School Psychology Forum:
RESEARCH IN PRACTICE

VOLUME 1 • ISSUE 1 • PAGES 28–45 • NOVEMBER 2006

Evaluating and Comparing Responsiveness to Two Interventions Designed to Enhance Math-Fact Fluency

Erin Carroll, Christopher H. Skinner, and Haley Turner
The University of Tennessee, Knoxville
Elizabeth McCallum
Duquesne University
Sarah Woodland
Knox County Schools

Abstract: Response-to-intervention models of service delivery are designed to identify, prevent, and remedy students' academic skill deficits, including mathematics skills deficits. Although educators have developed procedures for enhancing math skills, further research is needed to establish interventions that are both efficient and effective for students functioning at a range of abilities. Researchers used an adapted alternating treatments design to evaluate and compare responsiveness to two interventions intended to improve the addition—fact fluency of a student with mild mental retardation. During cover, copy, compare (CCC), a student was instructed to read a list of math problems and answers, cover each problem and answer, write the problem and answer, and check her response. During the taped-problems intervention (TP), the student received a packet of problems and was instructed to complete each problem before the answer was provided by a corresponding audiotape. A third set of problems served as a control set. To allow for precise comparison of learning rates across interventions, 7.5 minutes were allotted for each intervention. Results showed that both interventions affected increases in addition fluency, with TP yielding more rapid increases. Discussion focuses on the need to provide practitioners with empirically validated interventions as well as the need to compare interventions using precise measures of learning rates in order to identify more effective and efficient interventions.

Overview

With the reauthorization of the Individuals with Disabilities Education Improvement Act of 2004, models for identifying, preventing, and remedying learning problems have shifted toward a new model, known as responsiveness to intervention (RTI). The traditional model of determining the presence of a learning disability involves looking for a within-child discrepancy, which occurs when a student's achievement level is significantly lower than would be expected given his or her cognitive ability. Like the discrepancy model, the RTI approach consists of examining low-achieving students but instead mandates the use of a process to determine if these students respond to research-based instruction or interventions. Under an RTI model, the failure to respond sufficiently

*Correspondence regarding this article should be directed to Erin Carroll, Department of Educational Psychology and Counseling, The University of Tennessee, Knoxville, TN 37996 (**erincarroll@hotmail.com**).*

to high-quality instruction is one of the primary factors for determining whether a child has or is at risk for a learning disability (Fuchs, Mock, Morgan, & Young, 2003).

A primary purpose of the RTI model of service delivery is to help identify and remedy skill deficits before they become debilitating. Early development of math skills may enhance development of more complex math skills and students' success in other academic and career areas (e.g., science, computers, engineering). Unfortunately, a recent report by the National Assessment of Educational Progress found that merely 36% of fourth-grade students and 30% of eighth-grade students are performing at a proficient level in math (Perie, Grigg, & Dion, 2005). Although much RTI research has focused on preventing and remedying reading difficulties, less attention has been paid to mathematics. A search of the PsycINFO database using the terms *response/responsiveness to intervention* and *math* produced 20 results whereas a search of *response/responsiveness to intervention* and *reading* produced 179 results. The shortage of research on how to apply the RTI model to math skill deficits may hinder practitioners' attempts to identify, prevent, and remedy math difficulties.

Researchers have identified and validated some basic components needed to implement an RTI model designed to prevent and remedy math problems. Specifically, researchers have (a) collected evidence that identifies central target behaviors, (b) developed efficient and sensitive procedures that allow for repeated measurement of target behavior development, and (c) validated numerous interventions that may enhance these target behaviors (Deno & Mirkin, 1977; Martson, 1989).

Target Behavior

Evidence suggests that developing rapid and accurate responding to basic math facts, including single-digit addition, subtraction, multiplication, and division, may prevent and remedy learning problems (Skiba, Magneusson, Mastron, & Erickson, 1986). The terms *fluency* and *automaticity* describe this level of skill development (Deno & Mirkin, 1977; Haring & Eaton, 1978; Hasselbring, Goin, & Bransford, 1987). More advanced math tasks require the application of cognitive resources such as working memory and attention, both of which are limited (Carpenter & Moser, 1982; Gagne, 1982; Pellegrino & Goldman, 1986). Because many of these complex math concepts and tasks require the use of basic facts, students who must expend their cognitive resources by correctly responding to basic facts will have fewer resources available to apply to the development of the advanced skills and concepts (Rathvon, 1999). Additionally, students who can complete basic facts rapidly and accurately may have less math anxiety and be more likely to choose to do advanced math tasks because they require less effort and can be completed more quickly, resulting in a thicker schedule of reinforcement (Cates & Rhymer, 2003; Skiba et al., 1986; Skinner, 2002). Because advanced skill development requires students to choose to engage in academic behaviors, these studies clearly support the need to enhance basic-fact fluency or automaticity (Skinner, Pappas, & Davis, 2005).

Measurement Procedures

In addition to identifying target behaviors, researchers have developed procedures for measuring basic-fact fluency. Nearly 30 years ago, Deno and Mirkin (1977) described curriculum-based measurement (CBM) procedures for assessing math fluency. These procedures are brief, allow for repeated measurement, and are highly sensitive to small increases in basic-fact fluency. CBM probes allow educators to measure students' fluency by assessing digits correct per minute (DCM). Using these sensitive DCM data, researchers have been able to develop and empirically validate many fluency-building interventions including cover, copy, compare (CCC), taped problems (TP), flashcard drill, timed assessment procedures (e.g., brief 1-minute assessments), and peer tutoring (Greenwood, Delquadri, & Carta, 1997; Hayden, & McLaughlin, 2004; Kroesbergen & Van Luit, 2003; McCallum, Skinner, & Hutchins, 2004; McLaughlin, Skinner, & Logan, 1997; Skinner, Ford, & Yunker, 1991; Skinner, Shapiro, Turco, Cole, & Brown, 1992).

Intervention Selection

Under previous models of service delivery, intervention selection has been guided by within-student strengths and weakness. Under RTI models, a major criterion for intervention selection is that interventions should be evidence-based or empirically validated (Stoiber & Kratochwill, 2000). Therefore, interventions should meet several criteria: appropriate for empirical study, relevant to the problem or difficulty a student is experiencing, effective, and efficient. One purpose of the current study was to continue the process of empirically validating a promising math fluency-building intervention (TP) by comparing it to an established procedure with a stronger evidence base (CCC).

