100%.

N/PS. For N/PS, second observers evaluated 20% of the baseline recordings, 14.2% of the post-inservice recordings, and 40.9% of the post-coaching recordings across teachers. During baseline for N/PS, mean interobserver reliability was 91.1% with a range of 82% to 100%. The mean post-inservice reliability for N/PS was 91.1% with a range of 79% to 100%. The mean post-coaching reliability for N/PS was 100%.

The majority of disagreements between the primary data collector and second data collectors were related to unison errors and teacher “leads”. This is important to note because it may have been a function of using recorded teaching sessions rather than live observations.

Procedural Fidelity

A second observer listened to recordings of the group inservice, and all coaching activities (i.e., preconferences, coaching sessions, feedback sessions) to mark the occurrence of each of the prescribed steps included by the researcher. Overall procedural fidelity was rated 96.2%, inservice fidelity was rated 100%, mean coaching fidelity was 94%. Coaching fidelity was 92% for Kristy, 95% for Megan, and 95% for Beth. Time was extremely limited during the preconferences. Therefore, the researcher did not use this time to explain and set up a date for the feedback sessions, but subsequently arranged these meetings via email. During the coaching session, the researcher also did not tell Kristy what to focus on for the feedback session. This information was not given because the researcher did not have an opportunity to speak privately with Kristy immediately after the coaching session (i.e., students were present). The researcher reminded Kristy what to focus on for the feedback session via email the same day.

Another observer scored fidelity on the inservice and one of the preconference, coaching, and feedback sessions to establish interobserver reliability. Interobserver reliability for procedural fidelity across all of these training activities was 100%.

Effects of Inservice and Coaching on Teachers' Instructional Unit Accuracy

What is the effect of enhancement training in the form of inservice and coaching on teachers' accurate delivery of instructional units in CM?

Figure 7 presents the percentage of correct instructional units in CM across all conditions and phases for three of the teachers. All four teachers received the 3 hr group inservice; however, only three of the four teachers received coaching. One teacher, Jade, did not receive coaching because she achieved and maintained a high level of instructional unit accuracy after the group inservice (see Figure 7). Data collection ended for Jade when her student teacher assumed responsibilities for teaching math. All four teachers increased the number of correct instructional units following inservice, however a causal relationship between inservice and improvements in instructional unit accuracy cannot be claimed for this change, given that all teachers received the inservice together. A second change in level and a substantial decrease in variability occurred following coaching for all three teachers who received individual training. Further, visual analysis of the graph indicated replication of the positive effects of coaching across three teachers since each teacher's instructional accuracy improved immediately following coaching, while the preceding teacher's data remained stable in the post-inservice phase. Therefore, a functional relationship between coaching and increased instructional unit accuracy in CM was demonstrated.

Kristy. During baseline, Kristy's instructional unit accuracy scores were low and stable, with a range of 0% to 21%, and a mean of 5.9%. In the post-inservice phase, Kristy's scores improved, but remained relatively low (i.e., majority of sessions were below 50% correct). Post-inservice scores ranged from 36.3% to 68.1% with a mean of

46%. After coaching, Kristy's scores improved again, and remained relatively high. Post-coaching scores ranged from 76.6% to 96.7% with a mean of 90.9%.

Megan. During baseline, Megan's instructional unit accuracy scores were low and variable, ranging from 0% to 40.6% with a mean of 21.1%. In the post-inservice phase, Megan's scores improved somewhat, but continued to have high variability, with a range from 29.4% to 73.6%, and a mean of 50.7%. After coaching, Megan's scores ranged from 77.4% to 100% with a mean of 92.7%, and variability decreased substantially.

Beth. During baseline, Beth's instructional unit accuracy scores were low and variable, with a range from 0% to 27%, and a mean of 11.2%. After inservice, Beth's scores improved but still demonstrated high variability, ranging from 0% to 85.7% with a mean of 64.3%. After coaching, her scores ranged from 80% to 100% with a mean of 92.3%, indicating a change in level and a substantial decrease in variability as compared to post-inservice and baseline.

Jade. During baseline, Jade's instructional unit accuracy scores were low and relatively stable, with a range of 1% to 27% and a mean of 13%. In the post-inservice phase, Jade's scores increased substantially and remained stable over time, with a range of 66.6% to 94%, and a mean of 84.3.

To what extent does enhancement training, focused on teaching skills addressed during CM, generalize to instructional delivery during a second untrained daily math session focused on numeracy and problem solving?

Figure 6 presents the percentage of correct instructional units in Numeracy/Problem Solving (i.e., N/PS, generalization) across all conditions and phases. N/PS data showed a pattern similar to the primary dependent variable, but with substantially more variability. Despite the variability, mean instructional unit accuracy for all teachers improved from baseline to post-inservice, then again during post-coaching for the three teachers who received individual training.

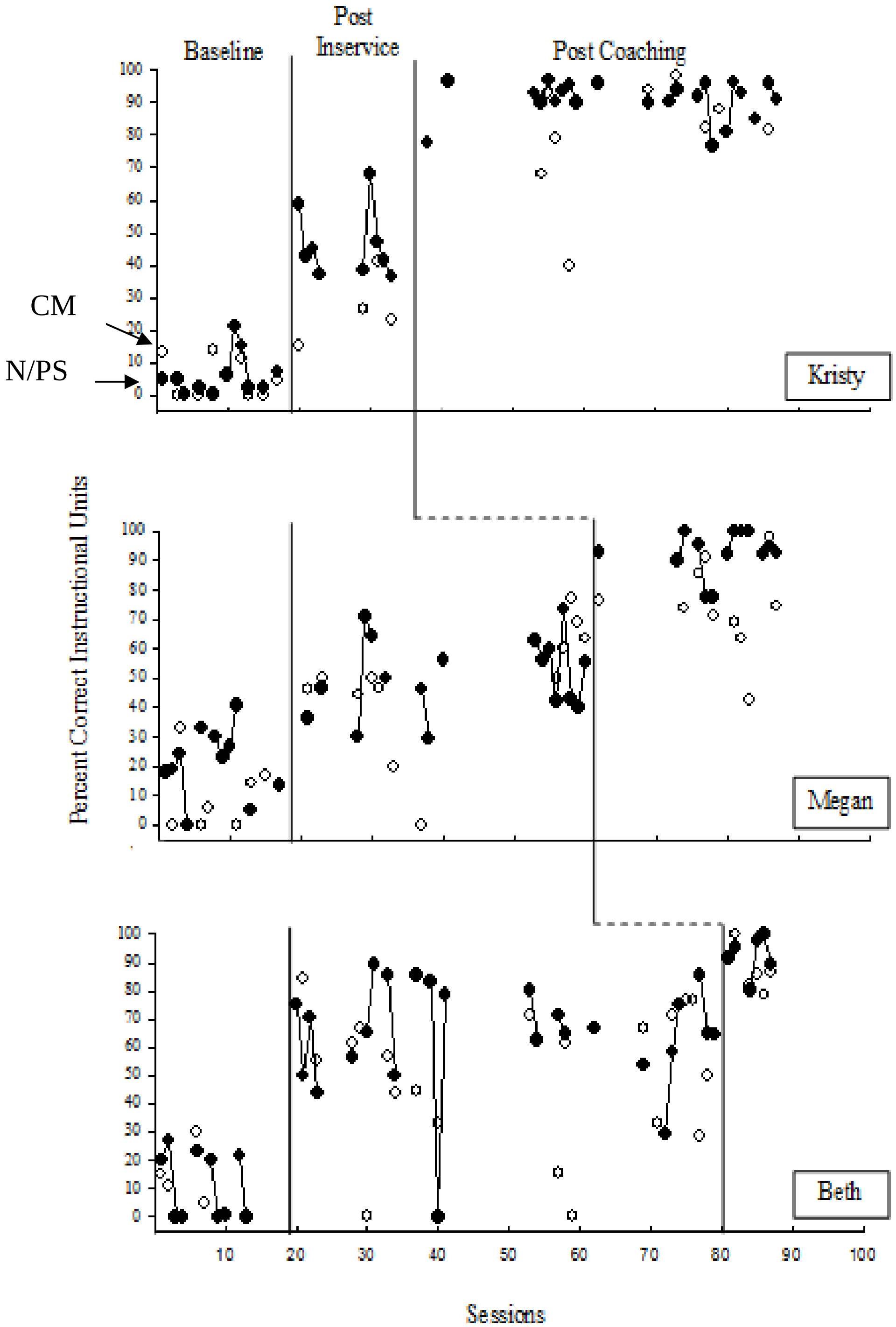
Kristy. During baseline, Kristy's instructional unit accuracy scores for N/PS were low and stable, with a range of 0% to 14%, and a mean of 5.3%. In the post-inservice phase, Kristy's scores improved, but remained relatively low, with a range of 15.3% to 41.1% and a mean of 26.5%. After coaching, Kristy's scores improved again, and remained relatively high. Post-coaching scores ranged from 40% to 98.4% with a mean of 80.4%. The majority of sessions during post-coaching were above 80%.

Megan. During baseline, Megan's instructional unit accuracy scores for N/PS were low and variable, ranging from 0% to 33% with a mean of 9.9%. In the post-inservice phase, Megan's scores improved somewhat, but continued to have high variability, with a range from 0% to 76.9%, and a mean of 48%. After coaching, Megan's scores ranged from 42.8% to 90% with a mean of 74.6%.

