UCONN LIBRARY

University of Connecticut
DigitalCommons@UConn

NERA Conference Proceedings 2008

Northeastern Educational Research Association (NERA) Annual Conference

10-23-2008

Improving Automaticity with Basic Addition Facts: Do Taped Problems Work Faster Than Cover, Copy, Compare?

James Cressey
University of Massachusetts - Amherst, jamescressey@gmail.com

Kristin Ezbicki
Hillsboro-Deering Cooperative School District, NH, krisezbicki@yahoo.com

Follow this and additional works at: http://digitalcommons.uconn.edu/nera_2008

Part of the Curriculum and Instruction Commons, and the Science and Mathematics Education Commons

Recommended Citation
Cressey, James and Ezbicki, Kristin, "Improving Automaticity with Basic Addition Facts: Do Taped Problems Work Faster Than Cover, Copy, Compare?" (2008). *NERA Conference Proceedings 2008*. 12.
http://digitalcommons.uconn.edu/nera_2008/12

Improving Automaticity with Basic Addition Facts:

Do Taped Problems Work Faster Than Cover, Copy, Compare?

James Cressey

University of Massachusetts Amherst

Kristin Ezbicki

Hillsboro-Deering Cooperative School District, NH

Basic Math Fact Automaticity

The recent report of the National Math Panel (2008) indicated that fostering automaticity with basic math facts should be a focus point for mathematics educators. Automaticity is an important skill set for students to develop as they progress towards more advanced applications of mathematics. Automaticity, or fluency, is defined as the accuracy and speed with which a student can perform a simple operation (addition, subtraction multiplication, or division), usually with numbers from 0-12.

A strong positive correlation exists between basic math fact automaticity and overall mathematics performance (Thurber, Shinn, & Smolkowski, 2002; Tronsky & Royer, 2003). Students who reach fourth grade and still rely on inefficient addition strategies that involve counting are unlikely to have reached automaticity in solving basic addition problems. As they begin to be expected to learn multiplication problems, these strategies become even more inefficient and prone to error. Further on, advanced mathematical challenges become more and more daunting without the basic skills that should have been mastered in earlier grades.

Response to Intervention for Mathematics

The emerging approach of Response to Intervention (RTI) as a model for providing mathematics instruction requires sensitive measurement tools for screening, progress monitoring, and decision-making. Mathematics Curriculum-Based Measurement (M-CBM) has been developed in a number of varieties, and has shown good psychometric properties, sensitivity, and predictive screening utility (Fuchs, Hamlett, & Fuchs, 1999; Fuchs, Fuchs, Compton, Bryant, Hamlett, & Seethaler, 2007; Hintze, Christ, & Keller, 2002; Shinn, 1989; Thurber, Shinn, & Smolkowski, 2002). M-CBM provides practitioners with a useful tool for evaluating the skills and progress of students with the mastery of basic mathematical calculations.

Data on students' basic math fact automaticity can also be used in order to make instructional decisions within a multi-tiered model of services. Schools that operate using a 3-tiered model provide differentiated mathematics instruction based on student's individual needs. The three tiers of intervention typically progress from Tier I to Tier III with increasing levels of intensity, frequency, duration, and with different approaches to instruction (Fuchs & Fuchs, 2006; Fuchs, et al., 2007). While Tier I consists of a universal core curriculum that all students are provided with, Tiers II and III require supplemental interventions targeting specific skill gaps or objectives. The RTI model of multi-tiered instruction emphasizes the importance of using interventions with a research basis, or other evidence or data indicating that the interventions will be effective.

A key ingredient for a successful system of mathematics RTI is a menu of supplemental Tier II and Tier III interventions, with research or practical evidence of their effectiveness. In the area of early literacy, teachers often have numerous interventions available to them, in the domains of phonological awareness, vocabulary, phonics, fluency, and comprehension. However, in mathematics there are fewer interventions for struggling students and less research to determine what types of interventions will be most likely to succeed. Teachers sometimes find that even their core mathematics curriculum does not emphasize basic math fact automaticity, and they must turn to teacher-created resources or supplemental workbooks. These circumstances make it almost impossible for teachers to find and implement research-based interventions for building automaticity.

