Utah State University
DigitalCommons@USU

All Graduate Plan B and other Reports | Graduate Studies

5-2012

A Synthesis of Math Interventions for Elementary Students

Stephanie Toone
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Part of the Psychology Commons

Recommended Citation

Toone, Stephanie, "A Synthesis of Math Interventions for Elementary Students" (2012). *All Graduate Plan B and other Reports*. 967.
https://digitalcommons.usu.edu/gradreports/967

This Report is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Plan B and other Reports by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.

A Synthesis of Math Interventions for Elementary Students

by

Stephanie Toone

A plan-B paper submitted in partial fulfillment

of the requirements for the degree

of

EDUCATIONAL SPECIALIST

in

School Psychology

Approved:

Donna Gilbertson, Ph.D. Committee Chair	Gretchen Gimpel Peacock, Ph.D. Committee Member

Renee Galliher, Ph.D.
Committee Member

UTAH STATE UNIVERSITY
Logan, UT
2012

Table of Contents

List of Tables

Abstract

One major concern when developing a Response to Intervention (RTI) multi-tier program is the procedural elements required to ensure effective instruction. Ten studies were reviewed to gain an understanding of the components that result in positive RTI outcomes in mathematics. Elements analyzed included RTI program procedures, tier administration and tier outcomes. Results suggest that several empirically-based math interventions are being used in RTI programs in the literature and the researchers are implementing options as part of the Classwide, Tier 2 and Tier 3 instructional programs. Intervention progress is monitored primarily using Curriculum Based Measurement procedures to evaluate performance levels and growth rates on computational math skills. Results from this review indicate, however, that among the students receiving intervention the percentage of students responding to the different interventions was not consistently 80% or more at each tier. Specific procedures including intervention training, frequency, session duration, group size and criterion to identify responders will be discussed as well as potential challenges in applied settings.

Introduction

Poor performance in mathematics is a problem in schools today. An estimated 70% of elementary school aged children in American schools experience difficulty successfully learning and applying mathematics principles (National Mathematics Advisory Panel, 2008). The 2003 National Assessment of Education Progress indicated that 31% of the nation's fourth-grade students scored at or above the proficiency standard in the area of mathematics (Manzo & Galley, 2003) and most students do not obtain mathematics proficiency standards by the end of their formal schooling (Perie, Grigg, &Dion, 2005; U.S. Department of Education, 2003). Learning math skills requires continually building on an understanding of procedures, which rely on each other for successful learning. Thus, when difficulties in math are not discovered early on, and interventions are not available to help them, children with academic problems continue to fall further behind each year in school (National Mathematics Advisory Panel, 2008; National Council of Teachers of Mathematics [NCTM], 2000). Often students' lack of basic skills continue to influence performance on more complex math applications in later grades and these students are often identified as a student with a learning disability due to severe difficulties in math performance (Silver, Pennet, Black, Fair, & Balise, 1999). Reported prevalence of mathematics learning disabilities (MLD) in the general population ranges from 3% to 8% (Gross-Tsur, Manor, & Shalev, 1996; Hanich, Jordan, Kaplan, & Dick, 2001; Grégoire, & Desoete, 2009; Mazzocco & Meyers, 2003). Students identified with specific learning disabilities perform lower and grow at a slower pace relative to their peers in learning mathematics.

Unfortunately difficulties in academics are not the only consequences of having a Learning Disorder. Often times low academic functioning is also associated with discouragement, low self-esteem, and social skills deficiencies. Individuals with Learning Disorders have been reported to drop out of school approximately 1.5 times more than those without learning disorders (American Psychiatric Association, 2000). Thus, educators need to provide effective academic intervention in the area of mathematics when difficulties first emerge to prevent the development of severe learning difficulties in mathematics.

To assist the number of students who are struggling in math, research clearly shows that early mathematics intervention can remediate skill gaps and prevent future deficits (Clements & Sarama, 2007; Fuchs, Fuchs, & Karns, 2001; Fuchs, Fuchs, Yazdian, & Powell, 2002; Griffin & Case, 1997; Sophian, 2004). Although early intervention often repairs academic problems for many children, often, a student's remediation needs are not met in the general education classroom when problems first emerge because special education services for severe deficits is seen as the only available option for intervention (Donovan & Cross, 2002). Traditionally, the accepted procedure for identifying and placing children with a learning disability in special education has been a refer-test-place procedure (VanDerHeyden, Witt, & Barnett, 2005). This procedure results in a lack of services until a student shows a severe deficit in math performance to be eligible for testing for special education services. The students who struggle but do not show a severe enough deficit to warrant eligibility as a student with a learning disability in the area of math, would continue to struggle with fewer avenues for supplemental support.

Multi-tiered Response to Intervention Approach

Given the problems with meeting all students' needs within a two tier model consisting of general or special education, IDEA 2004 allows schools to implement an alternative multitier approach for eligibility assessment for LD. This approach, Response to Intervention (RTI), reforms the two tier approach by expanding the current assessment practice to focus on identification of students for early intervention/prevention and allocate at least one additional tier of intervention support to identified at-risk students before evaluating for special education services (Burns & Gibbons, 2013; Burns, Riley-Tillman, & Vanderheyden, 2013; Riccomini & Witzel, 2010; Vanderheyden & Burns, 2010). To accomplish this goal, the level of performance and learning rate over time of every student within the entire school population is monitored throughout the school year. Students who are not adequately responding within the general education curriculum will be given more intensive instructional support. The purpose of the more intense interventions are to increase the identified low performing students' performance within the range of their peer's performance or a benchmark criterion indicating adequate low risk performance within a reasonable period of time. Students' response to these supplemental remedial interventions then becomes an assessment tool that can be used with additional data to determine whether or not the child needs special education services due to a learning disability. A poor response to a series of quality interventions provides data suggesting that the student has a chronic learning disability and his/her poor performance is not better explained by some other factor such as low motivation or prior poor instruction (Noell, Gilbertson, VanDerHeyden, & Witt, 2005; Mellard, Byrd, Johnson, Tollefson, & Boesche, 2004). Special education services are then considered if a student's performance level and learning rate are both significantly less than that of his or her peers even with additional supports.

The concept of responsiveness to intervention within a multitier approach appears to be a promising method to obtain data to help identify math disabilities while providing intervention to remediate academic deficits. The focus on frequent evaluation of all students' learning within the RTI approach provides a solid data based approach for making educational judgments about distribution of instructional resources, the effectiveness of resources allocation for promoting expected learning for the greatest number of students. Although results from a few studies indicate that RTI may successfully identify and support at-risk students, most studies have suggested that RTI effectively enhances outcomes in reading. Although many students experience difficulties in math, less is empirically known about the specific application of RTI as a potential alternative for preventing serious math delays. Empirical support for RTI multitier support on math performance is just emerging to guide the selection of screening measures, progress monitoring measures and intervention protocols that may be appropriate for each tier of instruction (Newman-Gonchar, Clarke, & Gersten, 2009). The following sections will provide brief background on this literature.