Beth. During baseline, Beth's instructional unit accuracy scores for N/PS were low and variable, with a range from 0% to 27%, and a mean of 11.2%. After inservice, Beth's scores improved with a mean of 49.8% but still demonstrated high variability, ranging from 0% to 85.7%. After coaching, her scores ranged from 78.5% to 100% with a mean of 86.5%, indicating a change in level and a substantial decrease in variability.

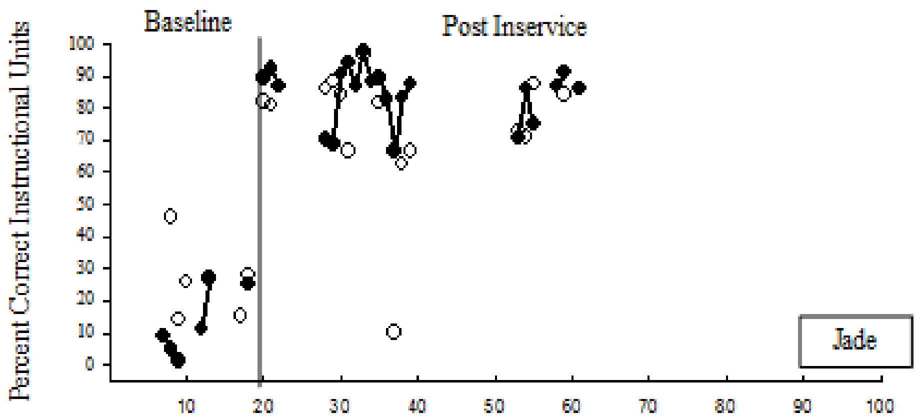
Jade. During baseline, Jade's instructional unit accuracy scores for N/PS were low and relatively stable, with a range of 14% to 28% and a mean of 25.7%. In the post-inservice phase, Jade's scores increased substantially and remained stable over time, despite one low score of 10%. Jade had a range of 10% to 96.8% and mean of 71.7%.

Figure 6. Percent correct instructional units in CM and N/PS for Kristy, Megan, and Beth.



Note: Closed data points represent CM, open data points represent N/PS.

Figure 7. Percent correct instructional units for Jade.



In addition to collecting data on instructional unit accuracy in CM and N/PS, the number of group responses were calculated for each 10 min session and averaged by phase for each teacher (i.e., responses using choral responding and/or response cards). Table 1 and Table 2 depict the mean and range for each teacher across the study. These data indicate that three teachers used group responding more during CM after the inservice. The largest difference in use of group responding for CM occurred from baseline to post-inservice. Kristy demonstrated an additional gain in use of group responding after coaching. Megan's use of group responding remained relatively consistent throughout all conditions and phases of the study.

A slightly different pattern of group responses emerged for N/PS. All four teachers increased group responding after the inservice, and the three teachers who received coaching increased group responding again after coaching. All teachers used group responding more during CM than N/PS during the post-inservice and post-coaching phases.

Table 1. Mean number of group responses per 10 min session in CM.

	Baseline	Post-Inservice	Post-Coaching
Kristy	x=11.8, r=0-28	x=23.4, r=11-29	x=34.6, r=22-49
Megan	x=22, r=15-30	x=18.6, r=7-28	x=19.9, r=8-29
Beth	x=9.1, r=5-14	x=17.3, r=7-29	x=22, r=12-41
Jade	x=11.8, r=7-17	x=29.3, r=11-51	

Table 2. Mean number of group responses per 10 min session in numeracy/problem solving.

	Baseline	Post-Inservice	Post-Coaching
Kristy	x=11.3, r=5-15	x=13.3, r=10-27	x=26.3, r=5-45
Megan	x=11.6, r=10-13	x=13.1, r=3-26	x=19.1, 12-33
Beth	x=7.3, r=1-18	x=12.5, r=0-26	x=23, r=17-33
Jade	x=12.6, r=6-17	x=21.2, r=4-36	

How many 1st grade students who receive enhancements during Tier 1 insruction change risk level (i.e., from "some risk" to "low risk" or "low risk" to "some risk")?

How many 1st grade students who do not receive enhancements during Tier 1 instruction change risk level?

The researcher intended to use student data from the AIMSWeb benchmark assessments as a measure of social validity. However, the school's assessment team administered the second benchmark assessment in early January, when only one teacher had received coaching. The timing of the assessment did not allow for a meaningful interpretation of the importance of the intervention relative to any changes in student performance. Therefore, these data were not examined.

Social Validity Results

What additive value do teachers place on coaching in relation to enhancement training through group inservice alone?

Teachers responded to two questions related to the differential impact of inservice and coaching on their instruction in CM and N/PS. Of the three teachers who received both inservice and coaching, two responded that the workshop was “somewhat helpful."All three teachers rated the coaching session "very helpful." Teachers agreed that the inservice was helpful in covering the initial information, and all three cited the video clips and practice as the most helpful aspect of inservice. Teachers reported that the coaching session helped build their confidence, allowed them to ask questions specific to their individual students, and receive feedback on their instruction. In addition, all three teachers reported that coaching in calendar carried over to N/PS. However, one teacher reported that it took her longer to get used to using the strategies in N/PS.

To what degree do teachers attribute enhancements in lessons to improvements in student achievement and progress through the curriculum?

All teachers noted increased academic engagement as an improvement directly related to using enhancements. Teachers also noted that students were able to apply more of learning from group instruction to individual work. All teachers reported that the enhancements motivated students because they enjoyed the interactive nature of the lessons. One teacher also suggested that students enjoyed seeing objectives checked off on the Visible Curriculum.

During coaching sessions, all teachers stated that students had mastered much of the CM content, and asked questions related to how to begin introducing new concepts during this instructional time. These comments suggest that teachers' attributed mastery of curriculum in part to using the enhancements. However, none of the teachers made remarks related to progress in the curriculum on the social validity questionnaire. In addition, only Jade used the Visible Curriculum to determine what topics to teach during the study. When listening to the recordings, the researcher observed teachers introducing new and/or more difficult content during CM after the introduction of enhancements post-inservice. However, the same pattern was not observed for N/PS. Teachers adhered to teaching one math topic in N/PS per week, as previously decided in grade-level planning at the beginning of the school year.

CHAPTER 5: DISCUSSION

The purpose of this study was to evaluate the effects of inservice and coaching on teachers' instructional unit accuracy in Calendar Math (CM) and to determine the extent to which changes generalized to an untrained math session (Numeracy/Problem Solving; N/PS). A multiple-baseline across teachers design evaluated changes in CM and N/PS across three teachers, with an additional teacher receiving inservice only. Results indicated a change in level for instructional unit accuracy after inservice, and a second change in level and decrease in variability after coaching, indicating a functional relationship between coaching and an increase in instructional unit accuracy. Results were replicated for N/PS, although with more variability in the post-coaching phase. In addition, teachers reported the training was valuable and feasible. Discussion points related to these results will be presented in this chapter, along with limitations of the study, and recommendations for research and practice.

Effects of the Intervention on Dependent Variables

What is the effect of enhancement training in the form of inservice and coaching on teachers' accurate delivery of instructional units in CM and N/PS?

Teachers in this study were trained to use techniques consistent with effective Tier 1 instruction found in current literature. Previous studies suggest that explicit instruction (Cirino et al., 2007), high levels of student engagement (Foorman et al., 1998), monitoring (Foorman et al., 2006), and planning

(McIntosh, Graves, & Gersten, 2007) are critical features of Tier 1 instruction. First, teachers were trained to use MLT, an instructional technique in which teachers introduce new material and correct errors by providing explicit modeling, supported practice, and testing using independent responses. Second, teachers were trained to use MLT in tandem with unison responding, which increased student engagement, as shown by changes in group responses (see Tables 1 and 2) and instructional unit accuracy (see Figure 6). Third, MLT and unison responding provided teachers frequent opportunities to monitor student responses and provide feedback. Finally, at the inservice, teachers were trained to plan instructional delivery using the Daily Enhancement Checklist, a document which was intended to prompt consistent, accurate use of the combined enhancements. Although this planning method was included in the training, none of the teachers used this tool, or found it particularly helpful. Potential reasons for this and implications will be discussed later in this chapter, as they relate to social validity and recommendations for research and practice.

While the intervention consisted of effective Tier 1 intervention components, it also combined strategies that have demonstrated strong effects on student achievement in previous investigations. First, MLT, choral responding, response cards, and systematic error correction have demonstrated positive effects on student achievement in a substantial number of previous studies (Barbetta et al., 1993; Barbetta et al., 1994; Gettinger, 1993; Hollingsworth & Woodward, 1993; Idol-Maestas, 1995; Nelson et al., 2004; Parks, Weber, & McGlaughlin, 2007; Van Houten, 1993). Second, all of the strategies used by teachers in this study are found in DI program design (Watkins & Slocum, 2004), which is supported by substantial research as well (Adams & Engelmann, 1996; Przychodzin-Havis et al., 2005; Przychodzin-Havis, 2004). Third, the combined use of MLT, unison responding, and systematic error correction as a Tier 1 intervention (as implemented in this study) has shown promise (Bursuck et al., 2004).

Not only were teachers trained to use empirically supported instructional techniques, they were also trained using an empirically supported professional development model, consisting of high quality inservice and coaching. The two-level training model used in this study meets three of the four criteria provided by Yoon et al. (2007) in the IES review of professional development studies. First, teachers received follow-up support after initial training (i.e., inservice), in one preconference, one coaching session, and one feedback session.