Automaticity Research

One practice method that has been researched is Cover, Copy, Compare (CCC; Poncy, Skinner, & Jaspers, 2006). Students are taught a self-managed independent practice procedure with five steps: (1) look at a mathematics problem with the answer and study it; (2) cover the mathematics problem with the answer; (3) record the answer on the other side of the page; (4) uncover and compare your answer; (5) correct your answer if it was incorrect. Several variations on the CCC intervention have been studied, including overcorrection of errors, contingent reinforcement for correct responses, and group vs. individual delivery. CCC has been demonstrated to increase accuracy and automaticity in students with learning and behavioral problems (Skinner, McLaughlin, et al., 1997; Skinner, et al., 1989; Skinner, et al., 1993; Stading & Williams, 1996; Poncy, Skinner, & Jaspers, 2006; Codding, et al., 2007).

Taped Problems (TP) is another method of providing basic skill practice with immediate error correction (McCallum, Skinner, Turner & Saeckler, 2006). In the TP method, students listen to an audiotape of math problems and attempt to write the correct answer quickly, before it is provided by the tape. If they answered incorrectly, or did not have time to write an answer, students write down the correct response after they hear it on the tape. A "varying time delay" method was used by the researchers in repeated trials. Different time delays between the problem and the answer are used in each set of problems. At first, the delay is very short (< 1 second) in order to discourage students from using inefficient strategies such as counting. Then, the delay is increased to 4 seconds in order to encourage more independent responding. Finally, the delay is shortened to 2 seconds for the final two sets to promote faster, more automatic responding.

McCallum, Skinner, Turner & Saeckler (2006) conducted the TP intervention solely with single digit multiplication facts from 2-9. Using single-skills CBM probes as an outcome measure, the researchers found large effect sizes ranging between .86 and 1.6 (Cohen's *d*) with a general education classroom of 3rd graders. The TP intervention has been found to improve accuracy and automaticity with basic division in a 4th grade student (McCallum, Skinner, & Hutchins, 2004). The TP method was also recently compared to the CCC intervention by Poncy, Skinner, & Jaspers (2007) with single digit addition. Both the TP and CCC interventions were found to be effective for the student in this study, although the TP intervention took significantly less time (by approximately 29%).

Further research is needed to evaluate the effectiveness of these interventions under different conditions, with different participants, and in different settings. The comparison of CCC and TP interventions produced positive results for both methods, but was only conducted with one subject, a 10 year old girl with mental retardation (Poncy, Skinner, & Jaspers, 2007). The question remains whether or not group administration and group research design would yield similar results.

The purpose of this preliminary pilot study was to extend the work of Poncy, Skinner, & Jaspers (2007) comparing the CCC and TP interventions, using a group design method rather than a single subject method. In particular, we sought to apply these methods to the promotion of automaticity with 0-12 addition facts. We wanted to investigate the relative effectiveness of CCC and TP in comparison to a control condition that did not involve basic math fact practice, and in comparison to each other over time. We were curious to observe whether or not the TP condition might produce faster results than CCC or control conditions, because of the time pressure element of the TP intervention. Stated formally, the hypotheses of this study are as follows: Given 20 days of intervention (80 problems/day, 5 days/week),

1. Students in the CCC and TP groups will make more progress towards automaticity with basic addition facts than students in the control group.
2. Students in the TP condition will develop higher levels of automaticity more rapidly than students in the CCC and control conditions.

Method

Setting and Participants

This study was conducted at a short-term residential treatment program for children with serious emotional and behavioral challenges. Drawn from an original sample of 81 students, 51 participants were included in the final analysis, including 8 girls and 43 boys. The participants ranged in age from 6 to 14, with a mean age of 10.2, and represented a heterogeneous population of students, many of whom come from a large urban public school district. Close to 100% of the students are eligible for special education services under the category of emotional and behavioral disability.

Fourteen classroom groups were selected for the study and assigned randomly to the three conditions of the experiment, assuring that each of the three conditions would have roughly the same number of students. The TP group included students from three different classrooms, the CCC group consisted of students from four different classrooms, and the control group consisted of students from five different classrooms.

The classroom teachers administered all assessments and interventions. Participating teachers had all been hired as summer teaching interns for the program, ranged in age from 19-27, and had between 0 and 3 years of experience working with children.