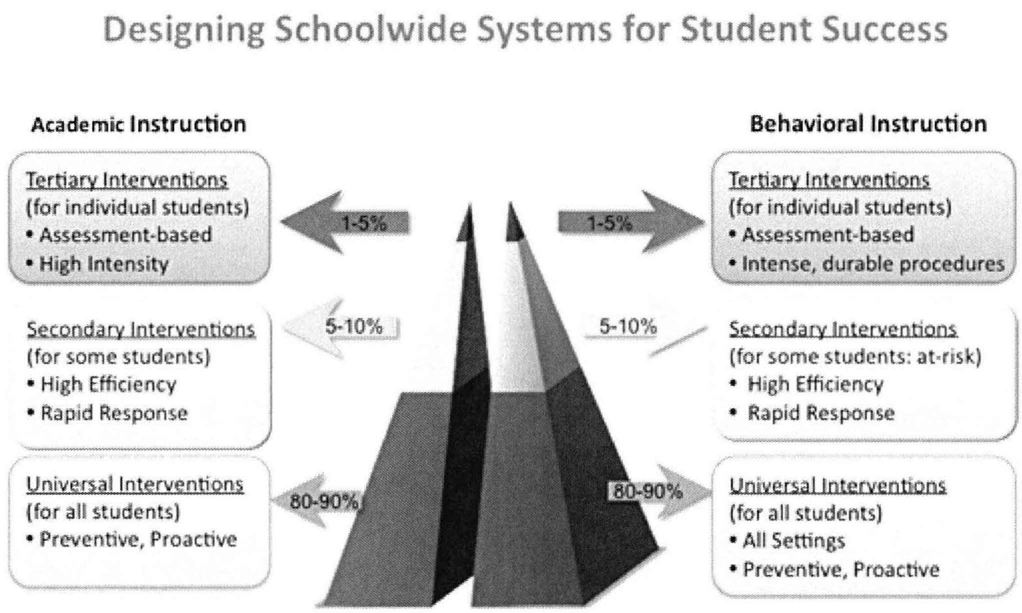
Response to Intervention Approaches

Clearly, adequate support for students struggling in math will only be accomplished when it is known which interventions and RTI strategies best optimize performance for at-risk students (Fuchs, Fuchs, & Hollenbeck, 2007; VanDerHeyden & Witt, 2005). In schools, one definition of "best" may be ease in training and implementation that increases the likelihood that the intervention will be used in a manner that effectively supports students. Two models of RTI have been proposed by several researchers that differentially balance program efficiency and effectiveness: a standardized protocol or a problem solving approach (Fuchs & Fuchs, 2006). To enhance feasibility and efficiency, the standard protocol approach to RTI involves selecting a few effective validated interventions for all identified at-risk children with comparable problems in a given area. This approach attempts to simplify the process by preparing intervention and progress monitoring materials, training those involved in implementation of the intervention, and setting up procedures to determine whether the intervention has been applied accurately in just a few interventions. By organizing a few standardized protocols to utilize in an RTI program, it is more likely that a well-organized intervention program with adequate support can be effectively implemented with a significant number of students. The goal of this approach is to effectively enhance the majority of the identified at risk student's academic progress by selecting the best intervention available for the most common math problems.

Alternatively, to enhance intervention effectiveness, the problem-solving model relies on the assumption that there are numerous individual reasons for academic problems. These individual differences make it difficult to determine a small set of interventions that will work for the majority of students based only on a set of defining characteristics. Thus, solutions to problems are developed by conducting a four-stage problem solving process including problem identification, problem analysis, plan implementation, and problem evaluation (Fuchs, Mock, Morgan, & Young, 2003). This approach takes more time than the standardized protocol and requires problem solving skills and extensive expertise in probe. However, the identification of individualized interventions may be adequately efficient and effective if used with those students who do not respond to the standardized protocol. The more labor and time intensive problem solving would then be conducted with a few identified at-risk children who did not adequately respond to the standard intervention protocol.

Given the advantages of both types of programs, schools may first implement and evaluate the more feasible standard protocol approach for students who are performing below their classmates and in the risk range. The more intensive problem solving approach may then be employed with a smaller subset of students who show more extensive skill deficits to identify specific individual causes of poor performance in mathematics. Incorporating both approaches into an RTI models allows educators to use a systematic yet efficient approach to effectively address students' problems.

In both RTI approaches, the teaching interventions of strengthening intensity are often referred to as tiers (see figure 1from http://www.pbis.org/school/rti.aspx). Often the tiers are presented as a triangle with three overlapping tiers, which together represent a continuum of interventions that increase in intensity based on the subsequent responsiveness of the learner (Sugai, 2007). The first tier involves whole-group instruction and general screening. This tier is used for core instructional interventions for problems in basic skill areas, and for interventions that general education teachers may take on in the general education class setting. This tier generally addresses the needs of approximately 80% of students. Tier 2 involves approximately 15% of the student population and uses targeted, small-group intervention. Here, students who are at risk are assisted with more intensive, research-based interventions with close progress monitoring in conjunction with the primary instruction received by all students. The third tier includes the most intensive intervention setting and usually provides for the needs of approximately 5% of the student population (Berkeley, Bender, Peaster, & Saunders, 2009).



Curriculum Based Measurements in Mathematics

To assess student response to intervention, curriculum-based measurement (CBM) is one evaluation tool that enables educators to frequently progress monitor student response to intervention several times a year, month, or week (Deno, 2003; Shinn, 1989). CBM of mathematics (M-CBM) involves the administration of brief one to eight minute timed tests consisting of grade level math problems to assess student level and trend of academic performance on skills taught within the math curriculum. Students' performance can be measured repeatedly over time to monitor trends whereas traditional standardized measures are typically less sensitive to change in performance and are administered once throughout the school year.

Christ, Scullin, Tolbize and Jiban (2008) reviewed the research and psychometric evidence for M-CBM on computation assessment. The reliability data of CBM for computational skills is acceptable, based on the findings reviewed in this study with Cronbach's alpha coefficients of M-CBM internal consistency and alternate form reliabilities reported at levels higher than .80 and coefficients of inter-rater reliability coefficients that ranged from .60 to 1.00. Criterion-related validity coefficients between M-CBM and standardized math tests vary vastly (between .38 to .83) with much higher validity estimates among computation assessments. Thus, Christ and colleagues concluded that the research literature supports the use of CBM to guide screening-type decisions about math computation skills. Stecker, Fuchs, and Fuchs (2005) reviewed the literature on the utility of M-CBM to improve student achievement. These authors identified five studies that investigated the relationship between implementing MCBM and subsequent improvement in math achievement relative to the math achievement of students in classes that did not use CBM. In all five studies, general education or special education teachers included in the CBM trained group were asked to administer CBM probes weekly and review the data with consultants in order to make instructional changes at both the class and individual levels. Results indicated that students within the experimental CBM groups outperformed students in the control groups (i.e., no CBM training) on CBM probes at the end of the study. Results from this series of studies indicate that CBM math assessments increased teacher's ability to frequently track student progress, give specific suggestions for planning instruction, match students up for tutoring, and give students immediate feedback on their development (Stecker, Fuchs, & Fuchs, 2005).