When using an RTI model, the knowledge that a student's skills are in need of remediation comes with the recognition that the student has already fallen behind his or her peers. To prevent the student from falling farther behind, the goal should be to remedy those skills as quickly as possible. Therefore, when considering which intervention to apply, practitioners must consider which intervention results in the *most rapid* increases in skill development (Skinner, Belfiore, & Watson, 2002/1995). For example, assume that one intervention caused an increase in fluency of 10 DCM in 2 weeks, whereas a different intervention caused an increase of 20 DCM in 2 weeks. It would appear that the second intervention is more effective. However, now consider that the first intervention required a total of 50 minutes (5 minutes per day) and the second required a total of 200 minutes (20 minutes per day). The first intervention yielded an increase of .2 DCM per minute of instructional time (e.g., 10 digits correct per minute increase/50 minutes of instructional time). However, the second intervention, which appeared to be superior when instructional time was not taken into account, yielded an increase of only .1 DCM per minute of instructional/learning time. Thus, if the practitioner's goal is to remedy the problem as quickly as possible, the intervention that results in the most rapid increase skill development (e.g., the increase of .2 DCM for each minute of instruction) is superior and should be selected.

The Two Interventions

The TP intervention is similar to the taped-words intervention, a procedure that occasions high rates of accurate, active, academic responding by having students read words along with an audiotape (Bliss, Skinner, & Adams, in press; Freeman & McLaughlin, 1984). McCallum et al. (2004) adapted the taped-words intervention for use with math facts and incorporated varying time-delay procedures to prevent and/or discourage the use of time-consuming finger counting strategies. Only two studies have evaluated the TP procedure. In the first study, researchers used a multiple-probes-across-tasks design to demonstrate that TP could increase the division-fact fluency of a general education fourth-grade student (McCallum et al., 2004). More recently, TP was applied class-wide to improve the multiplication-fact fluency in a third-grade classroom (McCallum, Skinner, Turner, & Saecker, in press). Although TP has been used to enhance basic math-fact fluency both individually and on a class-wide basis, researchers have not investigated the efficacy of TP for enhancing math fluency in students with disabilities.

Another intervention shown to enhance math-fact fluency is CCC. Originally used to increase spelling rates, CCC has been adapted for a number of academic subjects (McLaughlin & Skinner, 1996). CCC requires a student to look at an academic stimulus (e.g., a math problem and answer), cover the stimulus, make an academic response (e.g., write the problem and answer), and check the response against the original stimulus. If the response is correct, the student moves on to the next problem. If the response is incorrect, the student copies the problem and answer three times. Researchers have shown that CCC is effective for enhancing math fluency in general education students (Skinner et al., 1992), students with behavior disorders (Skinner, Bamberg, Smith, & Powell, 1993; Skinner, Belfiore, Mace, Williams-Wilson, & Johns, 1997; Skinner, Turco, Beatty, & Rasavage, 1989), and students with learning disabilities (Hayden & McLaughlin, 2004; Smith, Dittmer, & Skinner, 2002).

Purpose

One purpose of the current study was to demonstrate how researchers could use CBM procedures to evaluate responsiveness to interventions targeting basic math facts. Curriculum-based procedures provide frequent and direct measures of a math skill, such as the number of digits answered correctly per minute on a series of multiplication problems. Increases in these data over time (and exposure to intervention) suggest an improvement in skill whereas no improvement or a decreasing trend in the data suggest that an alternative intervention may be necessary (Shapiro, 1996).

A second purpose of the current study was to contribute to research on empirically validated interventions. When academic interventions are first developed, the goal is often to establish their effectiveness (a students' responsiveness to that intervention) by demonstrating that they enhance skill development. Once this occurs, subsequent studies are then needed to determine if the intervention is superior to other empirically validated procedures. Thus, we sought to compare the effects of a promising intervention taped problems (TP) with moderate empirical support (e.g., only two studies) with an established intervention cover copy compare (CCC). By holding instructional time constant across the TP and CCC interventions (7.5 minutes per day), we provide practitioners and researchers with a demonstration of how to precisely measure and compare learning rates across interventions. This procedure may allow educators and school psychologists who are implementing RTI models to select, from available evidence-based interventions, the intervention that results in the *most rapid* remediation of skill deficits.

Method

Student and Setting

Amber, a 12-year-old African-American student enrolled in a general education fifth-grade classroom, participated in the study. Prior to the study, a child study team had begun the process of assessing Amber for special education. Shortly after the study began, the child study team determined that Amber was mildly mentally retarded. Amber was referred for intervention services by her teacher, a female with approximately 3 years experience. Intervention services were provided by the primary researcher, a second-year school psychology doctoral student. The teacher indicated that Amber had strong multiplication and division skills (i.e., was automatic with those basic facts) but often used finger-counting strategies when computing addition problems. This behavior made it difficult for Amber to work quickly and efficiently through multistep math problems that required addition (e.g., complex multiplication problems). Her teacher also expressed concern that Amber's math skill development was being hindered by her reliance on finger counting.

Prior to beginning the study, the primary researcher collected class-wide data on multiplication- and division-fact fluency. Three probes were administered for each skill (multiplication and division). The data indicated that Amber was completing a median (median of the three probes) of 29 (range = 22–32) DCM in division and 53 (range = 45–53) DCM in multiplication. Based on Deno and Mirkin's (1977) criteria, these skills were at the instructional (i.e., 20–39 DCM) and mastery levels (i.e., over 40 DCM), respectively. Furthermore, her division and multiplication skills were above her class's median scores (22 DCM for division and 34 DCM for multiplication). However, Amber's median scores for addition and subtraction were less well developed: 14 DCM for addition (range = 13–23) and 16 DCM for subtraction (range = 14–20), both of which fall in Deno and Mirkin's frustration range (i.e., less than 20 DCM). The class medians were 33 DCM for addition and 27 DCM for subtraction, both of which fall within the instructional range. These data supported the teacher's assertion that Amber had not developed fluent addition skills and was achieving at a level below her peers.