Second, training was provided directly to teachers by an individual with expert level knowledge of the methods. The researcher had substantial experience instructing young students using the enhancements, as well as experience training adult learners to use the enhancements using side-by-side coaching. Third, the effects of the professional development were measured behaviorally (i.e., instructional unit accuracy) in a way that allowed for causal inference about the effectiveness of training. Frequent observation and data collection across all teachers in the multiple-baseline design allowed the researcher to suggest causality between training activities and improvements in instruction.

In contrast, the professional development provided in this study did not meet the fourth criteria provided by Yoon et al., which was to provide training for a minimum of 14 contact hours. Given the high levels of instructional unit accuracy reached by teachers in this study after only 4 hours, it may be possible that this component is not as salient to producing teacher change as others found in the review of literature.

In addition to adhering closely to the IES guidelines for professional development, this study also closely followed previous research on the importance of general education "buy-in" for designing and implementing professional development. From results of numerous professional development studies, Maheady (2008) recommended that strategies presented to teachers should be clear, easy to implement, inexpensive, take little preparation time, and benefit all students in the class. On the social validity questionnaire, teachers rated the ease of implementation for all enhancements between "very easy" and "medium", which suggests that the strategies were clear and fairly easy to use during lessons. In addition, none of the strategies required extensive preparation by teachers. For example, students kept white boards and erasers at their individual work stations, that could be used during any lesson for a flexible response.

Also, the training and materials were inexpensive. The preparation of the training materials cost approximately 100 dollars (i.e., printing and lamination of the Visible Curriculum and Daily Enhancement Checklist). In addition, the only cost to the school was the price of four one-half day substitutes for the inservice, at 50 dollars per teacher. Teachers used already-existing materials during daily lessons (i.e., no additional cost), including the white boards and markers. If training were provided by qualified school staff, it is possible that this training could be provided to teachers for no cost in future staff development.

Finally, as recommended by Maheady (2008) the professional development was tied to socially important organizational outcomes. Specifically, this study's outcomes were tied to the school's goal of improving achievement for students at risk in an RTI model, using DI techniques. DI is a preferred method of instruction in the school district.

The results of this study were also different from existing professional development research because a change in teaching practice occurred immediately after the inservice. Garet et al. (2001) reported that one-day inservices typically produce little or no change in teaching practice. In this study teachers demonstrated an increase in the level of instructional unit accuracy, but still had substantial variability. The quality and structure of the inservice may have been a factor in producing some teaching changes. The inservice was provided to a very small, homogeneous group of teachers, which could be different from “typical” inservice delivery to larger groups of teachers (i.e., whole school/district) from various grade levels. In addition, the researcher provided multiple video clips of the instructional strategies, coupled with opportunities for teachers to practice teaching in pairs with feedback from the researcher. In a larger group, this level of individualization may not be possible, or would require additional trainers. Also, teachers knew the researcher would be listening to their post-inservice instruction, which could have influenced them to focus more intently on implementing the strategies correctly after the initial training. Daily observation would likely not follow a typical inservice.

Training provided in this study aligned well with standards for professional development (Yoon et al., 2007) and critical features of sustainability, in particular, teacher buy-in (Maheady, 2008). The primary finding of the study was that inservice and coaching effectively promoted sustained improvements in instruction, which further aligns the results to professional development literature. Specifically, the study combined the two models of coaching found effective in the literature, supervisory coaching (Fuchs et al., 1992; Joyce & Showers, 1995; Kohler et al., 1999) and side-by-side coaching (Gleason & Hall, 1991). Using supervisory coaching, the researcher observed (i.e., listened to recordings) lessons in which teachers used methods they had recently learned to implement, and recorded the presence or absence of particular instructional characteristics. Then, consistent with previous studies the researcher conducted a preconference, during which she gave nonevaluative feedback (i.e., termed, “suggestions to improve student responses”) to individual teachers.

Next, using side-by-side coaching, the researcher directly intervened using in-vivo demonstration during CM lessons. Direct intervention allowed teachers to see a model, then try the same format again with additional feedback from the researcher. Similar to other coaching studies, supervisory and side-by-side coaching improved the acquisition of new teaching behaviors (Coulter & Grossen, 1997; Kohler et al., 2001; Kretlow et al., submitted; O’Reilly et al., 1992).

In addition, the results of this study were consistent with results of previous coaching studies in that (a) coaching following inservice improved academic instruction and error correction (Kohler et al., 1999), (b) coaching following inservice improved student engagement (i.e., group responses as shown in Tables 1 and 2) and teacher behaviors (i.e., instructional unit accuracy, as seen in Figure 6, Kohler et al., 2001), and (c) coaching improved specific DI teaching procedures (Morgan et al., 1994). Tables 3 and 4 demonstrate specific changes in four types of errors that occurred after inservice and coaching. These data numerically depict how individualized coaching changed instruction for each teacher. For example, Kristy demonstrated a substantial decrease in the number of unison errors that were not corrected from inservice to coaching. In contrast, Beth's decrease in the number of scaffolding errors was more substantial. These descriptive data support that teachers benefited from the individualized feedback and support coaching provided.

Table 3. Mean number of errors per 10 min session for Calendar Math

Teacher/Phase	Individual turns (handraising)	No error correction for unison error	No error correction for task error	Scaffolding error (modeling or leading review skills, not modeling or leading new skills)
Kristy				
Baseline	5.75	8.25	1.25	2.87
Post Inservice	.83	9.83	2.33	.5
Post Coaching	0	3.5	.44	.33
Megan				
Baseline	7	7.55	2.55	5.22
Post Inservice	1.35	3	2.14	2.10
Post coaching	.77	.66	.11	.33
Beth				
Baseline	9.22	2.44	1.11	4.77
Post Inservice	.65	1.4	.9	2.45
Post coaching	0	0	0	1.2
Jade				
Baseline	11.6	1.75	1.6	7
Post Inservice	.13	.73	.96	1.26
Post coaching				

Table 4. Mean number of errors per 10 min session for Numeracy/Problem Solving.

Teacher/Phase	Individual turns (handraising)	No error correction-Unison	No error correction-Task	Scaffolding error (modeling or leading review skills, not modeling/leading new skills)
Kristy				
Baseline	5.25	5.75	3.50	2.75
Post Inservice	1.5	6.5	1.75	.75
Post Coaching	0	3	1	.71
Megan				
Baseline	8.75	1.75	6.25	.75
Post Inservice	4	2.33	2	1.1
Post coaching	3.8	1.12	1.3	.37
Beth				
Baseline	8.25	4	1.25	2
Post Inservice	2	4	1.87	1.75
Post coaching	.6	1.2	.80	.20
Jade				
Baseline	9.5	1	8.5	.5
Post Inservice	2.56	1.75	2.18	1.12
Post coaching				

Despite many similar findings with previous coaching studies, this study was also inconsistent with previous coaching literature in several ways. First, Kohler et al., (1999), Kohler et al., (2001), and Morgan et al., (1994) used extensive coaching, specifically a minimum of four coaching sessions to a maximum of two 25 min coaching sessions per week. This study found substantial changes in instruction after only four short sessions, one group inservice, one preconference, one coaching session, and one feedback session. The total time for all training activities was approximately 4 hrs. In relation to total training time, this study is more consistent with the results found by Maheady et al., (2004) and Kretlow et al. (submitted), which found high levels of accuracy in teaching procedures (i.e., 85% or higher) with approximately 4 hrs of inservice and coaching. In this study, teachers may have needed little coaching given their previous experience teaching DI programs. Also, recording lessons daily may have provided additional motivation for teachers to use enhancements correctly after coaching sessions.

The second dissonant finding of this study was that teachers demonstrated some generalization to an uncoached area. Kohler et al., (2001) found little or no changes in instruction for areas that were not routinely discussed with a coach. In this study, the researcher only directly coached skills taught during CM; however, to some extent in N/PS teachers' performance showed changes in level and decreased variability.

Finally, this study found similar results to Kretlow et al. (submitted) in regard to teachers' use of mastery in instructional decision making. Kretlow et al. found even when provided with consistent student feedback on performance via choral responding and response cards, kindergarten teachers taught the same content (i.e., routine format) each day during CM, regardless of student mastery. For example, teachers taught counting to 100 every day for the four month study duration, despite student mastery of the skill within the first few weeks of the study.

The researcher sought to prompt teachers to use mastery more efficiently in this study by adding "teaching to mastery" as an enhancement. The Visible Curriculum was intended to serve as a visual prompt for teaching to mastery, by making curriculum objectives public, so that teachers and students could see when skills were mastered, and what skills should be taught next. The researcher explained how to use the Visible Curriculum at the inservice.

Jade was the only teacher to use the Visible Curriculum to some extent. She used it only in CM. Jade posted the visible curriculum on her bulletin board and consistently checked off mastered skills. She routinely discussed mastery with students using the Visible Curriculum as a guide during recorded lessons. Jade expressed to the researcher that she consistently used the Visible Curriculum during CM because she thought it motivated students to see what they had accomplished. Interestingly, Jade had the largest and most stable change in level after inservice, although this change cannot be experimentally linked to her use of the Visible Curriculum.

When it was clear that the other teachers were not using the Visible Curriculum and were not using high levels of student accuracy on skills as an indicator to move forward to more difficult skills, the researcher addressed this in the individual preconferences focused on CM. Two of the three teachers (i.e., Kristy and Beth) introduced new and/or more difficult concepts in CM after prompting by the researcher during preconference, but still did not use the Visible Curriculum as an instructional guide even after coaching. For N/PS, all teachers taught one math topic per week (e.g., shapes for one week, addition the next week, graphing the next week), regardless of student performance during lessons.