School-based Interventions in Mathematics

Several recent reviews of interventions in the areas of mathematics may also indicate a collection of interventions that may best optimize performance for at-risk students in Tier 2 and 3. Many different types of interventions have improved math performance for at-risk students including: cooperative learning, peer tutoring, reward/motivation strategies, goal-directed strategies, schema/cognitive based strategies, and drill & practice (Powell, Fuchs, Fuchs, Cirino, & Fletcher, 2009). Codding, Panahon, Panahon, and Benson (2009) reviewed 37 studies investigating intervention effects on math computation rates to identify interventions of simple and moderate intensity. Studies, published between 1980 and 2007, included kindergarten to twelfth grade students experiencing difficulties in math. Simple interventions were defined as interventions that improved the academic environment without changing the instructional process. These interventions included contingent reinforcement, performance feedback, goal setting, cue cards, altering instructions or timing, or changing the form of practice opportunities. Moderate interventions were defined as interventions that enhanced the existing classroom instruction such as direct instruction on the weakest skills, increased skill opportunities, or increased pace of instruction. Identified simple interventions that included earning free time, flash card practice, goal setting with contingent reinforcement, and count-bys resulted in large effects sizes for most participants (d = 0.33 to 4.74). Moderate interventions that included peer tutoring, Cover-Copy-Compare self-instruction, review of taped problems, and incremental rehearsal, resulted in mostly large effects sizes for most participants (d = 0.17 to 8.59). All intervention methods were effective for improving at least one academic variable (i.e., accuracy, rate, score).

In a later review of 17 single case studies, Codding, Burns, and Lukito (2011) found that drill and practice with modeling generated the largest effects (phi coefficient of .71) and self-management showing moderate treatment effects (phi coefficient of .55) that targeted fluency gains on computational math skills as compared practice without modeling. Adding modeling to standard drill and practice procedures resulted in better outcomes for students with and without learning disabilities. Baker, Gersten, and Lee (2002) observed that giving teachers and students specific information on how each student is achieving appeared to improve mathematics achievement consistently. Also using peers as tutors improved success. Providing clear, specific feedback to parents of low achievers on their children's successes in mathematics also proved to have the potential to enhance achievement modestly. Finally, Kroesbergen and Van Luit (2003) found that students can learn effectively through both self-instruction and direct instruction methods. For the learning of basic math facts, direct instruction seems to be the most effective. For the learning of problem-solving skills, self-instruction methods can be very effective.

Summary

Clearly there is a need to intervene on math difficulties that are experienced by many students but there are many unknown answers to questions about how to intervene. Educators' awareness of and readiness to implement interventions that address early math difficulties may increase the likelihood that math deficit problems are remediated for most students. Adopting Response to Intervention (RTI) as an assessment tool is one alternative approach that has been recently recommended for providing multiple tiers of services to students that include services for at-risk students as well as students with learning disabilities (Fuchs, Fuchs, & Speece, 2002). One important goal of this program is to allocate school based intervention to minimize the number of students experiencing math difficulties by providing a series of more intensive interventions when math problems are first emerging. Given that this type of program is being done in complex school environments, a current challenge is to select effective practices that will be successfully implemented and sustained with adequate organizational guidance and support. An assumption of an RTI assessment is that a student's poor responses to an effective general education curriculum and to several well-implemented intensive interventions may be an indication of a learning disability (Fuchs, 2003). A small but growing amount of literature on RTI programs in the area of math is one source that may provide information on RTI procedures and intervention options to resolve math deficits. Thus, the purpose of this paper will be to examine this existing literature to summarize empirically supported intervention strategies that may be used in RTI. Given that early intervention is needed in elementary school to prevent math difficulties in later grades, intervention studies conducted with elementary children will be reviewed. One purpose of this synthesis of the literature on RTI programs targeting math performance is to gain an understanding of the RTI components that have potential positive effects on math outcomes. Specific elements of the RTI model that included at least one tier in addition to Tier I will be reviewed and coded for universal Tier I, Tier II and Tier III.

Method

Study Selection

Computer searches of articles published between 2000 and 2009 were conducted. PsycInfo and Psychology and the Behavioral Sciences Collection databases were used as the primary source for locating studies on the effects of an RTI program on student math performance. Studies that were included in this literature review met the following inclusion criteria: 1) implemented any type of RTI program with at least one Tier to provide effective math instruction in elementary, middle and/or high schools and 2) evaluated math numeracy, computation, and problem solving performance. The following descriptors were used in the database search: response to intervention, responsiveness to intervention, arithmetic, procedures, math, integrity and fidelity. The references of all selected studies were reviewed in an effort to find other potential studies that met the inclusion criteria.

Areas of Evaluation and Coding Procedures

Each identified study was coded within four sections: (1) Description of Studies, (2) RTI Program Procedures, (3) Tier Administration and (4) Tier Outcome. The author reviewed and coded dimensions within each of these three sections for each of the identified studies as described in the following sections. A description of specific coding procedures used within each section follows.

Description of Studies

The first broad area of this study pertained to the demographics of each school setting included in each study. First, the included grade levels were categorized as K to 3rd grade, 4th to 5th grade, 6th to 8th grade, or 9th to 12th grade. Second, the inclusion of a Title one school was also coded for each study. Third, the demographic information of study participants was coded by evaluating whether or not the percentages of the following demographics of participants were reported in the methods section: race, socio-economic status as reported by reduced lunch program, and English language learners.

Two study methods were also coded. First, the study design was coded as (a) Experimental, (b) Quasi-experimental, (c) Single subject design, or (d) a validity study of RTI. Second, as a high level of treatment integrity is required to draw accurate inferences about the relationship between an intervention and behavior change, all studies were coded "yes" if treatment integrity was measured and "no" if treatment integrity was not measured or there was no mention. The procedure used to assess the fidelity of RTI implementation was also coded as (a) direct observation, (b) permanent products, (c) daily logs, (d) audio tapes or (e) not measured.

RTI Program Procedures

The general procedure to evaluate the RTI program was the second broad area examined in this study. First, the Tier levels, which were the independent variables that were implemented and evaluated, were coded for each study. Tier 1 level was coded as a level when screening data was reported or if a classwide intervention was investigated. Tier 2 was coded when an evaluated intervention was reported by authors as a Tier 2 level and was reported as a supplemental intervention to a general education program. Tier 3 was coded when an evaluated intervention was reported by authors as a Tier 3 level and was reported as an additional intervention to the supplemental intervention and a general education program.

Second, the math outcome dependent variable was categorized as either a computational math skill measure or as an applied or word problem measure. The specific assessment measure used to evaluate progress on the math skill outcome was recorded and tallied.

Third, the reported training activities described prior to or as part of the implementation process of the program were coded. Training activities were coded as verbal and written instructions, coaching, in class coaching, role play, follow up problem solving, follow up performance feedback, or no mention of training methods.