Researchers conducted the current study in Amber's classroom after her classmates had left the room for nonclassroom educational activities (e.g., gym, music). The primary researcher conducted all procedures sitting near the student, who was seated at a desk. Amber's teacher was present in the room during most sessions and typically remained at her desk. Upon completion of each session, the researcher escorted Amber to the appropriate activity.

Materials

Researchers used a tape recorder, headphones, and stopwatch throughout the experiment. Baseline and intervention assessment data were collected via experimenter-constructed probes. Basic addition facts (adding digits 1–10) were divided into three mutually exclusive sets: A, B, and C (see Appendix A). The sets did not contain any problems with a digit of 0 (e.g., 0 + 5 = __). Problems were randomly assigned to each set, and the smaller number was always listed first (e.g., 2 + 4, 5 + 8, 3 + 7). Set A and Set B contained 18 problems and Set C contained 19 problems.

Four different assessment probes were constructed for each set of problems. Each of the four probes for the respective set of problems (Set A, Set B, Set C) was created by randomly ordering the problems. After listing each problem once, the problems were repeated again in the same order. Corresponding intervention sheets were constructed for each set of problems by randomly sequencing the problems and repeating the set of problems four times. Thus, there were 72 problems on Set A and Set B and 76 problems on Set C. Set B was randomly assigned to the TP intervention, Set C to CCC, and Set A to a no-treatment control condition.

During the TP, the student received a four-page packet containing 72 problems. The problems were numbered 1–72, and a blank line for the student's response followed each problem. The primary researcher constructed an audiotape by recording instructions for the TP procedure and then reading the 72 problems and answers in the same sequence in which they were presented in the packet. The number of each problem was read immediately before the problem. The interval between reading the answer to one problem and the number of the next problem was held constant at 3 seconds. A varying time-delay procedure was employed on this audiotape. The first time the problems were read (i.e., problems numbered 1–18) there was no delay between the problem and answer. The second time the set was read (i.e., problems 19–36) there was a 3-second delay between problem and answer, and the third time the set was read (i.e., problems 37–54) there was a 2-second delay between problem and answer. The fourth and final time the set was read (i.e., problems 55–72), there was a 1-second delay between problem and answer. The instruction to stop was recorded after the last (i.e., problem 72) answer was read. The tape ran for 7.5 minutes.

Researchers constructed similar packets for the CCC intervention. Each four-page packet contained the 19 problems assigned to Set C. Each packet contained the set of problems repeated four times in the same order for a total of 76 problems. CCC intervention worksheets contained each problem and its answer as well as three blank spaces for responding. Although CCC does not require an audiotape, a tape was constructed for the current study. The tape provided instructions for the CCC intervention, instructed the student to begin, and then instructed the student to stop after 7.5 minutes. The audiotapes allowed the researchers to keep procedures (e.g., instructions, headphones, and use of audiotape) and time spent working constant across CCC and TP. This allowed for a true comparison of learning rates (Skinner et al., 1997; Skinner et al., 2002/1995).

Experimental Design and Dependent Measures

Researchers used an adapted alternating treatments design (AATD) to evaluate responsiveness to the two interventions (Sindelar, Rosenberg, & Wilson, 1985). This design allows for the comparison of two distinct interventions on equivalent sets of instructional items during the same time period. Because time is held constant across both interventions, precise differences in learning rates can be established when performance on one set of stimuli (e.g., mutually exclusive sets of math problems) exceeds performance on another set of stimuli (Skinner et al., 2002/1995). Within an AATD, a control set of stimuli assigned to an assessment-only condition allows for an additional evaluation of each intervention. In the current study, researchers intermittently assessed performance on the control set of problems (Set A) in order to obtain measures of history effects (e.g., variables outside the study) and spillover effects (learning under one condition that carried over to performance on the other condition) that were less susceptible to testing effects caused by repeated measurement (Skinner & Shapiro, 1989).

DCM, errors per minute, and percent of problems correct on 1-minute assessment probes served as the dependent measures. The primary researcher calculated digits correct using Deno and Mirkin's (1977) scoring procedure. To be scored as correct, a digit had to be written in the correct place; that is, an answer of 12 to the problem 4 + 8= ___ would be scored as two digits correct. An answer of 10 or 22 would be scored as one digit correct, and an answer of 21 or 5 would be scored as zero digits correct. A digit was counted as an error if it was an incorrect response or if the digit was omitted. For example, the problem and response of 6 + 10 = 6 would be counted as one digit correct and one error. The problem and response of 3 + 5 = 13 would be scored as two errors. Researchers calculated percent of problems correct by dividing the number of correct problems (problems with no errors), by the number of problems attempted. If the last problem was only partially completed it was omitted from percent of problems correct scoring.

Procedures

Procedures were run on Monday, Wednesday, and Friday mornings. If Amber was absent or if school was not in session on that day, procedures were conducted on the next possible day. The primary researcher entered Amber's math class shortly before 11:00 a.m. when the class was dismissed to non-classroom educational activities. The researcher directed Amber to a desk, conducted the intervention activities, and then walked Amber to the appropriate activity.

During baseline, Amber completed assessments of all three sets of problems. On the first day of intervention, both CCC and TP procedures were described in detail. Amber completed both interventions, and her performance was assessed immediately after each intervention. These baseline procedures took approximately 30 minutes. Following this initial session, interventions were counterbalanced across days, and Amber received only one (either TP or CCC) intervention each day. The primary researcher assessed Amber's performance on the appropriate set of problems immediately after the intervention. Amber's performance on the control set of problems (Set A) was assessed once per week prior to administering the intervention for that day.

Assessment procedures. Baseline data were collected during the first three sessions. The primary researcher measured Amber's performance on each of the three sets of problems. She received three assessment probes each day: one for Set A, one for Set B, and one for Set C. The sequence of probes was counterbalanced across days. Amber was instructed to work from left to right and to complete as many problems as possible in 1 minute. A stopwatch was used to time each assessment and Amber was instructed to stop after 1 minute.