There are several reasons teachers may have attended to factors other than student performance. First, all 1st grade teachers (i.e., four who participated in the study and three additional teachers) agreed to teach one math topic per week at the beginning of the year grade level planning meeting. One topic per week was decided in order to make certain all objectives on the North Carolina Standard Course of Study were introduced, and given equal time. Because this was a previously agreed upon schedule, teachers in the study may have been reluctant to deviate from it. Second, the school district provides teachers with pacing guides for math content standards, which outlines the preferred order in which math topics should be introduced and the amount of time teachers should spend on each topic (i.e., days and weeks). Again, teachers may have been reluctant to use a method that was inconsistent with school level and district level guidelines. In summary, teachers in this study progressed through the curriculum, but did not use mastery to do so.

In addition to extending the coaching literature, this study contributed to the literature related to measuring change in teacher behavior in several ways. First, this study added to a very limited set of studies that have used an objective measure to examine changes in teaching that occurred after professional development. In their review of professional development literature, Yoon et al. (2007) eliminated over 900 studies because the measure used to examine change was subjective (e.g., self-report, qualitative measures). Second, the majority of measures developed for observing change in previous studies were designed for use with highly structured or scripted programs (e.g., DI, CWPT). The instructional unit measure used in this study was not designed for use with a specific program, which may make it more transferrable across content areas, grade levels, and instructional settings.

Finally, in previous studies the instructional trial (Heward, 1994) and the learn unit (Greer and colleagues, 1985; 1991; 1992; 1996), have both been used to measure instructional behaviors. The instructional unit used in this study is similar to both measures in that it allows for a quality rating for each step of a three-term contingency. However, the instructional unit extends both the instructional trial and the learn unit because it focused on measuring instructional accuracy during group instruction, rather than with an individual student, and added quality ratings of unison responding and scaffolding, both of which are critical aspects of increasing student engagement and improving achievement (Godfrey et al., 2003; Hollingsworth & Woodward, 1993; Idol-Maestas, 1995; Parks, Weber & McGlaughlin, 2007; Sindelar, 1986).

Discussion of Social Validity Data

Acceptability of professional development. Overwhelmingly, teachers placed an additive value on coaching in relation to group inservice alone. These results are consistent with research on teacher preferences regarding side-by-side coaching. Blakely (2001) found the majority of teachers rated side-by-side coaching most effective in helping them acquire new DI teaching procedures. Teachers rated side-by-side coaching higher than group training and verbal feedback. Teachers in this study reported that coaching allowed them to “see and talk about things to improve on”, and that coaching provided “an opportunity to ask questions and get feedback.” Teachers also reported that coaching helped them “feel more confident” and “willing to try the strategies outside the coached lessons.”

During this study, three teachers requested additional coaching sessions, which suggests teachers found coaching acceptable and helpful. Also, as seen in the instructional unit accuracy data, coaching was extremely efficient. Some immediate gains occurred after inservice, but all teachers reached higher, more consistent levels of accuracy after only one side-by-side coaching session. Coaching allows the teacher to focus on specific aspects of instruction unique to their classroom and individual difficulties. For example, several of Kristy’s students had substantial difficulty sustaining attention during lessons, which led to many unison errors during instruction. Kristy’s coaching focused mostly on managing student responses by refining her signal and pace.

Neither Megan nor Beth had many unison errors prior to coaching, but each had other specific difficulties such as continuing to model when no longer necessary, or using error correction procedures incorrectly.

However, the effectiveness of coaching does not supplant the need for high quality inservice. It would not be as efficient to use coaching time to introduce instructional strategies to teachers individually than to use group inservice. A group inservice may be a more efficient way to share initial information, including steps for specific teaching procedures. Teachers also rated the inservice "very helpful" on the social validity questionnaire, primarily citing explanations of the enhancements, video clips, and opportunities to practice teaching as critical factors.

Although improvements in instructional unit accuracy occurred immediately after inservice, teachers' data for both CM and N/PS were highly variable. Empirical evidence suggests that optimal gains in student achievement may be diminished when teaching procedures are inconsistently used. For example, Furtak et al. (2008) and Kovaleski, Gickling, Morrow, and Swank (1999) found that low or variable levels of fidelity with teaching procedures correlated with lower gains in student achievement. In summary, both inservice and follow up are recommended by Yoon et al. (2007) based on an extensive review of empirical literature. Based on the data in this study and teacher responses on social validity questions, it appears that each may serve a very distinct function in the development of new teaching skills. The function of the inservice is to introduce teachers to new teaching procedures, model examples, and give teachers opportunities to practice with feedback. Coaching provides teachers with individualized feedback specific to difficulties they encounter after trying the procedures learned at the inservice. The primary value of coaching in relation to inservice is in vivo opportunities to practice and receive feedback.

Feasibility of using enhancements. Overall, teachers indicated that all enhancements were relatively easy to implement. Teachers were asked to use a Likert scale to respond to questions related to the level of difficulty they had implementing each of the enhancements (i.e., very easy, somewhat easy, medium, difficult, very difficult). All of the teachers rated choral responding and response cards “very easy” or “somewhat easy.” Teachers cited getting used to consistent signaling as the only challenge related to using choral responding and response cards.

Teachers rated MLT “somewhat easy” or “medium”. Several teachers reported that they had difficulty designing the wording, but once they had used it a few times it became more “simple” and “stress free.” Two of the three teachers cited coaching as an important factor in helping them master MLT. The mean group responding data in Tables 1 and 2 also support this for Megan and Beth. According to these data, Megan and Beth showed little change in their use of group responding from post-inservice to post-coaching, but demonstrated substantial changes in instructional unit accuracy after coaching. This suggests that Megan and Beth did not struggle with implementing unison responding, but did have some difficulty with MLT post-inservice, which was resolved after coaching. All teachers reported that they planned to continue using enhancements, and some reported that they were already using them in other subjects (e.g., writing). Given these comments, teachers may perceive the enhancements as a transportable way of teaching, which could lead to generalization beyond math.

Despite high ratings of acceptability and feasibility, teachers reported some challenges related to enhancement implementation. First, all four teachers reported that coaching made it easier to implement enhancements in N/PS. For example, one teacher noted that watching a demonstration of MLT helped her understand the format better, and therefore helped her feel more confident to use it in N/PS. In contrast, all four teachers also reported that it was easier to use enhancements in CM than in N/PS, because the content during N/PS was “constantly changing”, whereas CM was more routine. Three teachers also listed difficulty getting a few students to respond consistently. The issue of helping reluctant students to respond was discussed in individual coaching sessions. Teachers were given two recommendations (a) instruct all students to whisper their answer to a partner, then give a group response, and/or (b) give individual turns to students who are reluctant to respond after the group response (i.e., set students up for successful responses). Following coaching sessions, teachers reported that use of these strategies improved responding for the target students.

Perceived impact of enhancements. When asked how enhancements impacted students, all four teachers responded positively. Teachers listed benefits to students such as (a) improved engagement, (b) constant assessment, (c) increased student accountability for learning, (d) improved on-task behavior, (e) increased lesson structure to sustain attention, and (f) better student performance on independent work. In addition, 3 teachers suggested that other teachers should receive this training, which suggests that teachers felt enhancements could help other students as well.

Limitations

Several limitations in this study are important to discuss. First, although changes in teacher performance were clearly demonstrated, no student data were collected. As previously discussed, MLT, unison responding, and systematic error correction have substantial empirical support for improving achievement. However, there are limited data demonstrating the effectiveness of combining enhancements as a Tier 1 intervention. Only one study, Bursuck et al. (2004) has examined the combined use of Tier 1 enhancements on changes in student performance, and the effects were measured for reading.

Second, all teachers received DI training prior to the study, and had been using MLT, choral responding, and systematic error correction every day during Reading Mastery lessons. Previous DI training may have made it much easier to embed DI components in other lessons. Therefore, results of this study may not generalize to teachers with no DI experience. Third, although results indicated generalization of instructional unit accuracy to an untrained instructional session, both sessions measured were math. Results may not translate to other content areas (e.g., reading, writing, science).

Fourth, no long term maintenance data were collected. Teachers sustained use of enhancements for the duration of the study, while they were recorded daily, and all reported intentions to continue using them. But, it is unknown whether teachers' instructional unit accuracy would maintain for longer, unrecorded periods of time. Finally, teachers did not use the Visible Curriculum or the Daily Enhancement Checklist.

These materials may not have been perceived as relevant or useful to them, nor did they solve issues related to using student mastery noted by Kretlow et al. (submitted).

Recommendations for Future Research

Results of this study lead to several recommendations for future research. First, due to the timing of the AIMSweb administration, changes in student performance were not assessed. Future investigations should examine the impact of Tier 1 enhancements in math on changes in risk level, as measured by a reliable and valid progress monitoring system. In particular, it seems likely that the enhancements used in this study could be helpful in preventing academic failure for students identified in need of Tier 2 support, although changes in risk level should be evaluated for students in all Tiers.