Dependent Measure

Automaticity was measured using single-skill addition Mathematics Curriculum-Based Measurement (M-CBM) probes (Shinn, 2005). These M-CBM probes are two-minute timed, group-administered tests of computation. Each probe presents 84 addition problems using addends from 0-12, arranged in 6 rows of 7 problems on the front and back of the page. The total number of correct digits (CD) written in two minutes is calculated. Thus, if a student has an incorrect response, but some of the digits are written in the correct place, the student gets partial credit for their correct digits. The student's CD score is then divided by 2 to calculate their number of digits correct per minute (DCPM), which was the score used for the purposes of this study.

M-CBM has demonstrated internal consistency, test-retest reliability, and interscorer agreement of .93 (Shinn, 2005; Fuchs, Fuchs, & Hamlett, 1988). Additional research found that M-CBM has an alternate form reliability of .91 (Thurber, Shinn, & Smolkowski, 2002). We felt that this test was a valid choice for measuring automaticity with basic addition facts, based on the existing body of research as well as school-based applications that make use of M-CBM for this purpose.

Procedure

Taped Problems

The Taped Problems (TP) condition was modeled after the procedure used by researchers who originally developed and studied it (McCallum, Skinner, & Hutchins, 2004; McCallum, Skinner, Turner & Saeckler, 2006; Poncy, Skinner, & Jaspers, 2007). An independently created version of the taped problems and answers was developed by the researchers, using a sound recording and editing software program. The audiorecordings were made by the researchers using software that provided a digital measurement of elapsed time, in order to make the time delays as precise as possible. The audiorecordings were rendered onto CDs for use in the classrooms. Sets of 20 addition problems and answers were generated from an online worksheet creator (www.themathworksheetsite.com). Each addition problem consisted of two randomly selected numbers from 0-12.

The same 20 problems were used to make four parallel sets of taped problems with problem-answer delays of 1, 4, 2, and 1 seconds. Teachers implemented the TP intervention 5 days a week, at the same time every day. The TP audio files were played aloud for the students to listen to as a group and complete their answer sheets simultaneously, but independently. The four audio files were played one after the other. In total, the TP intervention lasted for about 10 minutes each day. Students listened to a scripted set of instructions (see Appendix A) explaining that they should try to write the answer to each problem before it is given on the CD. Students were told to write the correct answer down if they wrote an incorrect answer or did not have time to answer it. Students were monitored by teachers and staff during the intervention to ensure that they understood and were following instructions.

Cover, Copy, Compare

The Cover, Copy, Compare (CCC) condition was designed to deliver practice in a different method from the TP intervention, while keeping other factors between the conditions the same. Thus, the same four 20-problem sets were used in the CCC and TP conditions. The CCC intervention was also delivered 5 days a week, for about 10 minutes each day. Students were given the same pages of addition problems that TP students were given (four copies of the same page). In addition, CCC students were given one answer key to be used for the intervention. They were given scripted instructions (see Appendix B) that explained the following procedure. Students were told to read the answer to the first problem, cover the answer sheet, and write the answer to the problem on their worksheet. Then, the students should uncover the answer sheet to check their answers. If the answer is incorrect, they should write the correct answer in. These steps are repeated for each problem on the page. Students worked on the CCC pages for about 10 minutes each day, with the expectation of completing all 80 problems. Teachers and staff monitored them to ensure that students understood and followed directions.

Control Group

The control group condition was matched to the other two groups on the amount of time spent on mathematics instruction each day, in order to control for the effects of time spent on mathematics in general. Teachers in these classes implemented a math warm-up activity that took about 10 minutes as well. Most of the warm-ups were drawn from a book called *Read it, Draw it, Solve it* that presents students with a word problem, which could involve addition, subtraction, multiplication, or division. Students read the problem, drew a picture to illustrate the word problem scenario, and then solved the problem. At times, students were also prompted to write a sentence to describe the illustration and the solution. There were no components of timed practice or explicit basic addition fact practice in these activities, in order to avoid similarities between the control condition and the experimental conditions on this important factor.

Treatment Integrity

Teachers were trained on the CCC and TP intervention procedures before the students arrived. Both interventions were scripted in order to standardize the intervention delivery. Teachers were trained on the assessment procedures during the staff orientation week at the beginning of the summer, before the students arrived for the program. The researcher modeled the procedures, guided teachers through the steps, and gave teachers opportunities to practice. Teachers were given positive feedback as well as corrective feedback and additional practice opportunities in training.