Taped problems. Amber received a four-page TP intervention packet and was instructed to put on the headphones and turn on the tape recorder. The tape instructed Amber to write down her response after a problem was read on the tape but before she heard the answer. Amber was instructed to evaluate her response after hearing the correct answer on the tape. If her response was incorrect, she was to put a line through the wrong answer and write in the correct one. She was told to follow along with the tape and not to skip ahead. If she did not answer on time (i.e., before the answer was read on the tape), she was instructed to write in the correct answer as given on the tape and wait for the next problem. After the last problem had been read, the recording instructed her to press the stop button on the tape recorder. During the initial session, the primary researcher provided these instructions to Amber, answered any questions, and then turned on the tape where the instructions were repeated. Following this initial session, only the tape provided the instructions. Immediately following the tape, the primary researcher conducted an assessment of Set B. The researcher randomly selected one of four Set B probes, which was administered and scored in the same manner used during baseline. See Appendix B for a guide to conducting the TP intervention.

Cover, copy, compare. Amber received the four-page CCC packet and was instructed to put on the headphones and turn on the tape recorder. Instructions on the tape explained the CCC procedure: Amber was instructed to read a problem and answer, cover the problem and answer with her hand, copy the problem and answer, and compare her response with the original problem. If the response was correct, Amber was to move on to the next problem. If the response was incorrect, Amber was instructed to copy the problem and answer in the three spaces provided. She was instructed to continue working until the tape said to stop. Although CCC does not require an audiotape, instructions were tape recorded and then the tape was silent until 7.5 minutes elapsed. At this time, the tape instructed the student to stop. This assured that the time was held constant across both interventions. During the first session, the researcher supplied instructions, answered any questions, and then turned on the tape, where the instructions were repeated. For all subsequent sessions, only the tape provided the instructions. Immediately after the tape, the researcher administered a randomly chosen Set C assessment probe. Procedures were identical to those used in baseline and following the TP intervention. See Appendix C for a guide to conducting the CCC intervention.

In order to enhance cooperation, the primary researcher provided Amber with a sticker after each session and a new pencil each week. Amber was praised after each session for trying her best but did not receive specific feedback about her performance during the study. At the conclusion of the intervention, the primary researcher showed Amber a graph of her performance, which the researcher shared with Amber's teachers. However, Amber did not see any graphs or records of her performance until after the intervention. Researchers were interested in Amber's opinion of the interventions, specifically whether she found one to be more enjoyable than the other. Thus, 2 days after the last intervention session, the primary researcher asked Amber to complete a brief student acceptability form (see Table 1).

Table 1. Student Acceptability Form

Please answer the following questions by checking how much you agree or disagree with the statement. If you strongly agree with it, check the line under "I agree." If you strongly disagree with the statement, check the line under "I do not agree." If you are not sure how you feel, check one of the middle three lines.

	I agree				I do not agree
1. I enjoyed working on my math skills.	X	____	____	____	____
2. I think the strategies improved my addition skills.	X	____	____	____	____
3. I missed out on class when working on my addition skills.	____	____	____	____	X
4. I think other students would like these strategies.	X	____	____	____	____
5. I liked seeing the graph of my results.	X	____	____	____	____

Which strategy did you like better, the taped problems or the cover, copy, compare?

____ Taped problems ____

Interscorer Agreement and Procedural Integrity

A second researcher independently scored digits correct, errors, and percent of problems correct for 22% of the assessment probes. Interscorer agreement was calculated by dividing the number of agreements on digits correct by the number of disagreements plus agreements and multiplying by 100. The same procedure was used to calculate agreement for errors and for percent of problems correct. Interscorer agreement was 100% across all measures.

The primary researcher followed a sequence of 12 procedures when conducting sessions (see Appendix D). The researcher's self-recording of steps suggested that all steps were completed in the prescribed sequence, with the exception of one CCC session (between sessions 11 and 12 on the figures). On this day, Amber was called out of the room following the CCC intervention, but prior to administering the assessment probe. Therefore, steps 7–11 (assessment procedures) were not implemented.

Results

Table 2 displays the phase means and standard deviations across each of the three dependent variables. Figures 1–3 display data on DCM, errors per minute, and percent of problems correct across baseline and intervention phases, respectively. Figure 1 shows that during the baseline phase, Amber's DCM for all three sets of problems were much more variable than during the intervention phase. After the interventions were applied, immediate increases in DCM were evident across all three sets. A comparison of Amber's performance on the two treatment sets of problems with the control set of problems showed greater gains on both TP and CCC problems than the assessment-only problems. A comparison of performance on TP and CCC problems shows that TP resulted in higher gains in digits correct per minute. Furthermore, as the intervention progressed, Amber's DCM on the TP problems showed an increasing trend, while her performance on the CCC problems leveled out and showed no improvement following the first two intervention sessions. Although Amber's performance on the control set showed an initial increase, subsequent assessments showed a decreasing trend, which suggests that history, carryover, or testing effects did not account for the increasing performance on the TP or CCC problems. These results suggest that both CCC and TP increased DCM. However, TP was more effective than CCC and continued to result in increases in DCM as the intervention progressed.

Table 2. Means, standard deviations, and effect sizes for digits correct per minute, errors per minute, and percent of problems correct across the taped problems intervention, cover, copy, compare, and the control condition.

	Dependent measure		
Intervention condition	DCM	EM	PPC
Control (Set A)			
Baseline	21.67 (1.25)	11.33 (5.44)	55.67 (17.99)
Intervention	25 (7.07)	9.0 (0.82)	71.67 (8.26)
TP (Set B)			
Baseline	18.33 (7.59)	11 (2.45)	49 (2.94)
Intervention	42.5 (2.99)	0.83 (1.07)	96.67 (4.27)
CCC (Set C)			
Baseline	15.33 (4.92)	5.67 (2.05)	57.33 (13.02)
Intervention	34.8 (.98)	2.2 (2.4)	92.2 (8.35)

Note. DCM, digests correct per minute; EM, errors per minute; PPC, percent of problems correct; TP, taped problems intervention; CCC, cover, copy, compare.

Figure 1. Digits correct per minute.