Tier Administration

For the third broad area of this study, dimensions were coded and evaluated for each of the Tiers implemented within a study to gain specific knowledge on procedural aspects of each Tier. First, the specific math program used in each Tier was recorded and tallied. Second, the teaching components of the intervention (e.g., modeling, guided practice, independent practice, feedback, error correction, goal setting, Reinforcement, Graphing) were coded. Third, personnel implementing the program was also coded as (a) general education teacher, (b) special education teacher, (c) peer tutors, (d) researchers or (e) not mentioned. Fourth, program administration procedures such as how often the program was implemented (i.e., 2 to 3 times a week, 4 to 5 times a week or not mentioned), the duration of each session (i.e., 15 minutes or less, 16 to 30 minutes, 31 minutes to 60 minutes, or not mentioned) and group size (classwide, 4 to 6 students, 2 to 4 students, individual, or not mentioned) were also coded. Finally, how often student response was evaluated per Tier was recorded as (a) daily, (b) weekly, (c) 2 to 3 times a week, (d) 4 to 5 times a week, (e) monthly, (f) 3 times a year, (g) one time for study, (h) pre-post, or (i) not mentioned.

Tier Outcome

The final broad category pertained to the outcome of each Tier. First, the criterion used to identify responders was coded including (a) slope, level, or dual discrepancy, (b) benchmark or standardized test percentile, and (c) mastery, instructional, or frustrational levels. Also the number of students participating in a Tier and number students reported as responding to the intervention were also evaluated to estimate effectiveness of each Tier. These numbers were used to calculate the percentage of students responding to the administered Tier intervention in the study.

Results

A total of 10 studies were identified as meeting the inclusion criteria for this review (see studies with * in Reference section). All of the studies were published between 2000 and 2009. Coded results of these studies will be presented for each of the four broad areas examined in this study: (1) Description of Studies, (2) RTI Program Procedures, (3) Tier Administration, and (4) Tier Outcome.

Description of Studies

Demographics. Table 1 presents the reported demographic data in each study. Of the ten studies, all of the studies (n = 10; 100%) were implemented in an elementary setting. More specifically, 70% were conducted in grades Kindergarten through third grade and 30% were conducted in grades four and five.

The racial and ethnic composition of the sample in each study was reported for 60% of the ten studies. Information regarding the percentage of English language learners and reduced or free lunch among participants in the sample was reported by 30% and 40% of the studies. Student population percentages for race, ELL and Reduced/free lunch are reported in Table 2.

Table 1

Demographics of Students in Studies.

Demographics		Frequency	Percentage
Required parent permission		3	30%
Grade level	K-3	7	70%
	4 to 5 grade	3	30%
	Middle	0	0%
	High School	0	0%
Race	Reported	6	60%
ELL	Reported	3	30%
Reduced/Free Lunch	Reported	4	40%
Title I school status	Yes	4	40%
	Unclear	6	60%
	no	0	0%

Table 2

Reported Percent Population of Sample

	White		Black		Latino(a)		Asian	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0% - 10%			1	16.7	1	16.7	3	50.0
11% - 20%			3	50.0	1	16.7		
21% - 30%					1	16.7		
31% - 40%	1	16.7	1	16.7	2	33.3		
41% - 50%	1	16.7			1	16.7		
51% - 60%	1	16.7						
61% - 70%	2	33.3						
71% - 80%	1	16.7						

	ELL		Reduced Lunch	
	Frequency	Percent	Frequency	Percent
0% - 10%	2	33.3		
11% - 20%	1	16.7	1	16.7
21% - 30%				
31% - 40%			1	16.7
41% - 50%			1	16.7
51% - 60%			1	16.7

Study design. Three types of study designs were used to examine the effects of the RTI program interventions on math skills. Specifically, studies used experimental (n = 5), single subject design (n = 4), and validity of RTI (n = 1).

Assessment of treatment integrity. The assessment of program fidelity was reported by all of the reviewed studies (N = 10). Two types of integrity measures were used: direct observation and audiotapes. Direct observation was used the most with 60% (n = 6) implementing this method. The use of audiotapes was implemented in 40% (n = 4) of the studies reviewed. All studies that measured treatment integrity reported high levels of integrity indicating that treatment steps were accurately implemented 80% to 100% of the time that integrity levels were observed.

RTI Program Procedures

Within the broad area of RTI program procedures, results of four dimensions will be presented. These include Tier levels implemented, the type of math dependent outcome variable, training procedures, and assessment of treatment integrity.

Tier levels implemented. When examining which Tier levels were evaluated in studies, the coded results showed that the majority (n = 9; 90%), of the included studies reported screening students per grade level in Tier 1 to identify students exhibiting mathematical difficulties (See Table 3). A smaller percentage, 40% (n = 4) reported Tier 1 students response to a classwide level intervention to identify students exhibiting mathematical difficulties. Tier 2 students were evaluated in 70% (n = 7) of the studies, and Tier 3 students were evaluated in 20% (n = 2) of the studies. Of the 10 ten studies, two examined outcomes of three Tiers, five examined both Tier 1 and 2, two examined both Tier 1 and a classwide and one only examined the outcome of a classwide Tier intervention.

Table 3

Tiers Included in Studies.

Tier Involved	Frequency	Percentage
Tier 1 Screened	9	90%
Tier 1 Classwide	4	40%
Tier 2	7	70%
Tier 3	2	20%

Dependent variable. Table 4 presents the specific assessment tools used to monitor student progress within the RTI programs in each of the studies (N = 10). There were two types of skills assessed across the studies reviewed: computation and applied/word problems. Of the two types, computation was assessed the most with 80% (n = 8) focusing on this skill where as the applied/word problems skill was assessed in 20% (n = 2) of the studies.

The most frequent method used to measure mathematic skills in the reviewed studies (N = 10) was M-CBM (see Table 3), and this was used in 70% (n = 7) of the studies. Other measurements used in at least 2 studies included immediate transfer, near transfer, and Woodcock Johnson III Applied Problems.

Table 4

Evaluation of RTI Programs

		Frequency	Percentage
Type of Skills Assessed	Computation	8	80%
	Applied / word problems	2	20%
Measures and Assessments	M-CBM	7	70%
	Immediate transfer	2	20%
	Near transfer	2	20%
	WJ III Applied Problems	2	20%
	Can't do / Won't do	2	20%
	Far transfer	1	10%
	WJ III Computation	1	10%
	Texas Early Mathematics Inventory-Progress Monitoring (TEMI-PM)	1	10%
	Stanford Acheivement Test - 9th edition (SAT-9)	1	10%
	Stanford Achievement Test - 10th edition (SAT-10)	1	10%
	Brigance Comprehensive Inventory of Basic Skills-Revised (CIBS-R)	1	10%
	District Math Screener Test	1	10%
	Knowing Math Test	1	10%
	Statewide Accountability Test in Mathematics	1	10%
	First-Grade Concept/Application	1	10%
	Story Problems	1	10%

Training procedure. More than half of the studies (60%, n = 6) reported training of the staff implementing the RTI procedures. The most common training provided was verbal and written instructions training (60% n = 6), followed by role play (n = 5, 50%), follow up problem solving (n = 5, 50%), coaching (40%, n = 4), performance feedback (40%, n = 4).

Tier Administration

Within the broad area of Tier program administration procedures, results of five dimensions for each Tier will be presented. These include the specific math intervention program used and the effective teaching components of the intervention, implementing agent, frequency of the program implemented, the duration of each session, and group size.

Specific math intervention program. Twelve different mathematic interventions were utilized across the ten studies. Table 5 presents data on the types of interventions used in the studies. As noted in this Table, there is a wide variety of program options and few listed programs have been replicated within a Tier.