Second, future studies should examine what percentage of instructional unit accuracy leads to critical gains in student performance. If research can identify the level at which students benefit substantially, coaching can be more focused on increasing accuracy to this level (e.g., 85%, 90%). This study demonstrated that inservice and coaching improved the level and consistency of instructional unit accuracy, but no guidelines for mastery criteria were set for teachers. Correlation or regression analysis could be used to identify specific levels of accuracy that lead to critical student gains, and to determine if student performance continues to improve as instructional unit accuracy improves (i.e., determine if a ceiling in student effects exists). Similarly, future research should also investigate the effects of using teachers' instructional unit accuracy data during preconference and feedback sessions, to see if additional improvements in accuracy will follow.

Third, future studies should investigate the added effects of enhancements when used with an empirically supported program, implemented with fidelity. In the only existing study that found positive effects for Tier 1 enhancements on student achievement, teachers used a structured reading program and regular fidelity data were collected (Bursuck et al., 2004). The enhancements used in the present study related only to the delivery of instruction, not the design of content. Further, teachers in this study did not use any program with any consistency, and the school's adopted program currently lacks empirical evidence for effectiveness with students at risk (IES, 2007). A critical feature of RTI is an empirically supported program, used with high fidelity (Fuchs & Deshler, 2007; Kame'enui, 2007). Enhancements alone may not produce salient changes in student performance, rather enhancements combined with high quality content design may be more effective. Such factors should be investigated experimentally. For example, future research should investigate whether additional gains in student achievement exist when enhancements are added to a nonscripted, evidence-based math program, similar to the method used by Bursuck et al. (2004) in reading.

Fourth, in addition to investigations related to student achievement, variables related to coaching also warrant future investigation. For example, in this study the researcher had extensive knowledge of instructional design and coaching procedures, she was not a member of the school staff. IES recommends professional development provided by an "expert" (Yoon et al., 2007), but other research also emphasizes the effectiveness of "bottom up" teacher training, in which teachers or instructional leaders (e.g., literacy facilitator, lead teacher) who are members of the school staff provide inservice and follow-up support to teachers (Desimone et al., 2001). Therefore, future studies should investigate the impact of teacher coaches on their peers' instructional unit accuracy. In particular, investigations in which special education teachers serve as peer coaches to general education teachers may be especially beneficial, given the consultative role special educators are likely to assume in RTI models (Haager, Klingner, & Vaughn, 2007).

Fifth, given the difficulty Kretlow et al. (submitted) and the researcher in this study had effectively training teachers to use mastery to make instructional decisions, future studies are needed to determine what factors impact teachers' decisions about what sequence of content to teach and when to introduce new or more difficult skills. This information will likely be gathered using qualitative methodology, perhaps through individual interviews and/or focus groups with general educators, attending grade level planning meetings, and examining existing documentation (e.g., weekly/monthly group planning tools, pacing guides, content standards). Answering questions related to teachers' instructional decision making will be helpful in determining an intervention to target teaching to mastery when combined with the other enhancements used in this study.

Similarly, given that teachers in this study did not find the Visible Curriculum or the Daily Enhancement Checklist useful tools, future investigations should revise these materials and training related to them, in order to promote progress through the curriculum and effective planning to use enhancements. More specifically, an investigation of what level of support (i.e., materials) is necessary to lead to changes in instructional decision making (i.e., teaching to mastery) and lesson planning. For example, using a multiple-baseline-across-teachers design, low level, teacher-centered materials such as the ones used in this study could be compared to more explicit, prescribed materials, such as the teaching formats used by Bursuck et al. (2004). In any investigation of materials, teacher recommendations should be considered prior to design.

Finally, a number of single subject studies have demonstrated the effectiveness of coaching on improvements in instructional quality across settings and content areas. These studies should be further investigated using quality indicators for single subject research (Horner et al., 2005), in order to determine whether coaching is an evidence-based professional development practice. Similarly, a randomized group experimental study in which an inservice and coaching professional development model is compared to "typical" inservice would be helpful in potentially providing additional, more generalizable support for coaching that would meet IES indicators for evidence-based practice as well.

Implications for Practice

Professional development. First, results from this study combined with several others including Kretlow et al. (submitted), Kohler et al. (1999), Kohler et al. (2001), Maheady et al. (2004), and Morgan et al. (1994) demonstrated the positive effects of inservice and coaching on instructional quality. These results combined with the recommendations of the IES review of professional development (Yoon et al., 2007) and the NCLB (2001) requirements for high quality professional development, suggest that when designing professional development, a two-level training model that includes high quality inservice and coaching should be included. Specifically, the combination of supervisory and side-by-side coaching has shown promise. In addition, given the changes seen after inservice alone in this study, it may be beneficial to train small, homogenous groups of teachers, with demonstrations of teaching procedures and opportunities to practice instruction with feedback embedded throughout training activities.

Second, teachers with experience using DI programs may be primed to use the strategies in other instructional sessions. In this study, the two teachers with the most DI experience (i.e., Beth and Jade) showed the greatest improvement in instructional unit accuracy after inservice alone. Although these results are not causal in nature, they suggest that a small amount of prompting may be helpful in getting teachers to use salient components of DI (e.g., choral responding) in other lessons. Similarly, the results of this study show that all teachers may not need individual coaching. More specifically, teachers who either (a) already demonstrate mastery of DI components in other lessons or, (b) are able to incorporate enhancements accurately after inservice alone may not need coaching. Further, these teachers may be good candidates for coaching roles. In addition, since some improvement occurred after initial inservice, it may be worthwhile to devote time during initial DI training to demonstrating ways to incorporate DI features into non-scripted lessons.

Teachers in this study consistently suggested some changes to inservice and coaching that may be helpful in practice as well. First, all teachers suggested more video clips of teachers using individual enhancements across math topics. Second, all teachers also suggested more feedback immediately following inservice. Interestingly, several suggested this feedback be provided in the form of email communication. Providing more frequent verbal feedback to teachers, sharing data, and answering questions via email is a feasible addition to the inservice and coaching model. Finally, several teachers suggested providing a set of generic, preprinted response cards at the initial inservice. Preprinted response cards were not provided in this study because in the investigation by Kretlow et al. (submitted), the teachers did not use the sets of cards provided. The cards given to teachers by Kretlow et al. were more specific to CM, and as the teachers suggest, more generic pre-printed response cards (e.g., ABCD, Yes/No) may be more useful.

Enhancements implementation. The enhancements used in this study are highly transportable, because they are very low cost in regard to training and materials. For this reason, they can be used in any content area, with any grade level, across a vast array of student achievement levels. For example, it costs nothing to implement choral responding on a consistent basis. Teachers can also create a set of write-on response cards for approximately 10 dollars, and can create sets of preprinted response cards using materials already available in schools (e.g., computer, construction paper, laminator). If “bottom up” peer coaching is implemented, training to use enhancements could also be cost free. A low cost, low tech intervention such as this may lend itself more toward sustainability.

In addition to limited cost, enhancements can be easily incorporated into lessons because they do not require intensive teacher preparation. The most time-consuming activity for teachers is planning MLT for introducing new concepts. Creating preprinted response cards may initially require some time, but if teachers choose to create generic sets of pre-printed response cards as some in this study did, they can laminate them and reuse across subjects and classes.

REFERENCES

Adams, G. L., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR.* Seattle, WA: Educational Achievement Systems.

Arc Center (2003). *The ARC Center tri-state student achievement study: Executive summary.* Retrieved October 6, 2008 from http://www.comap.com/elementary/projects/art//stories/portageprint.htm

Baker, S., Gersten, R., Lee, D. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal, 203, 51*-73.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching*, *44*, 375-395.

Barbetta, P. M., Heward, W. L., Bradley, D. M., & Miller, A. D. (1994). Effects of immediate and delayed error correction on the acquisition and maintenance of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis, 27,* 177-178.

Barbetta, P. M., Heron, T. E., & Heron, W. L. (1993). Relative effects of whole-word and phonetic-prompt error correction on the acquisition and maintenance of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis, 26,* 99-110.

Baxter, J. Woodward, J. & Olson, D. (2001). Effects of reform-based mathematics instruction on low achievers in five third-grade classrooms. *The Elementary School Journal*, *101,* 529-547.

Bell, J., & Bell., M. (2004). *Everyday mathematics. Grade 1.* Chicago, IL: SRA/McGraw-Hill.

Berninger, V. W., Rutberg, J. E., & Abbott, R. D. (2006). Tier 1 and tier 2 early intervention for handwriting and composition. *Journal of School Psychology, 44,* 3-30.

Blakely, M. R. (2001). A survey of levels of supervisory support and maintenance of effects reported by educators involved in Direct Instruction implementation. *Journal of Direct Instruction, 1,* 73-83.

Boardman, A. G., Arguelles, M. E., & Vaughn, S. (2005). Special education teachers' views of research-based practices. *Journal of Special Education, 39,* 168-180.

Bradley, R., Danielson, L., & Doolittle, J. (2007). Professional development, capacity building, and research needs: Critical issues for response to intervention implementation. *School Psychology Review, 36*, 632-637.

Brownell, M. T., Ross, d. D., & Colon, E. P. (2005). Critical features of special education teacher preparation: A comparison with general teacher education. *Journal of Special Education, 38*, 242-252.

Bursuck, W. D., & Damer, M. (2007). *Reading instruction for students who are at risk or have disabilities.* Boston, MA: Pearson Education, Inc.

Bursuck, W. D., Munk, D. D., Nelson, C., & Curran, M. (2002). Research on the prevention of reading problems: Are kindergarten and 1st grade teachers listening? *Preventing School Failure, 47,* 4-9.