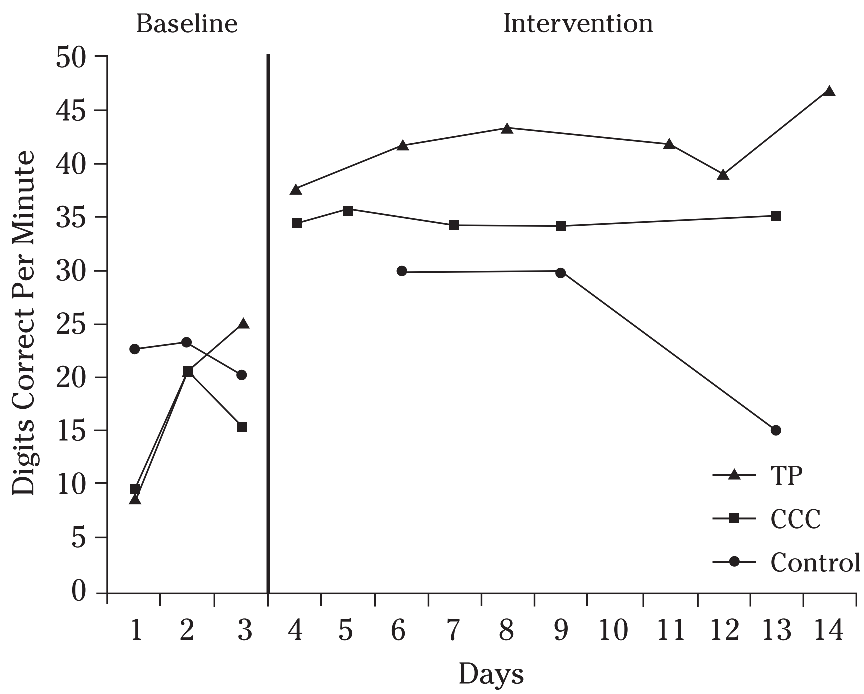

Figure 2. Errors per minute.

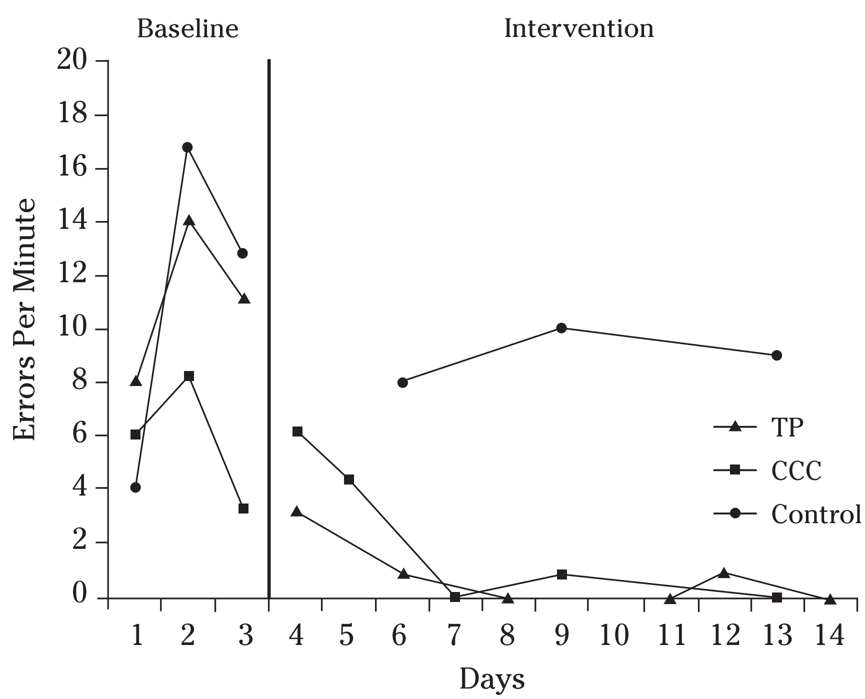

Figure 3. Percent problems correct.

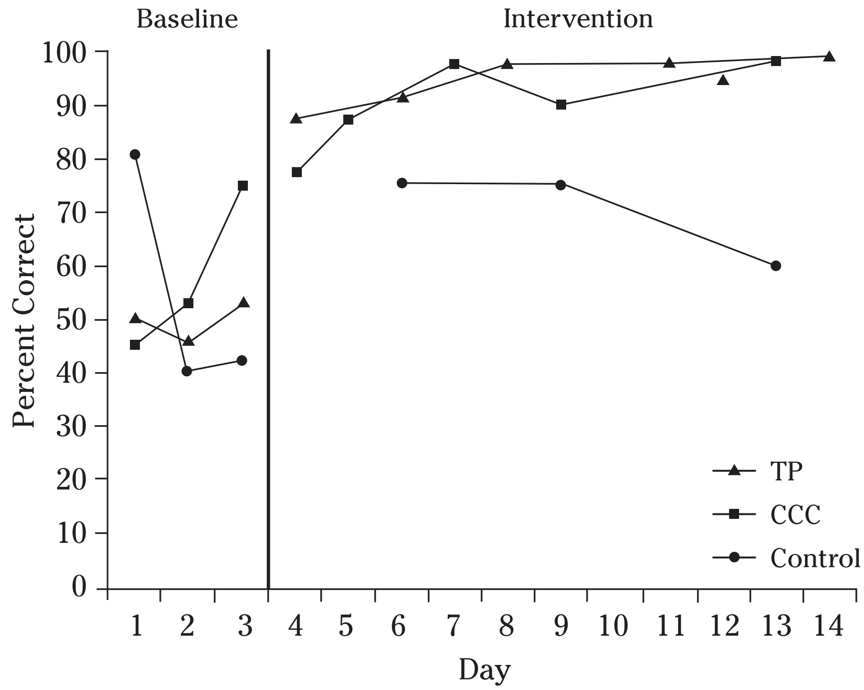

Figure 2 shows variable levels of errors per minute during baseline across all three sets of problems, with Amber making the fewest errors per minute on the CCC set of problems during baseline. Both TP and CCC resulted in reductions in errors per minute with TP resulting in a more rapid decrease during the first two intervention sessions. Although the difference in learning rates across CCC and TP was small, Figure 2 shows that both TP and CCC were superior to the assessment-only condition, suggesting that both interventions were effective in decreasing errors per minute.

Data on percent of problems correct (Figure 3) are similar to those found for errors per minute . Baseline data were highly variable, but Amber showed rapid and similar improvements in accuracy on Set B and Set C (TP and CCC, respectively). Although Amber showed initial improvements in accuracy on the assessment-only set of problems (Set A), the overall decreasing trend suggests that carryover, practice, and history effects did not account for the increases in accuracy under problems assigned to either intervention. Thus, these data suggest that both interventions resulted in increased problem accuracy, with neither being superior. Amber's responses to the student acceptability form administered after the intervention phase ended are displayed in Table 1. Amber's responses suggested that she found both TP and CCC acceptable, but preferred TP.