Data Analytic Plan

This study used a mixed design with between-groups and within-subjects comparisons to test our hypotheses. M-CBM probes were administered three times: as a Pretest, at the end of Week 2, and at the end of Week 4. The data analysis for the study began with calculating three difference scores for each subject to show changes in automaticity from Pretest to Week 2, from Pretest to Week 4, and from Week 2 to Week 4. Pairwise comparisons were conducted to test these difference scores for significant differences between groups.

Because the subjects were not randomly assigned to groups, we chose to use *t* test comparisons of the difference scores, rather than a covariate analysis such as ANCOVA. The use of ANCOVA with nonequivalent groups results in an inflated Type I error rate, and although a procedure correcting for unreliability can correct this bias, it also reduces power. Using *t* test comparisons of the difference scores avoids these problems with bias and power when working with nonequivalent groups and repeated measures (Peller, Wells, & Matthews, 2006).

The Holm procedure to control family-wise error was selected as the best method for evaluating the significance of these comparisons. In addition to these tests of statistical significance, effect sizes were calculated using Cohen's *d* in order to evaluate and compare the magnitude of intervention effects. These analyses allowed us to evaluate the two hypotheses of the study: that TP and CCC would yield significantly higher levels of automaticity than the control condition, and that TP would produce faster results than the other conditions.

Results

Descriptive Statistics

Table 1 presents the results of the M-CBM probes at Pretest, after two weeks of intervention, and after four weeks of intervention. Mean and standard deviation scores are reported for each of the three conditions at each time period, and are also presented in visual form in Figure 1.

Table 1
Mean M-CBM Scores and Standard Deviations

	Pretest	Week 2	Week 4
Taped Problems (*n*=14)	19.43(14.53)	22(10.79)	22.04(15.26)
Cover, Copy, Compare (*n*=18)	27.39 (13.93)	23.06(9.52)	26.08(11.16)
Control Group (*n*=19)	30.26(15.83)	23.04(14.83)	32.42(13.96)

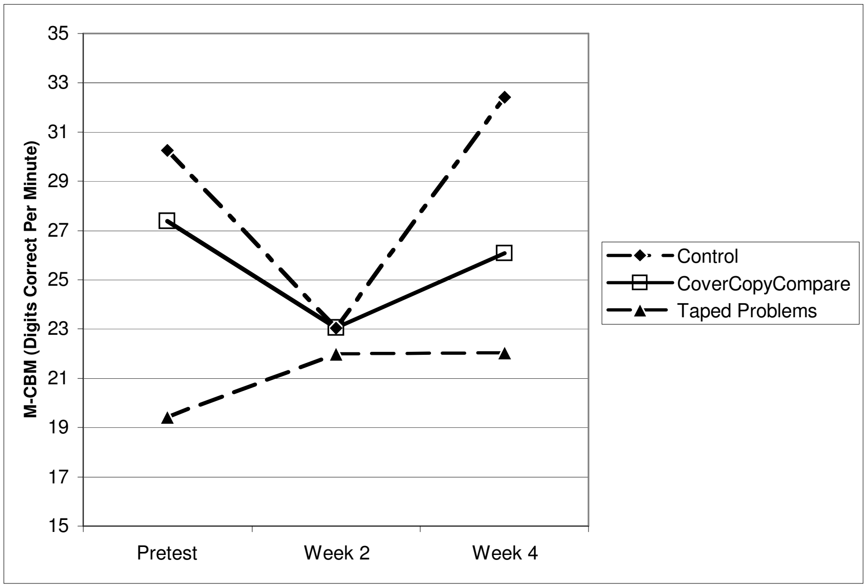


Figure 1
Means Plot

The next step in our analysis was to calculate three difference scores for each subject. These three difference scores represent the change in each subject's M-CBM scores from Pretest to Week 2, from Pretest to Week 4, and from Week 2 to Week 4. The mean and standard deviation of the difference scores were calculated within each group. These results are shown in Table 2.