Bursuck, W. D. et al. (2004). Evaluating the impact of a prevention-based model of reading on children who are at risk. *Remedial and Special Education, 25*, 303-313.

Christle, C. A., & Schuster, J. W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class math instruction. *Journal of Behavioral Education, 12,* 147-165.

Cirino, P. T., Pollard-durodola, S. D., & Foorman, B. R. (2007). Teacher characteristics, classroom instruction, and student literacy and language outcomes in bilingual kindergarten. *Elementary School Journal, 107,* 341-364.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in 1st grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394-409.

Coyne, M. D., Kame'enui, E. J., & Carnine, D. W. (2007). *Effective teaching strategies that accommodate diverse learners.* Upper Saddle River, NJ: Pearson Education.

Davis, L. L. , & O'Neil, R. E. (2004). Use of response cards with a group of students with learning disabilities including those for whom English is a second language. *Journal of Applied Behavior Analysis, 37*, 219-222.

Dixon, R. C., Carnine, D. W., Lee, D., Wallin, J., & Chard, D. (1998). *Review of high quality experimental mathematics research: Executive summary.* Report to the California state board of education. Retrieved October, 6, 2008 from http://idea.uoregon.edu/ ~ncite.

Edformation (2008). *AIMSWeb Tests of Early Numeracy*. Eden Prairie, MN: Pearson Psych Corp.

Engelmann, S., & Bruner, E. C. (1995). *Reading mastery classic I*. Columbus, OH: SRA/McGraw-Hill.

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37-55.

Foorman, B. R., Schatscheider, C., Eakin, M. N. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology, 31*, 1-29.

Fuchs, D., Compton, D. L., Fuchs, L. S., Bryant, J., & Davis, N. G. (2008). Making “secondary intervention” work in a three-tier responsiveness-to-intervention model: Findings from the first-grade longitudinal reading study of the national research center on learning disabilities. *Reading and Writing: An Interdisciplinary Journal, 21*, 413-436.

Fuchs, D. & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn’t be afraid to ask). *Learning Disabilities Research & Practice*, *22*, 129-136.

Fuchs, D. & Fuchs, L. S. (2006). Introduction to Response to Intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93-99.

Fuchs, D., Fuchs, L. S., & Burish, P. (2000). Peer-Assisted Learning Strategies: An evidence-based practice to promote reading achievement. *Learning Disabilities Research and Practice, 15,* 85-91.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement, using a reading maze task. *Exceptional Children, 58,* 436-450.

Fuchs, L. S., Fuchs, D., & Hollenbeck. K. N. (2007). Extending responsiveness to intervention to mathematics at first and third grades. *Learning Disabilities Research & Practice, 22*, 13-24.

Fuchs, L. S., Fuchs, D., & Karns, K. (2001). Enhancing kindergartners’ mathematical development: Effects of peer-assisted learning strategies. *Elementary School Journal, 101,* 495-510.

Fuchs, D., Fuchs, L. S., & Otaiba. A. (2001). K-PALS: Helping kindergarteners with reading readiness: Teachers and researchers in partnerships. *TEACHING Exceptional Children, 33,* 76-80.

Fuchs, L. S. et al. (2006). The effects of computer-assisted instruction on number combination skill in at-risk 1st graders. *Journal of Learning Disabilities, 39,* 467-475.

Fuchs, L. S. et al. (2008). Effects of preventative tutoring on the mathematic problem solving of third-grade students with math and reading difficulties. *Exceptional Children, 74*, 155-173.

Furtak, E. M. et al. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education, 21,* 360-389.

Gagnon, J. C., & Maccini, P. (2007). Teacher-reported use of empirically validated and standards-based instructional approaches in secondary mathematics. *Remedial and Special Education, 28,* 43-56.

Garet, M. S., Porter, A. C., & Garet, L. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 28,* 915-945.

Gersten, R. et al. (2008) *Report of the task group on instructional practices*. National Mathematics Advisory Panel. Retrieved September 16, 2007 from http://www.ed.gov/about/bdscomm/list/mathpanel/reports.html

Gersten, R., Chard, D., & Baker, S. Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities, 33,* 445-457.

Gettinger, M. (1993). Effects of error correction on third graders' spelling. *Journal of Educational Research, 87, 39*-45.

Gleason, M. M. & Hall, T. E., (1991). Focusing on instructional design to implement a performance-based teacher training program. *Education and Treatment of Children, 14,* 316-332.

Godfrey, S. A., Grisham-Brown, J., Schuster, J. W., & Hemmeter, M. L. (2003). The effects of three techniques on student participation with preschool children with attending problems. *Education and Treatment of Children, 26,* 255-272.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Greenwood, C. R. & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education, 24,* 276-289.

Greenwood, C. R., Hart, B., Walker, D., & Risley, T. R. (1994). The opportunity to respond revisited: A behavioral theory of developmental retardation and its prevention. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshleman, & T. A. Grossi (Eds), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 213-223). Pacific Grove, CA: Brooks/Cole.

Greenwood, C. R. & Maheady, L. (2001). Are future teachers aware of the gap between research and practice and what should they know. *Teacher Education and Special Education, 24,* 333-347.

Haager, D., Gersten, R., Baker, S., & Graves, A. (2003). *The English language learner observation instrument for beginning readers.* In S. Vaughn & K. L. Briggs (Eds.). Reading in the classroom: Systems for the observation of teaching and learning (pp. 200-215). Baltimore: Paul H. Brookes.

Haager. D., Klingner, J., & Vaughn, S. (2007). *Evidence-based reading practices for response to intervention.* Baltimore, MD: Brookes Publishing Company.

Haager, D. & Mahdavi, J. (2007). Teacher roles in implementing intervention. In D. Haager, J. Klingner, & S. Vaughn (Eds.) *Evidence-based reading practices for response to intervention,* p. 245-265.

Harniss, M. K., Carnine, D. W., Silbert, J., & Dixon, R. C. (2007). *Effective strategies for teaching mathematics*. In M. D. Coyne, E. J. Kame'enui, & D. W. Carnine (Eds.) Effective Teaching Strategies that Accommodate Diverse Learners (pp. 139-168). Upper Saddle River, NJ: Pearson Education.

Hasbrouck, J. E. (1997). Mediated peer coaching for training preservice teachers. *Journal of Special Education Teachers, 31,* 251-271.

Heward, W. L. (1994). Three "low-tech" strategies for increasing the frequency of active student response during group instruction. In R. Gardner, D. M. Saintao, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 283-320) . Monterey, CA: Brooks/Cole.

Hollingsworth, M., & Woodward, J. (1993). Integrated learning: Explicit strategies and their role in problem-solving instruction for students with learning disabilities. *Exceptional Children 59,* 444-455.

Horner, R. H., Carr, C. G., Halle, J., Gail, M., Odom, S., et al. (2005). The use of single-subject research to identify evidence-based practices in special education. *Exceptional Children, 71,* 165-179.

Idol-Maestas, L. (1995). *Group story mapping: A comprehension strategies for both skilled and unskilled readers. Technical Report.* Washington, DC: Center for the study of Reading.

Ingersoll (2001). Teacher turnover and shortages: An organizational analysis. *American Educational Research Journal, 37,* 499-534.

Institute for Education Sciences (2007). *What Works Clearinghouse intervention report: Everyday Mathematics.* Retrieved October 6, 2008 from http://ies.ed.gov/ncee/wwc/pdf/WWC_Everyday_Math_043007.pdf

Iver, M. A., & Kemper, E. The impact of direct instruction of elementary students' reading achievement in an urban school district. *Journal of Education for Students Placed at Risk*, 7, 197-220.

Ingham, P. & Greer, D. R. (1992). Changes in student and teacher responses in observed and generalized settings as a function of supervisor observations. *Journal of Applied Behavior Analysis, 25*, 153-164.

Iver, M. & Kemper, E. (2002). The impact of Direct Instruction on elementary students' reading achievement in an urban school district. *Journal of Education for Students Placed at Risk, 7,* 197-220.

Jager, B., Reezigt, G. J., & Creemers, B. P. (2002). The effects of teacher training on new instructional behavior in reading comprehension. *Teaching and Teacher Education, 18,* 831-842.

Jackson, B., Larzelere, R., St. Clair, L., Corr, M., Fitcher, C., & Egertson (2006). The impact of HeadsUp! Reading on early childhood educators' literacy practices and preschool children's literacy skills. *Early Childhood Research Quarterly, 21*, 213-226.

Joyce, B. & Showers, B. (1995). *Student achievement through staff development: Fundamentals of school renewal.* New York: Longman.

Kame'enui, E. J. (2007). Responsiveness to intervention: A new paradigm. *TEACHING Exceptional Children, 39*, 6-7.

Kamps, D. M. (1994). Enhanced small group instruction using choral responding and student interaction for children with autism and developmental disabilities, *American Journal on Mental Retardation, 99,* 60-73.

Klingner, J. K., Vaughn, S., & Hughes, M. J. , & Arguelles, M. E. (1999). Sustaining research-based practices in reading: A 3-year follow up. *Remedial and Special Education, 20*, 263-274.

Knight, S. L., & Wiseman, D. L. (2005). Professional development for teachers of diverse students: A summary of the research. *Journal of Education for Students Placed at Risk, 1,* 387-405.

Kohler, D. W., Ezell, H. K., & Paluselli, M. (1999). Promoting changes in teachers' conduct of student pair activities: An examination of reciprocal peer coaching. *The Journal of Special Education, 33,* 154-165.