Overall, across all three sets of problems, the initial change on the assessment-only set of problems suggests that carryover effects or some other threat to internal validity (e.g., reactive arrangements) may have accounted for some of the enhanced performance under either or both of the interventions. However, because both CCC and TP consistently resulted in superior performance compared to the assessment-only problems, the current study suggests that both were effective for enhancing Amber's addition skills. Furthermore, across all three dependent variables, TP appeared to result in more rapid increases in addition-fact fluency than CCC.

Discussion

One purpose of the current study was to demonstrate CBM procedures in evaluating responsiveness to math fluency interventions. Although the current study appears to demonstrate the usefulness of CBM measures, the results also indicate several limitations when these measures are compared with CBM reading measures (such as words read correctly per minute). Words correct per minute (WCM) can be considered a general outcome measure of reading growth. Research suggests that this measure is sensitive to changes in general reading skill development across grades 3–6. Furthermore, although WCM may become less sensitive to reading skill development after grade 6, it still correlates well with general measures of reading skills (Neddenriep, Skinner, Hale, Oliver, & Winn, in press; Skinner, Neddenriep, Bradley-Klug, & Ziemann, 2002). The current study targeted basic addition facts. The problem with such targets is the limited pool of stimulus items (i.e., math facts) relative to the large pool of words that serve as stimuli for measuring WCM across grade levels. While the smaller and more specific pool of math-fact stimuli may make fluency development easier, it also suggests that more frequent assessments of math skill development may be needed to evaluate responsiveness to intervention. Additionally, as students master facts, educators will have to alter stimulus item sets (e.g., go from addition to subtraction facts) more frequently. While these issues are not insurmountable, they do suggest that RTI models of basic math-fact fluency may require different assessment procedures than RTI reading models.

The success of RTI models of prevention and remediation requires that we have high-quality, empirically validated, effective interventions. The current study extends the research on the TP procedure (McCallum et al., 2004; McCallum et al., in press) by demonstrating that TP can be used to enhance addition-fact accuracy and fluency in a student with mental retardation. Additionally, we added to the research on TP by comparing its effects with another empirically validated intervention, CCC. Although the difference between the two interventions was not large, TP resulted in more rapid learning than CCC across all three dependent variables.

The current study also provides additional data supporting the effectiveness of the CCC procedure (Hayden & McLaughlin, 2004; Skinner et al., 1989). Although the current study suggests that CCC caused sustained and cumulative increases on response accuracy (e.g., errors per minute and percent of problems correct), the immediate changes in fluency or speed of accurate responding (DCM) were not followed by additional increases in DCM as the study progressed.

There are several plausible causal mechanisms that may explain why TP resulted in greater increases in learning rates than CCC. One possible explanation for Amber's superior performance under TP may be the number of problems completed in that intervention. During each TP intervention session, Amber responded to 72 problems in 7.5 minutes. During CCC, Amber's average rate of responding during the intervention phase was 58 problems in 7.5 minutes. Previous research on opportunities to respond (Skinner et al., 1997) suggests that Amber's rapid learning rate under TP may have been caused by these additional opportunities to respond. However, this should be interpreted cautiously. In the current study, the student completed more problems (had more opportunities to respond) under TP than CCC, but this may be different for other students. Regardless, the more rapid rate of responding occasioned by TP allowed Amber more opportunities to respond which may account for the superior increases in fluency relative to CCC.

Amber's rate of responding was also higher under TP (approximately 10 trials per minute) relative to CCC (approximately 8 trials per minute). Researchers have shown that increasing the pace of responding can enhance accuracy (e.g., Van Houten & Little, 1982; Van Houten & Thompson, 1976). Future researchers should attempt to determine if the more rapid pace of responding, as opposed to the total number of responses, accounted for the increase in learning. Again, this may differ across students, as some students completing an intervention such as CCC may work at a faster pace compared to the pacing provided by an audiotape under TP. If pace is causally related to learning, researchers should determine if reducing the intervals between problems (i.e., reducing intertrial intervals) and/or reducing the time delays between problem and answer enhances the effectiveness and efficiency of TP (Skinner et al., 2002/1995).

TP was designed to incorporate varying time-delay procedures to prevent finger counting (McCallum et al., 2004). Time delay is not a component of CCC, so future researchers should conduct studies to determine if the time-delay procedures accounted for the superior increases in learning rate under TP. Also, future researchers should attempt to determine which time delays and sequence of delays result in the most rapid and sustained increases in fluency.

During the TP intervention, Amber's performance was paced by the audiotape. During CCC Amber paced herself through the work. Future researchers should investigate whether the audiotape pacing caused greater increases in learning rates than self-pacing. Under TP, the pace of the audiotape was constructed to prevent Amber from employing finger-counting strategies. Prior to implementing the intervention, Amber's teacher was concerned that she would attempt to employ finger counting strategies and fail to beat the tape, which could lead to frustration and her failing to keep pace with the audiotape. This did not occur as Amber worked steadily along with the audiotape. Furthermore, as opposed to being frustrated with TP procedures, acceptability data showed that Amber rated TP as acceptable and preferred it to CCC.

Although the current study showed that CCC improved performance on all three measures, researchers should attempt to determine why CCC failed to increase DCM following the first intervention session. Perhaps the rapid pacing occasioned during the TP intervention transferred (e.g., served as a setting event or establishing operation) to the assessment conditions. Regardless, educators should consider providing performance feedback on the intervention (e.g., feedback on how many CCC problem were completed) and/or performance feedback on assessment problems (e.g., feedback on DCM following assessment) to determine if such procedures enhance skill development (e.g., Eckert, Ardoin, Daly, & Martens, 2002; Skinner et al., 1991). Additional studies across students (e.g., general education students who do not finger count, students with Attention Deficit Hyperactivity Disorder), settings (e.g., class-wide applications), and tasks (e.g., multiplication, subtraction) are needed to evaluate the external validity of the current findings.