Table 5

Type of Intervention

	CW Tier 1 n = 4		Tier II n = 7		Tier III n = 2	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Cover Copy Compare	1	10%	0	0%	0	0%
Cover Copy Compare + Instruction	0	0%	1	10%	1	10%
Flashcard practice	1	10%	1	10%	0	0%
Hot Math Schema-broadening instruction	2	20%	0	0%	0	0%
Peer Tutoring	2	20%	2	20%	0	0%
Fluency building	1	10%			1	10%
Booster Lessons			1	10%		
Knowing Math			1	10%		
Fluency building			1	10%		
Detect Practice Repair			0	0%		
Concrete-representational-abstract tutoring			1	10%		
Math Flash			1	10%		

Table 6 presents data on the different effective teaching components of the interventions. These results indicate a few components that are more consistently implemented than others within each Tier including goal setting, immediate feedback, and reinforcement. Independent practice was also included in all Tiers but varied in whether or not the practice was timed. Interestingly, more studies have examined the effects of effective teaching strategies within a classwide intervention than a Tier 3 level.

Table 6

Intervention Components

Intervention Components	Tier 1 Classwide n= 4		Tier II n= 7		Tier III n = 2	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Modeling	2	20%	2	20%	1	10%
guided practice	3	30%	5	50%	1	10%
Independent practice	3	30%	2	20%	1	10%
untimed						
Fluency practice	3	30%	2	20%	1	10%
error correction	3	30%	4	40%	1	10%
immediate feedback	4	40%	2	20%	1	10%
self grading	3	30%	1	10%	1	10%
self graphing	3	30%	1	10%	1	10%
goal setting	5	50%	2	20%	1	10%
Reinforcement	4	40%	5	50%	2	20%
Transfer	1	10%				
Pacing			1	10%		

Person Implementing. Table 7 presents data from all studies on who implemented the intervention for each tier. Among the studies (n = 4) that implemented interventions in Tier 1 classwide, the interventions were primarily administered by peer tutors (75%). Among the studies that implemented interventions in Tier 2 (n= 7), the majority of interventions were administered by the researcher (85.7%), with teachers (28.6 %) and peer tutors (14.3%) also administrating interventions. In the studies that implemented interventions in Tier 3 (n = 2), researchers were the only administrators of interventions.

Table 7

Person Implementing

		Frequency	Percentage
Tier 1 CW	Researchers	1	10%
	Peer tutors	3	30%
Tier 2	Researchers	6	60%
	Teacher	2	20%
	Peer tutors	1	10%
Tier 3	Researchers	2	20%

How often implemented. Table 8 presents data from all studies on the number of times the intervention was administered per week. Of the studies implementing in a Tier 1 classwide setting, 75% were administered four to five times per week and 25% were administered two to three times per week. Of the studies implementing in a Tier 2 setting, 28.6% did not report how often the interventions were administered. The remaining studies reported 42.9% of interventions were administered four to five times per week, and 28.6% were administered two to three times per week. Of the studies implementing in a Tier 3 setting, 50% reported administering the interventions four to five times per week, and 50% reported administering two to three times per week.

Table 8

How Often Implemented

		Frequency	Percentage
Tier 1 CW	4-5 times /week	3	30%
	2-3 times/week	1	10%
Tier 2	4-5 times /week	3	30%
	2-3 times/week	2	20%
	not mentioned	2	20%
Tier 3	4-5 times /week	1	10%
	2-3 times/week	1	10%

Duration per session. Of the twelve mathematic interventions that were utilized, Table 9 presents data from all studies on the amount of time that an intervention was administered to a student. Of the studies implementing in a Tier 1 classwide setting, 20% were administered in 31-60 minutes, 10% were administered in 15-30 minutes, and 10% were administered in 15 minutes and under. 20% of the studies did not report the amount of time spent implementing interventions. Of the studies implementing in a Tier 2 setting, 20% were administered in 31-60 minutes, 20% were administered in 15-30 minutes, and 10 were administered in 15 minutes and under. 20% of the studies did not report the amount of time spent implementing interventions. Of the studies implementing in a Tier 3 setting, 10% were administered in 31-60 minutes and 10% were administered in 15-30 minutes.

Table 9

Duration per Session

		Frequency	Percentage
Tier I Classwide	15 min and Under	1	10%
	15-30 min	1	10%
	31-60 min	2	20%
	Not Mentioned	2	20%
Tier II	15 min and Under	1	10%
	15-30 min	2	20%
	31-60 min	2	20%
	Not Mentioned	2	20%
Tier III	15-30 min	1	10%
	31- 60 min	1	10%

Group size. The number of students included in each intervention group was also coded. Table 10 presents data on group size across the different Tier settings. Of the studies implemented in a Tier 1 classwide setting, the majority (75%) were administered with peer tutoring with 25% administered in a medium group (4-6 students). Of the studies implemented in a Tier 2 setting, the majority (57%) were administered in a small group (2-4 students) with 14% administered in a peer tutoring format, 14% in a classwide format, and 14% in an individual format. Of the studies implemented in a Tier 3 setting, 50% were administered in a small group (2-4 students) format, and 50% were administered in an individual format.

Table 10

Group Size

	Group	Frequency	Percentage
Tier 1 Classwide	peer tutoring	3	30%
	Medium group (4-6)	1	10%
Tier 2	peer tutoring	1	10%
	Small group (2-4)	4	40%
	Classwide	1	10%
	Individual	1	10%
Tier 3	Small group (2-4)	1	10%
	Individual	1	10%

Frequency of response evaluation. Table 11 presents data showing the frequency that researchers collected screening and progress data for each Tier level. As shown in this table, 5 of the 7 studies that conducted Tier 1 schoolwide level screening, collected and reviewed one screening and only one study conducted screening 3 times. The frequency of the collection and review of Tier I Classwide level progress data varied across studies although 50% (n = 4) of the studies collected and reviewed the data on a weekly basis. Frequency of data collection and review also varied for Tier 2 and Tier 3 with only 4 of the 9 studies collecting data daily or weekly.

Table 11

Frequency of Response Evaluation

Frequency	Tier 1 schoolwide		Tier 1 classwide		Tier 2		Tier 3	
	N	%	n	%	n	%	n	%
Daily	0	0%	1	25%	1	14%	1	50%
Weekly	0	0%	2	50%	2	29%	0	0%
4 -5 days a week	0	0%	1	25%	0	0%	0	0%
3 Times a Year	1	10%	0	0%	1	14%	0	0%
one time for study	5	50%	1	25%	1	14%	0	0%
pre-post for study	1	10%	3	75%	2	29%	0	0%
not mentioned					1	14%	1	50%

Tier Outcome

Criterion to Identify Responders. Table 12 shows the results for the criteria used to identify responders. While 8 of the 10 studies reported the criterion, there were no consistent patterns of the criteria used to indentify responders for each Tier. The most consistent criterion used to identify responders across all Tiers was a score within an instructional range (Deno & Mirkin, 1977). Moreover, there was variable use of CBM and standardized tests for decision-making.