Table 2
Mean Difference Scores and Standard Deviations

	Pretest to Week 2	Pretest to Week 4	Week 2 to Week 4
Taped Problems (*n*=14)	2.57(11.06)	2.61(9.41)	.04(7.35)
Cover, Copy, Compare (*n*=18)	-4.33(12.71)	-1.31(10.12)	3.03(13.25)
Control Group (*n*=12)	-2.09(5.49)	2.16(6.86)	3.84(6.14)

Comparing Mean Difference Scores

In order to answer our primary research questions, we needed to make pairwise comparisons between the group mean difference scores in the three conditions. We conducted multiple contrasts between the groups using each set of difference scores, yielding a total of nine *t* tests. Table 3 presents the results of these *t* tests and their *p* values. Comparisons that included the control group were evaluated with one-tailed significance tests, while comparisons between the intervention groups were evaluated using two-tailed tests.

Table 3
Pairwise Comparisons of Difference Scores

Comparison	*t*	*p*
(Pretest vs. Week 2)		
TP vs. CCC	1.186	.077
TP vs. Control group	1.111	.137*
CCC vs. Control group	-.564	.288*
(Pretest vs. Week 4)		
TP vs. CCC	1.244	.220
TP vs. Control group	.145	.443*
CCC vs. Control group	-1.193	.120*
(Week 2 vs. Week 4)		
TP vs. CCC	-.839	.406
TP vs. Control group	-.967	.170*
CCC vs. Control group	-.218	.414*

*1-tailed tests

None of these comparisons resulted in a significant difference between groups. Because none of the *p* values were less than .05, the Holm procedure was not required to determine significance.

Effect Sizes

Table 4 shows the magnitude of effect, or lack thereof, that was observed in the three conditions. These standardized mean difference effect sizes were calculated by comparing the group means at Pretest and Week 4, using the pooled standard deviation as a standardizer. While the Taped Problems condition resulted in a small positive effect size, the Cover, Copy, Compare condition resulted in a slightly negative effect size. The control group also experienced a small positive effect.

Table 4
Effect Sizes

Pretest vs. Week 4	*d*
Taped Problems	0.175
Cover, Copy, Compare	-0.10
Control Group	0.145

Discussion

Research Hypotheses

The results showed that our primary hypothesis was not supported by the data. We predicted that the intervention conditions would lead to significantly more improvement in automaticity than the control condition. The Taped Problems group did reach a higher level of automaticity after the intervention, increasing by 2.61 DCPM from Pretest to Week 4. However, the Cover, Copy, Compare group showed a slightly lower post-intervention mean score, decreasing by 1.31 DCPM over the course of four weeks. The control condition resulted in an increase of 2.16 DCPM during this time period. Although we can state that the Taped Problems condition increased by slightly more than the control condition, this was not a statistically significant difference. Effect sizes calculated to show the change in automaticity did not provide support for the hypothesis that the interventions led to improvement in automaticity after four weeks.

Only partial support was found for the second hypothesis, in which we predicted that the Taped Problems group would improve more rapidly than the Cover, Copy, Compare group. We can observe that from Pretest to Week 2, the midpoint of the intervention, the Taped Problems group improved by 2.57 DCPM, while the Cover, Copy, Compare group experienced a decrease in mean score by 4.33 DCPM. However, our data analysis revealed that there was no evidence of a statistically significant difference between the magnitude of these changes (p=.077).

Despite these findings, no group significantly regressed in their levels of basic addition fact automaticity. Knowing that summer recess often results in a regression of academic skills from June to September, perhaps we can conclude that these interventions (including the control condition) helped students to maintain their skills during the summer months and avoid regression.

Limitations

This pilot study was conducted under circumstances with a moderate amount of experimental control, but some factors went unchecked and most likely detracted from the internal validity of the study. Treatment integrity was not checked in a systematic way, due to a lack of resources. Anecdotally, at least four teachers reported that they did not do the intervention on some days. Treatment integrity was most likely low overall, because of the lack of systematic checks, as well as the challenges inherent to a population of students with behavioral problems and a staff of teachers with little to no prior experience.

Statistical power was another limitation. While the a priori power analysis indicated that the sample size would be adequate to find an interaction between time and group factors using an omnibus test, the amount of power in the study was inadequate for the purposes of multiple comparisons between the groups. The original sample of children included all of the students in the age range of 6-14, which was 83. This larger sample size would have yielded more statistical power; however, attrition proved to be another problem during data collection. Three entire classrooms were dropped from the study after the first M-CBM administration, when it became clear that a majority of the class was already proficient with basic addition facts. A number of individual students were not included in the final analysis because they were unwilling or unable to complete the second and third M-CBM probes. The final sample, N=51 students, yielded less power than was originally desired.