Implications for Practitioners

Demonstration of CBM procedures. In the current study we applied some of the tools necessary to implement RTI models of math skill remediation. We demonstrated how CBM procedures could be used to identify a math skill deficit and evaluate responsiveness to intervention. Although the current study suggests that the student was more responsive to the TP intervention (i.e., it resulted in the greatest increase in skill development), students may respond differently to interventions. Therefore, practitioners should be familiar with methods for evaluating the interventions in order to ensure selected remediation procedures are having the desired effect.

Empirically validating TP. When implementing RTI models, after skill deficits are identified, remediation requires effective (e.g., empirically validated), time-efficient interventions that enhance learning rates (Skinner et al., 2002/1995). Because skill development is unlikely to occur unless students choose to respond (e.g., choose to complete math problems), intervention procedures should be acceptable to students (Skinner et al., 2005; Turco & Elliott, 1986). Because educators must implement the interventions, they should require few resources and be easily applied across settings and students. The current study, as well as limited previous research, suggests that TP is an efficient, acceptable, easy-to-implement (very brief instruction to the student), and effective procedure for increasing basic math-fact fluency in students with skill deficits.

Intervention selection and responsiveness. RTI models allow for the identification of skills in need of remediation. Applied studies provide us with some confidence that validated interventions may remedy skill deficits. However, one concern with RTI models is intervention selection. Once problems are identified, how are practitioners to select from several available, acceptable, effective, science-based (e.g., empirically validated) remediation procedures? In order to prevent students with skill development problems from falling farther behind, educators should select the procedure that results in the *most rapid* remediation.

Although limited in scope, the current study also demonstrates how researchers can more precisely evaluate remediation procedures by comparing the effects of two or more interventions while holding instructional time constant across the interventions. Such research allows for a more precise evaluation of remediation effects, which may enable practitioners to select procedures that will result in more rapid increases in skill development. When a student is behind in skill development, it is important to know what works (e.g., what has been empirically validated), but we should be striving to identify what works *best*. If all else is constant across interventions (e.g., acceptability, resources, side effects), the one that is fastest or results in the most rapid increase in learning rates is best (Skinner et al., 2002/1995).

References

Bliss, S., Skinner, C. H., & Adams, R. (in press). Enhancing sight-word reading with a time delay taped-words intervention with an English language learning fifth-grade student. *School Psychology Review*.

Carpenter, P. A., & Moser, J. M. (1982). The development of addition and subtraction problem solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 9–24). Hillsdale, NJ: Erlbaum.

Cates, G. L., & Rhymer, K. N. (2003). Examining the relationship between mathematics anxiety and mathematics performance: An instructional hierarchy perspective. *Journal of Behavioral Education*, *12*, 23–34.

Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual*. Reston, VA: Council for Exceptional Children.

Eckert, T. L., Ardoin, S. P., Daly, E. J., III, & Martens, B. K. (2002). Improving oral reading fluency: An examination of the efficacy of combining skill-based and performance-based interventions. *Journal of Applied Behavior Analysis*, *35*, 271–281.

Freeman, T. J., & McLaughlin, T. F. (1984). Effects of a taped-words treatment procedure on learning disabled students' sight-word reading. *Learning Disability Quarterly*, *7*, 49–54.

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-intervention: Definition, evidence and implications for the learning disabilities construct. *Learning Disabilities Research & Practice*, *18*, 151–171.

Gagne, R. M. (1982). Some issues in psychology of mathematics instruction. *Journal of Research in Mathematics Education*, *14*, 7–18.

Greenwood, C. R., Delquadri, J. C., & Carta, J. J. (1997). *Together we can! Class-wide peer tutoring to improve basic academic skills*. Longmont, CO: Sopris West.

Haring, N. G., & Eaton, M. E. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, and C. L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23–40). Columbus, OH: Merrill.

Hasselbring, T. S., Goin, L. I., & Bransford, J. D. (1987). Developing automaticity. *Teaching Exceptional Children*, *1*, 30–33.

Hayden, J., & McLaughlin, T. F. (2004). The effects of cover, copy, compare and flash card drill on correct rate of math facts for a middle school student with learning disabilities. *Journal of Precision Teaching and Celeration*, *20*(1), 17–21.

Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special education needs: A meta-analysis. *Remedial and Special Education*, *24*, 97–114.

Marston, D. (1989). A curriculum-based measurement approach to assessing academic performance. What it is and why do it. In Shinn, M. R. (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18–78). New York: Guilford Press.

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*, 129–147.

McCallum, E., Skinner, C. H., Turner, H., & Saecker, L. (in press). The taped-problems intervention: Increasing multiplication fact fluency using a low-tech, class-wide, time-delay intervention. *School Psychology Review*.

McLaughlin, T. F., & Skinner, C. H. (1996). Improving academic performance through self-management: Cover, copy, and compare. *Intervention in School and Clinic*, *32*, 113–118.

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

Neddenriep, C. E., Skinner, C. H., Hale, H., Oliver, R., & Winn, B. (in press). An investigation of the validity of reading comprehension rate: A direct, dynamic measure of reading comprehension. *Psychology in the Schools*.

Pellegrino, J. W., & Goldman, S. R. (1986). Information processing in elementary mathematics. *Journal of Learning Disabilities*, *20*, 23–32.

Perie, M., Grigg, W., & Dion, G. (2005). *The nation's report card: Mathematics 2005* (National Center for Education Statistics 2006-453). Washington, DC: U.S. Government Printing Office.

Rathvon, N. (1999). *Effective school interventions: Strategies for enhancing academic achievement and social competence*. New York: Guilford Press.

Shapiro, E. S. (1996). *Academic skills problems: Direct assessment and intervention* (2nd ed.). New York: Guilford Press.

Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education and Treatment of Children*, *8*(1), 67–76.

Skiba, R., Magneusson, D., Marston, D., & Erickson, K. (1986). *The assessment of mathematics performance in special education: Achievement test, proficiency tests, and formative evaluation?* Minneapolis, MN: Special Services, Minneapolis Public Schools.

Skinner, C. H. (2002). An empirical analysis of interspersal research: Evidence, implications and applications of the discrete task completion hypothesis. *Journal of School Psychology*, *40*, 347–368.