Table 12

*Criterion to Identify Responders**

	Tier 1 screened N= 9		Tier I CW N= 4		Tier II N=6		Tier III N= 2		Total
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	
above Instructional level (40)	2	22%	2	50%	2	33%	2	100%	8
slope positive trend	1	11%			1	17%	1	100%	3
Standard test below 40%	1	11%	2	50%					3
standard below 25%	2	22%			1	17%			3
duel discrepancy	1	11%			1	17%			2
above mastery	1	11%	1	25%					2
above frustrational	1	11%	1	25%					2
above frustrational & < 16%	1	11%							1
scored above class mean					1	17%			1
lowest on CBMs	1	11%							1
slope median split					1	17%			1
standard ACH score <80					1	17%			1
Standard test below 35%	1	11%							1
Standard test below 16%	1	11%							1
no mention					2	33%			2

**Slope is growth in academic performance over time calculated as Time 2 – Time 1 or computed using ordinary least squares*

***Frustrational, instructional and mastery are proficiency levels from literature sources Dino and Mirken (19977) and Burns, VanDerHeyden and J (2006).*

**** Dual discrepancy is a requirement that a student's level and slope performance falls below peer comparison , a proficiency level and/or a benchmark.*

Percentage of Responders. Table 13 presents the percentage of responders per Tier to the RTI programs implemented. Percentages were calculated if the number of students participating and who responded to the intervention in a Tier or classwide intervention level Tier was reported. Percent of responders for 5 of the classwide interventions, 5 of the Tier 2 interventions and 2 of the Tier 3 level interventions could be calculated. In sum, 3 interventions (2 classwide and 1 Tier 3) had 80% or more students respond and all three of these effective interventions included fluency practice.

Table 13

Percentages of Reported Students Responding to Interventions per Tier and Study

Level	Intervention	Percent responding
Classwide Level	Flashcard practice (n = 477)	89.7%
	Flashcard practice (n = 345)	84.9%
	PS, Transfer, Self-regulation (n = 201)	76.1%
Tier 2	Schema-broadening instruction + tutor (n = 164)	72.0%
	Fluency building (n = 3)	66.7%
	Guided practice (n = 32)	56.3%
	Tutoring (n = 64)	68.8%
	Tutor package (n = 51)	49.0%
	Flashcard practice (n = 5)	20.0%
Tier 3	Fluency building (n = 1)	100.0%
	Cover Copy Compare (n = 4)	75.0%

Discussion

Responsiveness to intervention within a multitier approach appears to be a promising method to obtain data to help identify math deficits and disabilities while providing various levels of intervention to remediate academic deficits. The ultimate goal of RTI is to implement continuous monitoring and intervention to prevent students from falling behind. Recently, researchers have purported advantages of using the RTI process for math assessment that include early identification of and support for at-risk students that would reduce the number of students referred for Special Education services due to severe deficits in math performance (Fuchs, Fuchs, & Hollenbeck, 2007; VanDerHeyden &Witt, 2005). The results of this review show that there are only a few studies that have investigated the effects of a RTI program on math performance. Only ten studies were identified which included an RTI program to provide math instruction in schools and evaluated math numeracy, computation, or problem solving performance. This limited number was expected as most researchers and schools have first focused on RTI outcomes on reading performance.

The primary focus of the present review was on the characteristics of the RTI programs examined in the current literature. The degree that rigorous research methods were employed in the reviewed studies to clearly demonstrate a functional relationship between the intervention and academic change was also considered in this review. An examination of the study design used in the reviewed studies indicates that several experimental methods were employed that included both group and single subject designs. However, there are no studies that examined the model using a randomized design. The internal validity of study results were supported however by the reported high levels of treatment integrity measures in all studies. Moreover, more objective, accurate measures of treatment integrity such as direct observations or taped sessions were used as opposed to indirect measures, which are often subjective and unreliable measures due to observer biases (Witt, Gresham, & Noell, 1996). This is important because it allowed for the researchers to make the conclusion that the program was implemented as planned and that the independent variable, the RTI program, was the primary influence on any measured changes in math performance (McIntyre, Gresham, DiGennaro, & Reed, 2007).

Despite the limited studies, results from this review reveal some preliminary insight on current RTI practices in the area of mathematics to guide future research and practices. But this evidence is limited first to elementary settings given that no studies were conducted in secondary school settings. And second, due to the limited amount of studies indentified, conclusions could not be drawn by grade level. Finally, many of the studies included demographics such as race (90%) and ELL (80%), however, no studies reported specific subgroup's responses. Undoubtedly future replication studies or follow-up studies should include such demographic information in order to obtain more solid conclusions regarding the likelihood of intervention success with various classroom populations. It is also important to note that 60% of the studies did not include information on the school's Title I status. This is important because Title I schools may have access to more resources to implement and evaluate RTI programs as they receive extra funding.

In consideration of these limitations, these results reveal overall that the effects of different Tier levels have been explored to some extent and the effectiveness of a different intervention programs have been investigated as part of an RTI assessment program (Fuchs, Fuchs, & Hollenbeck, 2007; VanDerHeyden & Witt, 2005). Of the studies reviewed, it is important to note that the majority of the programs intervened on computational skills (80%). This is most likely due to the importance of identifying and intervening in early mathematics difficulties when students first learn computational skills. Young students having difficulties learning computational skills may not develop the level of automaticity that are needed for becoming proficient in more advanced hierarchical mathematic skills that are based on fluent computational math skills (Bryant, Bryant, Gersten, Scammacca, & Chavez, 2008). Given that frequent progress monitoring is a key element of the RTI process, it was not surprising that the majority of these studies monitored student progress on math skills using CBM. The use of CBM is important as it is one of the few assessments in the literature that provides the ability to assess skill change with intervention within a short period of time. The fact that most research on M-CBM psychometrics and validity provides some support for the utility of progress monitoring of computational skills may also have influenced the initial main focus on computational skills within RTI programs. But it is important to note that research on M-CBM has primarily supported screening decision making on computational skills and support for the use of broader interpretations of CBM results is scarce (Christ et al, 2008).

In theory, students who receive instruction at more intensive tiers are those who require more frequent monitoring and teacher support (Crawford & Ketterlin-Geller, 2008). Thus, the trend towards more frequent administrations of CBM when implementing more intensive Tiers was expected and was generally noted. Of the studies that conducted Tier 1 schoolwide level screening, 5 of the 7 studies, collected and reviewed one screening and only one study conducted screenings 3 times. The frequency of the collection and review of progress data of Classwide level intervention increased relative to screening and was collected at least on a weekly basis in all studies investigating this level of intervention. Frequency of data collection for Tier 2 was similar to Classwide level intervention. Only one studied reported frequency of progress monitoring for Tier 3 and data were collected in this study at the most intensive daily frequency.

RTI is based on the practice of using student learning outcomes in response to instructional programs to make instructional and important educational decisions (Batsche et al., 2005). Thus, outcome criteria used to identify responders to intervention programs were explored to summarize common assessments and criterions used per Tier. Data used to screen for at risk students in Tier 1 relied on both CBM and standard tests but specific criterion vastly varied across studies. CBM data was more consistently used to identify responders to the classwide and Tier level interventions but varied between mastery, instructional or frustrational level criterion. Slope was considered in more than half the studies for Screening, Tier 2 and Tier 3 level interventions; however, the method for estimating slope was varied among studies.