Another important limitation to note was the use of a single M-CBM probe at each administration, instead of three. Had it been feasible to administer three probes and take the median score at each administration, a more reliable estimate of automaticity would have been gathered. This is a common approach in research and in school practice that would have improved the strength of this study.

In looking at the Pretest M-CBM levels, we can see that the three groups did not start out with identical levels of automaticity. This is due, in part, to our inability to use randomization during group assignment. Certain classrooms of students were already at higher levels of automaticity than others. These differences may have resulted in a regression to the mean effect taking place over time. Other possible threats related to these differences would include selection bias, and a selection by treatment interaction. For example, students in the Taped Problems condition, who began with the lowest levels of automaticity, may have been placed in classrooms together based on their academic skill deficits. These particular students may have been slower to develop automaticity, regardless of practice method, than students from classrooms in the control group classrooms, who may have been placed in class together based on their higher levels of academic performance.

Future Directions

The limitations discussed here point the way for several improvements that can be made in future studies of this type. The research questions asked in the this study are still in need of exploration and research, using high quality methods and procedures. Mathematics interventions need to be explored and compared in many ways as we continue to develop resources for a Response to Intervention model of service delivery. Teachers are still looking for more research-based mathematics interventions that can help struggling students. Future research into the Taped Problems and Cover, Copy, Compare practice methods might attempt to zero in further on the differences between the methods. Aside from the time pressure difference, one method includes an audio component, while the other does not. Adding a time pressure to the Cover, Copy, Compare method would control for that factor, while exploring the difference between visual and auditory presentation of correct responses.

While the population of students with serious behavioral and emotional problems presented challenges to conducting a study with adequate treatment integrity and experimental control, this is still an important and worthwhile effort. Students from this population are often targeted for social, emotional, and behavioral research. However, it is equally important to discover teaching methods that will be successful with them in academic domains. Research demonstrating that this population of children can learn and can be taught academic skills is an important step towards improving educational services for all students.

Appendix A

Taped Problems Instructions

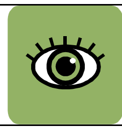
Listen,	Write,	Check
		

	1. **LISTEN** to the math problem on the recording.
	2. Try to **WRITE** the answer quickly, before you hear it on the recording.
	3. **CHECK** the answer when you hear it on the tape and fix any errors or **WRITE** the answer if you didn't have time to do it before.

- Problem set A….. HIGH SPEED! You only have one second to write each answer. This is to see which ones you have memorized.
- Problem set B….MEDIUM. You have 4 seconds to write each answer. This gives you more time to try answers on your own.
- Problem set C……FAST. You have 2 seconds to write each answer. This will help you start remembering answers faster.
- Problem set D….HIGH SPEED again. Did you get a little better at this with practice

Appendix B

Cover, Copy, Compare Instructions

Look,	**Say,**	**Cover,**	**Write,**	**Check**
				

	1. **LOOK** at the math problem with the answer.
	2. **SAY** the problem and answer quietly ("eight plus seven is fifteen")
	3. **COVER** up the problem and answer with an index card or by folding the paper over.
	4. **WRITE** the problem and the answer from memory.
	5. **CHECK** the problem and answer to see if you did it correctly. Fix any mistakes.

Repeat these steps for each math problem you are practicing.

References

Codding, R.S., Lewandowski, L., & Eckert, T. (2005). Examining the efficacy of performance feedback and goal-setting interventions in children with ADHD: A comparison of two methods. *Journal of Evidence-Based Practices for Schools, 6 (1),* 42-58.

Codding, R.S., Shiyko, M., Russo, M., Birch, S., Fanning, E., and Jaspen, D. (2007). Comparing mathematics intervention: does initial level of fluency predict intervention effectiveness? *Journal of School Psychology, 45,* 603-617.

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., Compton, D. L., Bryant, J. D., Hamlett, C. L., & Seethaler, P. M. (2007). Mathematics screening and progress monitoring at first grade: Implications for responsiveness to intervention. *Exceptional Children, 73*(3), 311-330.

Fuchs, L. S., Fuchs, D. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly 41*(1), 93-99.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Karns, K. (1998). High-achieving students' interactions and performance on complex mathematical tasks as a function of homogeneous and heterogeneous pairings. *American Educational Research Journal, 35*(2), 227-267.