Skinner, C. H., Bamberg, H., Smith, E. S., & Powell, S. (1993). Subvocal responding to increase division fact fluency. *Remedial and Special Education*, *14*, 49–56.

Skinner, C. H., Belfiore P. J., Mace, H. W., Williams-Wilson, S., & Johns, G. A. (1997). Altering response topography to increase response efficiency and learning rates. *School Psychology Quarterly*, *12*, 54–64.

Skinner, C. H., Belfiore, P. J., & Watson, T. S. (2002/1995). Assessing the relative effects of interventions in students with mild disabilities: Assessing instructional time. *Journal of Psychoeducational Assessment*, *20*, 346–357. Reprinted with permission from Skinner, C. H., Belfiore, P. J., & Watson, T. S. (1995) *Assessment in Rehabilitation and Exceptionality*, *2*, 207–220.

Skinner, C. H., Ford, J. M., & Yunker, B. D. (1991). An analysis of instructional response requirements on the multiplication performance of behavior disordered students. *Behavioral Disorders*, *17*, 56–65.

Skinner, C. H., Neddenriep, C. E., Bradley-Klug, K. L., & Ziemann, J. M. (2002). Advances in curriculum-based measurement: Alternative rate measures for assessing reading skills in pre- and advanced readers. *Behavior Analyst Today*, *3*, 270–281.

Skinner, C. H., Pappas, D. N., & Davis, K. A. (2005). Enhancing academic engagement: Providing opportunities for responding and influencing students to choose to respond. *Psychology in the Schools*, *42*, 389–403.

Skinner, C. H., & Shapiro, E. S. (1989). A comparison of a taped-words and drill interventions on reading fluency in adolescents with behavior disorders. *Education and Treatment of Children*, *12*, 123–133.

Skinner, C. H., Shapiro, E. S., Turco, T. L., Cole, C. L., & Brown, D. K. (1992). A comparison of self- and peer-delivered immediate corrective feedback on multiplication performance. *Journal of School Psychology*, *30*, 101–116.

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

Smith, T. J., Dittmer, K. I., & Skinner, C. H. (2002). Enhancing science performance in students with learning disabilities using cover, copy, and compare: A student shows the way. *Psychology in the Schools*, *39*, 417–425.

Stoiber, K. C., & Kratochwill, T. R. (2000). Empirically supported interventions and school psychology: Rationale and methodological issues: Part I. *School Psychology Quarterly*, *15*, 75–105.

Turco, T. L., & Elliott, S. N. (1986). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology*, *24*, 277–283.

Van Houten, R., & Little, G. (1982). Increased response rate in special education children following an abrupt reduction in the absence of a token economy. *Education and Treatment of Children*, *5*, 23–32.

Van Houten, R., & Thompson, C. (1976). The effects of explicit timing on math performance. *Journal of Applied Behavior Analysis*, *9*, 227–230.

Appendix A. Addition Problem Sets

List A	List B	List C
3 + 10 =	4 + 5 =	2 + 5 =
1 + 8 =	1 + 2 =	4 + 10 =
5 + 7 =	6 + 7 =	8 + 9 =
6 + 10 =	2 + 2 =	1 + 10 =
2 + 7 =	7 + 9 =	6 + 6 =
3 + 4 =	5 + 10 =	5 + 9 =
1 + 1 =	2 + 10 =	3 + 6 =
2 + 3 =	3 + 5 =	1 + 5 =
3 + 8 =	1 + 4 =	3 + 9 =
4 + 9 =	6 + 9 =	9 + 9 =
5 + 6 =	4 + 7 =	4 + 8 =
9 + 10 =	5 + 5 =	1 + 7 =
6 + 8 =	8 + 8 =	4 + 6 =
2 + 6 =	2 + 9 =	5 + 8 =
7 + 7 =	1 + 9 =	2 + 8 =
1 + 6 =	3 + 7 =	7 + 8 =
7 + 10 =	8 + 10 =	3 + 3 =
4 + 4 =	2 + 4 =	10 + 10 =
1 + 3 =		

Appendix B. Conducting the Taped Problems Intervention

1. Construct an audiotape by reading the number of each problem, the problem, and the answer. Incorporate the desired time delay between reading the problem and answer.

   Example: *Number one* ⟶ *two plus four equals* ⟶ *six*
   (3 second pause) (time delay)

2. Read the entire list of problems this way and then repeat the problems until the list has been repeated the desired number of times.

3. Provide the student with a sheet of problems that corresponds to the audiotape.

4. Provide the following instructions:

   - Follow along with the tape.

   - Try to beat the tape by writing down your response after the problem is read but before you hear the answer.

   - Check your response when you hear the answer on the tape. If it is incorrect, draw a line through it and write the correct answer.

   - If you don't write down a response on time (before the answer is read), write in the correct answer given by the tape and wait for the next problem.

Students sees	Student hears	Delay, seconds	Student responds
1. 5 + 2 = ___	Number one. Five plus two equals	None 3 2 1	Correctly or incorrectly

Appendix C. Conducting the Cover, Copy, Compare Intervention

Instructions to the student:

1. Read the problem and answer.
2. Cover the problem and answer.
3. Write down the problem and answer.
4. Check the problem against the original.
5. If your response is correct, move on to the next problem.
6. If your response is incorrect, copy the problem and answer in the three spaces provided.

Sample Cover, Copy, Compare Worksheet

1. 2 + 5 = 7	_______	_______	_______
2. 4 + 10 = 14	_______	_______	_______
3. 8 + 9 = 17	_______	_______	_______
4. 1 + 10 = 11	_______	_______	_______
5. 6 + 6 = 12	_______	_______	_______

Appendix D. Procedural Integrity Checklist

1. Give the student the appropriate folder.
2. Ask the student to sit at the desk with the tape player.
3. Make sure the student's packet corresponds to the tape before placing the tape in the tape player.
4. Instruct the student to follow along with the tape and not work ahead.
5. When the student is ready, start the tape.
6. When the first intervention ends, collect the follow-along packet.
7. Place the appropriate assessment probe face down in front of the student.
8. Start your stopwatch (or wait until the second hand reaches 12 if using a clock).
9. Instruct the student to turn over the paper and begin working.
10. After 1 minute, say, "Time's up."
11. Collect the probe.
12. Thank the student and escort the student to the next class.