An effective RTI model should consist of a quality core instruction and tiers of quality instruction that adequately addresses the needs of most of the students receiving the instruction. Educators may select from several empirically based intervention options to improve math performance (Baker et al., 2002; Codding et al., 2011; Codding et al., 2009, Fuchs, Fuchs, & Hollenbeck, 2007; VanDerHeyden & Witt, 2005). This review showed that several intervention options were examined but few were replicated. Although few specific interventions were replicated, there was positive support for general computational fluency based interventions (e.g., flashcard practice) and accurate based interventions (e.g., cover copy compare) and with peer tutoring.

Within an RTI program, it is also important that at least 80% of students receiving an intervention should be responding as expected. If fewer students are responding, then time and resources used to implement an intervention are not well utilized. Results from this review show that among the students receiving intervention the percentage of students responding to the different interventions was not consistently 80% or more at each intervention level. First, many reviewed studies implemented classwide interventions which is needed when a high percentage of students in a given classroom are not responding to the classwide instruction. Without initial attention to potential need for classwide interventions, schools will most likely struggle to find the resources necessary to address the needs of more than 20-25% of the student population who require additional support provided in more intensive Tiers. Based on study results, only flashcard practice resulted in the desired response rate (i.e., greated than 79%) for classwide level intervention. Confirming adequate response to classwide intervention also facilitates a more appropriate percentage of students who require placement in Tier 2 and Tier 3. Interestingly, the classwide interventions reported the inclusion of effective teaching components more consistently than the Tier level interventions.

Second, for Tier 2 level intervention, only there were no interventions that showed the majority of students (i.e., greater than 79%) responding within the positive response criterion set by the study. When examining intervention components, guided practice, error correction and reinforcement were key components of Tier 2 interventions. Perhaps additional effective teaching components (e.g., modeling, fluency practice) may be needed to increase the percentage of student who responds to this level of intervention.

Third, support for Tier 3 interventions was limited in this review. Only two interventions (cover copy compare and fluency building) were investigated in a Tier 3 program. Based on study results, only fluency building resulted in the desired response rate (i.e., greated than 79%) for Tier 3 intervention. The findings on specific intervention programs highlight a variety of options that address several types of problems such as targeting accuracy and fluency but clearly there is a need for more studies exploring and replicating the success of RTI programs within each intervention level. Sufficient support for students struggling in math will only be achieved when it is known which interventions best optimizes performance for at-risk students.

When selecting interventions, feasible implementation of interventions by teachers is a substantial concern of the RTI process. Thus, the reasonability of intervention implementation was examined by reviewing the percentage of interventions that were successfully implemented by teachers. Results reveal that researchers were very involved in the implementation of RTI programs. Given that only 28% of interventions occurred with a teacher, effective implementation in the classroom setting by the general education teacher still remains a significant concern in implementing RTI. There is some preliminary evidence, however, that interventions may be feasible when interventions are provided to students with peer tutors.

Another important feature of feasibility and success of intervention is the initial training provided to teachers that leads to high integrity levels. Research on intervention training suggests several effective teaching strategies that result in high treatment fidelity over time. For example, studies on the effects of training on intervention implementation suggest that classroom training with classroom rehearsal and feedback (Sterling-Turner, Watson, & Moore, 2002), immediate and faded delayed feedback during the in-class training sessions (LaFleur, Witt, Naquin, Harwell, & Gilbertson, 1998), and a brief weekly supportive feedback meeting to review implementation barriers, child progress data, and determine intervention modification or fading strategies (Noell et al., 2005) increases and maintains accurate implementation of a classroom intervention. While a majority of the studies trained using verbal and written instructions (60%) many of the studies trained using effective strategies that are more time intensive as well, that included role play (50%), follow up problem solving (50%), coaching (40%), and performance feedback (40%). It appears that fairly intensive training may be an important feature to an effective RTI program. Yet, the fact that many studies were able to achieve this intensive level suggests that this type of training may be feasible at least in elementary settings.

In this review of literature, intervention intensity, as defined by intervention session time and frequency, was also reviewed to examine feasibility. It was determined that intervention intensity did not increase as expected as students were provided intervention in higher Tiers for the number of sessions. That is, for classwide interventions the majority of interventions were administered 4 to 5 times a week while Tier 2 and Tier 3 was more evenly distributed between 4-5 and 2-3 times a week. Number of students served within a group and section duration generally increased with more intensive Tiers. The majority were administered in a small group (2-4 students) in Tier 2 whereas 50% were administered in a small group format and 50% were administered in an individual format in Tier 3. The majority of interventions in classwide sessions that reported session duration were completed between fifteen and sixty minutes, with only 10% being administered in 15 minutes or less. The majority of interventions in Tier II session that reported session duration were completed between fifteen and sixty minutes, with only 10% being administered in 15 minutes or less. Of the studies that reported session duration during Tier III administered interventions, all were between fifteen and sixty minutes.

In summary, a prospective advantage of the RTI approach is to replace the traditional refer-test-place model of identification and eligibility consisting of a dual general education and Special Education system to intervene-test-intervene-place consisting of multiple Tier programs (VanDerHeyden et al., 2004). This review was an attempt to summarize RTI math program research to provide tentative guidelines on effective program components and practices. Given that the research is just emerging there are still not definitive guidelines for schoolwide RTI programs. Researchers are making great efforts to assure that the researched RTI model is implemented with fidelity but studies using a randomized design and investigating RTI models for middle and high schools to increase math performance are still required. Yet, these results reveal that multiple Tiers are being conducted with improved student learning in the area of math using various types of interventions. Ensuring the development of mathematics competence for all students as they advance to more advance skills is essential to ongoing success.

References*

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision.* Washington, D.C.: American Psychiatric Association.

*Ardoin, S. P, Witt, J. C., Connell, J. E., & Koenig, J. E., (2005). Application of a three-tiered response to intervention model for instructional planning, decision making, and the identification of children in need of services. *Journal of Psychoeducational Assessment 23*, 362 – 380.

Baker, S., Gersten, R., & Lee, D., (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *Elementary School Journal, 103* (1), 51-73.

Batsche, G., Elliot, J., Graden, J. L., Grimes, J., Kovaleski, J. F., Prasse, D., Reschley, D. J., Schrag, J., & Tilly, W. D., III (2005). *Response to intervention: Policy considerations and implementation.* Alexandria, VA: National Association of State Directors of Special Education.

Berkeley, S., Bender, W. N., Peaster, L. G., & Saunders, L. (2009). Implementation of Response to Intervention: A Snapshot of Progress. *Journal of Learning Disabilities, 42* (1), 85-95.

*Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N., & Chavez, M. M. (2008). Mathematics Intervention for First- and Second-Grade Students With Mathematics Difficulties: The Effects of Tier 2 Intervention Delivered as Booster Lessons. *Remedial and Special Education, 29,* 20-32.

Burns, M. K. & Gibbons, K. (2013). *Response to intervention implementation in elementary and secondary schools: Procedures to assure scientific-based practices* (2nd edition). New York: Routledge.