Kohler, D. W., Crilley, K., & Shearer, D. D. (2001). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90,* 240-250.

Kovaleski, J. F., Gickling, E. E., Morrow, H., & Swank, P. R. (1999). High versus low implementation of instructional support teams: A case for maintaining program fidelity. *Remedial and Special Education, 20,* 170-183.

Kretlow, A. G., Wood, C. L., & Cooke, N. L. (submitted). Using inservice and coaching to improve kindergarten teachers accurate delivery of instructional units. *Journal of Special Education.*

Lane, K. L. (2007). Identifying and supporting students at risk for emotional and behavioral disorders within multi-level models: Data driven approaches to conducting secondary interventions with an academic emphasis. *Education and Treatment of Children, 30,* 135-164.

Maccini, P., & Gagnon, J. C. (2006). Mathematics instructional practices and assessment accommodations by secondary special and general educators. *Exceptional Children, 72,* 217-225.

Maheady, L. (2008, May). *Evidence-Based Practice Reviews*. Presentation at the annual Convention of the Association for Behavior Analysis International, Chicago, IL.

Maheady, L. Harper, G. F., Mallette, B., & Karnes, M. (2004). Preparing preservice teachers to implement class wide peer tutoring. *Teacher Education and Special Education, 27,* 52-62.

Marchand-Martella, T. A., Slocum, & R. C. Martella (Eds.) (2004). *Introduction to Direct Instruction* (pp. 28-65). Boston, MA: Allyn & Bacon.

McIntosh, A. S., Graves, A., & Gersten, R. (2007). The effects of response to intervention on literacy development in multiple-language settings. *Learning Disability Quarterly, 30,* 197-211.

Morgan, R. L., Menlove, R., Salzberg, C. L., & Hudson, P. (1994). Effects of peer coaching on the acquisition of Direct Instruction skills by low-performing teachers. *The Journal of Special Education, 28,* 59-76.

Miller, S., & Hudson, P. J. (2007). Using evidence-based practices to build mathematics competence related to conceptual, procedural, and declarative knowledge. *Learning Disabilities Research & Practice*, *22*, 47-57.

Narayan, J. S., Heward, W. L., Gardner, R., Courson, F. H., & Omness, C. K. (1990). Using response cards to increase student participation in an elementary classroom. *Journal of Applied Behavior Analysis, 23,* 483-490.

National Council on Teacher Quality (2006). *What education schools aren't teaching about reading and what elementary teachers aren't learning*. Washington, DC.

National Joint Committee on Learning Disabilities (1997). *Assistance to states for the education of children with disabilities, preschool grants for children with disabilities, and early intervention program for infants and toddlers with disabilities; proposed rule" 34 CRF 300, 301, and 303, Federal Register 62: 204 (Oct. 22, 1997).* Retrieved October 6, 2008 from http://www.ed.gov/legislation/FedRegister/propule/1997-4/102297a.html.

Nelson, J. S., Alber, S. R., & Gordy, A. (2004). Effects of systematic error correction and repeated readings on the reading accuracy and proficiency of second graders with disabilities. *Education and Treatment of Children*, *27*, 186-198.

North Carolina Department of Public Instruction (2008). *North Carolina standard course of study and grade level competencies*. Retrieved October 6, 2008 from http://community.learnnc.org/dpi/math/Kto5.FINAL.13May08.pdf

Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. In R. Case & Y. Okamoto (Eds.), *The role of central conceptual structures in the development of children's thought: Monographs of the Society for Research in Child Development,* (Vol. 1-2, pp. 27-58). Malden, MA: Blackwell Publishers.

O'Reilly, M. G. et al. (1992). Teaching systematic instruction competencies to special education student teachers: An applied behavioral supervision model. *Journal of the Association for the Severely Handicapped*, *17*, 104-111.

Parks, C., Weber, K. P., & McLaughlin, T. F. (2007). The effects of fading, modeling, prompting, and Direct Instruction on letter legibility for two preschool students with physical and developmental delays. *Child and Family Behavior Therapy, 29,* 13-21.

Przychodzin-Havis, A. M., et al. (2005). An analysis of corrective reading research. *Journal of Direct Instruction*, *5*, 37-65.

Randolph, J. J. (2007). Meta-analysis of the research on response cards: Effects on test achievement, quiz achievement, participation, and off-task behavior. *Journal of Positive Behavior Interventions*, *9*, 113-128.

Raney, P. & Robbins, P. (1989). Professional growth and support through peer coaching. *Educational Leadership, 35,* 35-38.

Rosenshine, B. (1983). Teaching functions in instructional programs. *Elementary School Journal, 83,* 335-351.

Ross, D. E., Greer, R. D., & Singer-Dudek, J. (2005). The teacher performance rate and accuracy scale (TPRA): Training as evaluation. *Education and Training in Developmental Disabilities, 40,* 411-423.

Schumaker, J. B., et al. (2002). *The educational context and outcomes for students with disabilities: General education and the satisfaction of general education teachers.* Report to the Office of Special Education Programs, U.S. Department of Education. Institute for Academic Access, Research Report 5. Retrieved March 24, 2008 from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1a/73/07.pdf

Simmons, D. C., Coyne, M. D., & Kwok, O. (2008). Indexing response to intervention: A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities, 41,* 158-173.

Simonsen, F., & Gunter, L. (2001). Best practices in spelling instruction: A research summary. *Journal of Direct Instruction, 2,* 97-105.

Sindelar, P. T., Bursuck, W. D., & Halle, J. W. (1986). The effect of two variations of teacher questioning on student performance. *Education and Treatment of Children, 9,* 56-65.

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research, 68,* 277-321.

Sweeney, W. J., Ehrhardt, A. M., Gardner, R., Jones, L., Greenfield, R., & Fribley, S. (1999). Using guided notes with academically at-risk high school students during a remedial summer social studies class. *Psychology in the Schools, 36,* 305-318.

Texas Education Agency. (1984). *Texas Teacher Appraisal System.* Austin, TX: Author.

U. S. Census Bureau. (2007b). Selected economic characteristics: 2007l. Retrieved June 16, 2008 from http://facfinder.census.gov.

U.S. Department of Education (2003). Identifying and implementing educational practices supported by rigorous evidence. Institute of Education Sciences National Center for Education Evaluation and Regional Assistance. Retrieved September

15, 2008 from http://www.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

U. S. Department of Education (2001). The No Child Left Behind Act. Washington, DC: U. S. Department of Education.

U.S. Department of Education (2007). Individuals with disabilities education improvement act. Washington, DC: U. S. Department of Education.

Van Houton, R. (1993). Rote vs. rules: A comparison of two teaching and correction strategies for teaching basic subtraction facts. *Education and Treatment of Children, 16,* 147-160.

Vaughn, S. Wanzek, J., Woodruff, A. L., & Linan-Thompson, S. (2008). *Prevention and early identification of students with reading disabilities*. In D. Haager, J. Klingner, & S. Vaughn (Eds.) Evidence-based Reading Practices for Response to Intervention (pp.11-28). Baltimore, MD: Paul H. Brookes.

Vellutino, F. R., Scanlon, D. M., & Small, S. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities, 39,* 157-169.

Watkins, C. L. & Slocum, T. A. (2004). The components of direct instruction. In N. E. Marchand-Martella, T. A. Slocum, & R. C. Martella (Eds.) *Introduction to Direct Instruction* (pp. 28-65). Boston, MA: Pearson Education Inc.

Williams, D., & Coles, L. (2007). Teachers' approaches to finding and using research evidence: An information literacy perspective. *Educational Research, 49,* 185-206.

Wood, C. L., Mabry, L., Kretlow, A. G., Lo, Y., & Galloway, T. (in press). Effects of preprinted response cards on students' participation and off-task behavior in rural kindergarten classroom. *Rural Special Education Quarterly*.

Woodcock, R. M., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery-Revised.* Allen, TX: DLM Teaching Resources.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from http://ies.ed.gov/ncee/edlabs

APPENDIX A: DATA COLLECTION FORM

Date ___ Data Collector _________ Teacher _________ CM/NPS Duration____

Activity	Opportunity for Group Response	Model, Lead, Test	Errors: Task Error, Unison Error, No Error	Controlled Response	Notes
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	
	CR RC	M L T	T U No	Y N	

Number of correct opportunities ___/___Number of total opportunities ___x 100 =___

APPENDIX B: SOCIAL VALIDITY QUESTIONNAIRE

1. How many years have you taught? How many years teaching 1st grade?

2. What teaching licenses do you currently hold? Circle/list all that apply.

   Elementary (K-5/K-6)

   Special Education: List category ________________________

   Other _______________

3. How many years have you taught a Direct Instruction program (e.g., Reading Mastery)?

4. Did you think the enhancements (i.e., choral responding, response cards, model/lead/test) improved your teaching during calendar? Please explain.

5. Do you think the enhancements improved your teaching during math? Please explain.

6. Please rank (by circling) the following aspects of the enhancements by the level of difficulty you had implementing them.