Fuchs, L.S., Hamlett, C.L., & Fuchs, D. (1999). *Monitoring Basic Skills Progress (MBSP): Concepts and Applications* [Blackline Masters] Austin, TX: PRO-ED.

Hintze, J.M., Christ, T.J., & Keller, L.A. (2002). The generalizability of CBM survey level mathematics assessments: Just how many samples do we need? *School Psychology Review*, 31, (4), 514-528.

McCallum, E., Skinner, C.H., & Hutchins, H. (2004). The taped-problems intervention: increasing division fact fluency using a low tech self managed time delay intervention. *Journal of Applied School Psychology, 20 (2),* 129-147.

McCallum, E., Skinner, C.H., Turner, H., & Saeckler, L. (2006). The taped-problems intervention: increasing multiplication fact fluency using a low tech, classwide, time delay intervention. *School Psychology Review, 35 (3),* 419-434.

National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education: Washington, DC.

Peller, S., Wells, C.S., & Matthews, W.J. (2007). *A comparison of ANOVA versus ANCOVA for analyzing pretest/posttest data.* Paper presented at the 2007 Northeast Educational Research Association conference.

Poncy, B.C., Skinner, C.H., & Jaspers, K.E. (2007). Evaluating and comparing interventions designed to enhance math-fact accuracy and fluency: cover, copy, and compare versus taped problems. *Journal of Behavioral Education 16 (1),* 27-37.

Poncy, B.C., Skinner, C.H., & O'Mara, T. (2006). Detect, practice, and repair: The effects of a classwide intervention on elementary students' math-fact fluency. *Journal of Evidence Based Practices for Schools, 7 (1),* 47-68.

Royer, J.M., and Tronsky, L.N. (1998). Addition practice with math disabled students improves subtraction and multiplication performance. *Advances in Learning and Behavioral Disabilities, 12*, 185-217.

Royer, J. M., & Tronsky, L. N. (1998). Addition practice with math disabled students improves subtraction and multiplication performance. In T. E. Scruggs, & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities, vol. 12.* (pp. 185-217). US: Elsevier Science/JAI Press.

Royer, J.M., Tronsky, L.N., Chan, Y., Jackson, S.J., & Marchant III, H.G. (1999). Math fact retrieval as the cognitive mechanism underlying gender differences in math achievement test performance. *Contemporary Educational Psychology, 24,* 181-266.

Shinn, M.R. (1989). *Curriculum based measurement: Testing special children.* New York, NY: Guilford.

Shinn, M. R. (2005). AIMSweb® Training Workbook: Administration and scoring of mathematics curriculum-based measurement (M-CBM) for use in general outcome measurement. [On-line]. Available: http://www.aimsweb.com.

Skinner, C. H., Bamberg, H. W., Smith, E. S., & Powell, S. S. (1993). Cognitive cover, copy, and compare: Subvocal responding to increase rates of accurate division responding. *RASE: Remedial & Special Education, 14*(1), 49-56.

Skinner, C. H., McLaughlin, T. F., & Logan, P. (1997). Cover, copy, and compare: A self-managed academic intervention effective across skills, students, and settings. *Journal of Behavioral Education, 7*(3), 295-306.

Skinner, C. H., Turco, T. L., Beatty, K. L., & Rasavage, C. (1989). Cover, copy, and compare: A method for increasing multiplication performance. *School Psychology Review, 18*(3), 412-420.

Stading, M., Williams, R. L., & McLaughlin, T. F. (1996). Effects of a copy, cover, and compare procedure on multiplication facts mastery with a third grade girl with learning disabilities in a home setting. *Education & Treatment of Children, 19*(4), 425-434.

Thurber, R.S., Shinn, M.R., and Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum based measures of mathematics. *School Psychology Review, 31 (4),* 498-513.

Tronsky, L.N. (2005). Strategy use, the development of automaticity, and working memory involvement in complex multiplication. *Memory & Cognition, 33* (5), 927-940.

Tronsky, L.N. & Royer, J.M. (2003). Relationships among basic computational automaticity, working memory, and complex mathematical problem solving. What we know and what we need to know. In J.M. Royer (Ed.) *Mathematical cognition* (pp. 117 - 146). Greenwich, CT: Information Age Publishing.