Burns, M. K., Riley-Tillman, T. C., & VanDerHeyden, A. M. (2013). *Advanced RTI applications: Intervention design and implementation.* New York: Guilford.

Christ, T. J., Scullin, S., Tolbize, A., & Jiban, C. L. (2008). Implications of Recent Research : Curriculum-Based Measurement of Math Computation. *Assessment for Effective Intervention, 33,* 198-205.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education,* 38, 136-163.

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disability Research & Practice, 26,* 36-47.

Codding, R. S., Hilt-Panahon, A., Panahon, C., & Benson, J. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and Treatment of Children,* 32, 279–312.

Deno, S. L. (2003). Developments in curriculum-based measurement. *Journal of Special Education*, *37*, 184–192.

Deno, S. L., & Mirkin, P. (1977). *Data-based Program Modification: A Manual.* Reston VA: Council for Exceptional Children.

Donovan, M. S., & Cross, C. T. (Eds.). (2002). *Minority students in special and gifted education.* Washington, DC: National Academy Press.

*Duhon, G. J., Mesmer, E. M., Atkins, M. E., Greguson, L. A., & Olinger, E. S. (2009). Quantifying intervention intensity: A systematic approach to evaluating student response to increasing intervention frequency. *Journal of Behavioral Education 18*, 101-118.

Freeland, J. & Noell, G. H., (1999). Maintaining accurate math responses in elementary school students: The effects of delayed intermittent reinforcement and programming common stimuli. *Journal of Applied Behavior Analysis, 32* (2), 211-15.

Fuchs, D., Mock, D., Morgan, P. L., Young, C. L., (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18* (3), 157-171.

Fuchs, L.S. (2003). Assessing Intervention Responsiveness: Conceptual and Technical Issues. Learning Disabilities Research & Practice, 18: 172-186.

*Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology 97* (3), 493-513.

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

*Fuchs, L.S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L, . & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology, 100,* 491–509. Grégoire, J. & Desoete, A. (2009). Mathematical Disabilities: An Underestimated Topic? *Journal of Psychoeducational Assessment 27*, 171-174

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* (1), 13-24.

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

*Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities, 37* (4), 293-306.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, 25, 33-45.

Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first grade children's mathematical development with peer-assisted strategies. School Psychology Review, 31, 569-583.

Griffin, S., & Case, R. (1997). Re-thinking the primary school math curriculum: An approach based on cognitive science. *Issues in Education*, 3, 1-49.

*Ketterlin-Geller, L. R., Chard, D. J., & Fien, H. (2008). Making connections in mathematics. *Remedial and Special Education, 29* (1), 33-45.

Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs. *Remedial and Special Education 24* (2), 97-114.

LaFleur, L., Witt, J. C., Naquin, G., Harwell, V., & Gilbertson, D. (1998). Use of coaching to enhance classroom management by improvement of student transitioning between classroom activities. *Effective School Practices,* 17, 70-82.

Manzo, K. K., & Galley, M. (2003). Math climbs, reading flat on '03 NAEP. Education Week, 23(12), 1-18.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of school-based interventions with children in the *Journal of Applied Behavior Analysis* 1991-2005. *Journal of Applied Behavior Analysis, 40*(4), 737-741.

Mellard, D., Byrd, S., Johnson, E., Tollefson, J., & Boesche, J. (2004). Foundations and research on identifying model responsiveness to intervention sites. *Learning Disability Quarterly,* 27, 1-14.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: National Council of Teachers of Mathematics.

National Mathematics Advisory Panel. *Foundations for Success: The Final Report of the National Mathematics Advisory Panel,* U.S. Department of Education: Washington, DC, 2008.

Newman-Gonchar, R., Clarke, B., & Gersten, R. (2009). A summary of nine key studies: Multi-tier intervention and response to interventions for students struggling in mathematics. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Noell, G. H., Gilbertson, D. M., VanDerHeyden, A. M., & Witt, J. C. (2005). Eco-Behavioral Assessment and Intervention for Culturally Diverse At-Risk Students. In C. L. Frisby & C. R. Reynolds (Eds.) *The Comprehensive Handbook of Multicultural School Psychology* (pp. 904-927). NJ: Wiley.

Perie, M., Grigg, W. S., & Dion, G. S. (2005). *The nation's report card: Mathematics 2005, national assessment of educational progress* (Publication NCES 2006453, U.S. Department of Education, National Center for Education Statistics). Washington, DC: Government Printing Office.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2009). Effects of fact retrieval tutoring on third-grade students with math difficulties with and without reading difficulties. *Learning Disabilities Research & Practice, 24*, 1–11.

Riccomini, Paul J.; Witzel, & Bradley S. (2010). Response to intervention in math. Thousand Oaks, CA, US: Corwin Press.

Shinn, M. R. (Ed.). (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.

Silver, C. H., Pennet, B. D., Black, J. L., Fair, G. W., & Balise, R. R. (1999). Stability of arithmetic disability subtypes. *Journal of Learning Disabilities, 32,* 108-109.

Sophian, C. (2004). Mathematics for the future: Developing a Head Start curriculum to support mathematics learning. *Early Childhood Research Quarterly*, 19, 59-81.

Stecker, P., Fuchs, L. S., Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42* (8),795-819.

Sterling-Turner, H. E., Watson, T. S., & Moore, J. W. (2002). Effects of training on treatment integrity and treatment outcomes in school-based consultation. *School Psychology Quarterly, 17*, 47-77.

Sugai, G. (2007). Promoting behavioral competence in schools: A commentary on exemplary practices. *Psychology in the Schools, 44* (1), 113-118.

U.S. Department of Education, National Center for Education Statistics. *The Condition of Education 2003*, NCES 2003-067. Washington, DC: 2003.

VanDerHeyden, A. M., Broussard, C., Fabre, M., Stanley, J., Legendre, J., & Creppel, R. (2004). Development and validation of curriculum-based measures of math performance for four-year old children. *Journal of Early Intervention,* 27, 27-41.

*VanDerHeyden, A. M. & Burns, M. K. (2005). Using Curriculum-Based Assessment and Curriculum-Based Measurement to guide elementary mathematics instruction. *Assessment of Effective intervention, 30,* 15-31.

*VanDerHeyden, A. M. & Burns, M. K. (2009). Performance indicators in math: Implications for brief experimental analysis of academic performance. *Journal of Behavioral Education 18*, 71-91.

VanDerHeyden, A. M. & Burns, M. K. (2010). *Essentials of response to intervention.* New York: Wiley.

VanDerHeyden, A. M., & Witt, J. C. (2005). Quantifying the context of assessment: Capturing the effect of base rates on teacher referral and a problem-solving model of identification. *School Psychology Review, 34*, 161-183.

*VanDerHeyden, A. M., Witt, J. C., & Barnett, D. W. (2005). The emergence and possible futures of response to intervention. *Journal of Psychoeducational Assessment*, 23, 339-361.

Witt, J. C., Gresham, F. M., & Noell, G. H. (1996). What's behavioral about behavioral consultation? *Journal of Educational and Psychological Consultation. 7, 327-344.*

*included in the reviewed studies