Model/Lead/Test Very Easy Somewhat Easy Medium Difficult Very Difficult
Please Explain
____
____

Choral Responding Very Easy Somewhat Easy Medium Difficult Very Difficult
Please Explain
____
____

Response Cards Very Easy Somewhat Easy Medium Difficult Very Difficult

Please Explain

____
____

7. How helpful were the following training activities you received?

Workshop Not helpful Somewhat helpful Very helpful

Please Explain

____
____

Demonstration/Coaching Session Not helpful Somewhat helpful Very helpful

Please Explain

____
____

8. In what ways did the enhancements impact your students? Please explain.

9. Please describe any added effects, if any, the individual coaching had on your implementation of the enhancements.

10. Did the demonstration and coaching session you received in Calendar help you implement the enhancement in Math? Please explain.

11. Do you plan to continue using any of the enhancements in the future? Why or why not?

12. What challenges did you face in implementing the enhancements?

13. What changes would you make to the training and follow-up support (i.e., workshop, coaching)?

14. What materials/feedback/support would have made it easier to implement enhancements in your classroom?

APPENDIX C: SAMPLE TEACHING FORMAT

Adapted from Stein, Kinder, Silbert & Carnine (2006)

Introducing new numbers

Teacher: You are going to count and end up with 10. Get ready (Signal)

Students: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

Teacher: I'm going to count and end up with 13. Listen: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13. Listen to the new part. 11, 12, 13. When I drop my hand, start with 10 and say the new part with me.

Teacher & Students: 11, 12, 13

Teacher: Say the new part all by yourselves starting with 10.

Students: 10, 11, 12, 13

Teacher: Now you're going to count and end up with 13. Start with 1. Get ready (signal).

Students: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13.

APPENDIX D: DAILY ENHANCEMENT CHECKLIST

Date_______

Skill 1 (e.g., counting by 2s)____________________

New skill? ___ **Review skill**? ___

If new, MLT___ or MT___ If review, begin with T (independent responses)

Unison responding: Response Cards___ Choral Responding ___ Other___

Date_______

Skill 2 (e.g., counting by 2s)____________________

New skill? ___ **Review skill**? ___

If new, MLT___ or MT___ If review, begin with T (independent responses)

Unison responding: Response Cards___ Choral Responding ___ Other___

Date_______

Skill 3 (e.g., counting by 2s)____________________

New skill? ___ **Review skill**? ___

If new, MLT___ or MT___ If review, begin with T (independent responses)

Unison responding: Response Cards___ Choral Responding ___ Other___

Date_______

Skill 4 (e.g., counting by 2s)____________________

New skill? ___ **Review skill**? ___

If new, MLT___ or MT___ If review, begin with T (independent responses)

Unison responding: Response Cards___ Choral Responding ___ Other___

APPENDIX E: INSERVICE FIDELITY CHECKLIST

	Yes	No	N/A
Researcher explains the rationale for increasing active student responding	___	___	
Researcher explains the critical features of choral responding	___	___	
Researcher demonstrates the choral responding procedures across content domains (live demo and video)	___	___	
Researcher provides opportunity for teachers to practice choral responding in pairs	___	___	
Researcher explains the critical features of write on response cards	___	___	
Researcher demonstrates the response card procedures across content domains (live demo and video)	___	___	
Researcher provides opportunity for teachers to practice using response cards in pairs	___	___	
Researcher provides specific praise to teachers during practice	___	___	
Researcher provides error correction to teachers during practice	___	___	___
Researcher explains the critical features of model-lead-test	___	___	
Researcher explains the rules for using MLT/MT/T	___	___	
Researcher demonstrates the MLT/MT/T procedures across content domains (live demo and video)	___	___	
Researcher provides opportunity for teachers to practice MLT/MT/T in pairs	___	___	
Researcher provides specific praise to teachers during practice	___	___	
Researcher provides error correction to teachers during practice	___	___	___
Researcher explains the critical features of error correction	___	___	

Researcher explains the rules for using MT, MLT, LT error corrections	___	___
Researcher explains the rules for correcting unison errors	___	___
Researcher demonstrates the error correction procedures across content domains (live demo and video)	___	___
Researcher provides opportunity for teachers to practice error correction in pairs	___	___
Researcher provides specific praise to teachers during practice	___	___
Researcher provides error correction to teachers during practice	___	___
Researcher leads teachers in identifying places to add all enhancements in teacher-provided math lessons	___	___
Researcher explains and demonstrates how to use the Daily Enhancement Checklist	___	___
Researcher explains the rationale and procedure for using the Visible Curriculum	___	___
Researcher provides decision making scenarios using the Visible Curriculum and asks teachers related questions	___	___
Researcher instructs teachers to begin using enhancements in both daily math sessions	___	___

APPENDIX F: COACHING FIDELITY CHECKLIST

Preconference	**Y**	**N**
Researcher states agenda of the meeting	__	__
Researcher provides specific praise for at least one enhancement the teacher is implementing	__	__
Researcher explains demonstration/coaching process	__	__
Researcher suggests 3 student objectives for calendar lesson (based on lesson recordings)	__	__
Researcher asks for teacher's feedback on suggested student objectives	__	__
Researcher explains calendar procedural facilitator, including a rationale for why particular enhancements are strategically placed for 2 of 4 activities for each skill	__	__
Researcher will ask guiding questions to lead the teacher to place enhancements for the remaining 1-2 activities for each skill listed on procedural facilitator	__	__
Researcher highlights/underlines up to 3 target behaviors for teacher to watch for during demonstration e.g., MLT, EC, signaling)	__	__
Demonstration/Coaching		
Researcher models each target behavior identified in preconference and when possible, across skills (e.g., EC for counting by 2s, EC for saying the wrong day, EC for saying the wrong color pattern)	__	__
Researcher models learning trials correctly (e.g., provides EC when necessary, does not lead on old skills)	__	__
Immediately after modeling each skill, researcher prompts teacher to try at least 2 learning trials within the same skill	__	__
Researcher gives specific praise to teacher at least once for each skill	__	__
Researcher provides corrective feedback if teacher does not implement instructional unit correctly	__	__
Researcher provides another opportunity for the teacher to implement an instructional unit after corrective feedback	__	__

APPENDIX G: VISIBLE CURRICULUM

Skill	Quarter Assessed	Check when Mastered	Plan for Periodic Review *Important for: Component skills, prerequisite skills, skills assessed across multiple quarters
BASIC NUMERACY			
Match a model to a numeral (e.g., match five circles to the numeral 5) up to 10	1		
Match a model to a numeral up to 20	2		
Match a model to a numeral up to 30 including tallies, pictures of base 10 blocks, number sentences and coins	3		
Match a model to a numeral up to 99 including tallies, pictures of base-10 blocks, number sentences, coins	4		
Write the number that tells how many objects	1		
Draw a number of items to match the numeral to 10	1		
Draw a number of items to match the numeral to 25	2		
Match a model to a number word to 10	1		
Match a model to a	2		

number word to 20			
Match a model to a number word up to 30	3		
Match a model to a number word up to 99	4		
Represent a numeral three different ways (up to 20)	4		
PROBLEM SOLVING			
Draw pictures to solve single digit addition problems	1		
Draw pictures to solve single digit subtraction problems	2		
Solve one and two digit addition story problems to 20	2		
Solve story problems using groupings of 2s, 5s, and 10s with models and pictures to count collections of objects	3, 4		
Solve addition story problems two different ways (e.g., numbers, pictures)			
Solve subtraction story problems two different ways (e.g., numbers, pictures)	3		
Using an analog clock, write the time to the hour	3		

TELLING TIME			
Using a digital clock, transfer the time to an analog clock by drawing the hands	1		
Using a analog clock, transfer the time to a digital clock	4		
Order 3 numbers from least to greatest (up to 20)	1		
Order three numbers from greatest to least (up to 20)	1		
Write the numerals 1 to 20 in order	1		
Identify the smallest number in a set (up to 50)	1		
Identify the largest set of tallies and cubes up to 60	2		
Order numbers from least to greatest (up to 50)	3		
ROTE COUNTING			
Rote count by 1s to 100	2		
Rote count by 10s to 100	1		
Rote count by 5s to 100	1		
Rote count by 2s to 40	1		
CALENDAR			

Answer questions using the calendar (e.g., If today's date is __, what day will Saturday be?	2, 4		
Write missing dates on a calendar	4		
Answer "day of the week" questions	4		
SHAPES			
Match a shapes to the written name (i.e., square, trapezoid, parallelogram, hexagon)	2		
Identify cylinders, rectangular prisms, and cones (e.g., everyday items)	2, 3, 4		
Identify parallelograms, trapezoids, and hexagons that are different sizes	4		
Identify the shape of everyday items	3		
Describe how 2 shapes are similar	2, 3		
Describe how 2 shapes are different	2, 3		
Draw shapes: parallelograms, squares, trapezoids, hexagons	2, 3		
Sort shapes by attribute	3		

GRAPHING/ORGANIZING DATA			
Identify “most” on a graph (e.g., tally, bar)	3		
Identify “least” on a graph (e.g., tally, bar)	2		
Create a line plot using raw data	2		
Chart raw data using tallies	3		
Use Venn diagrams to answer questions about similarities and differences	4		
PATTERNING			
Translate (change) patterns using letters, numbers, or words (up to ABC)	3		
MEASUREMENT			
Use non-standard units to determine the relative size of objects (e.g., which shape holds the most blocks, cubes, how many blocks are needed to cover a space)	2		
Use pattern blocks to fill a shape	3		
PROBABILITY			
Discriminate events that are more likely to happen (from a set of 2)	4		

Discriminate events that are less likely to happen (from a set of 2)	4		
Discriminate events that are certain to happen (from a set of 2)	**4**		
Discriminate events that are impossible to happen (from a set of 2)	**4**		