

An Experimental Analysis of the Effects of Reading Interventions in a Small Group Reading Instruction Context

Christine M. Bonfiglio · Edward J. Daly III ·
Michael Persampieri · Melissa Andersen

Published online: 15 June 2006
© Springer Science+Business Media, Inc. 2006

Abstract The current study examined the effects of several combinations of instructional and motivational interventions on oral reading fluency in the context of small group reading instruction. A treatment package consisting of acquisition, fluency, and motivational components was examined and then dismantled. Results were evaluated individually for each participant using a multiple-probe design in order to identify the most effective, yet efficient reading intervention package for each member of the group. The results indicated that, although all potential combinations of the package were effective, the experimental analysis was able to identify an effective yet more efficient package that produced substantial increases in performance across all students. Results are discussed in terms of effective instructional components in small group instruction and how to adapt methods for the experimental analysis of behavior to make decisions about the effectiveness of small group instruction in reading.

Keywords Academic interventions · Effective reading treatments · Experimental analysis · Oral reading fluency

In the year 2000, 37% of fourth graders read at or below the “Basic” level (National Center on Educational Statistics, 2001). Furthermore, the National Center on Educational Statistics (2001) indicates that the reading scores of the lowest 25th percentile of fourth graders have steadily declined. To help remediate these problems, educators must have substantial

This study is based on a dissertation submitted by the first author in partial fulfillment of requirements for the doctoral degree.

C. M. Bonfiglio (✉)
Elkhart Community Schools,
3317 Greenleaf Blvd., Elkhart, IN 46514
e-mail: cmbonfig@aol.com

E. J. Daly III · M. Persampieri · M. Andersen
University of Nebraska-Lincoln,
NE

sophistication with intervention design, especially in the use of assessment information for selecting effective interventions.

The field of applied behavior analysis brings a unique perspective and set of methods in helping to resolve academic performance problems. While most educational settings generally group students for instruction due to limited resources, students progress, or fail to progress, one at a time. Because growth in academic skills is an individual phenomenon and decisions about student performance must be made at the individual level, the functional assessment of behavior offers a conceptual and methodological approach that can greatly assist educators in making good instructional decisions for students.

Experimental analysis allows researchers to establish functional relationships between variables (Lerman & Iwata, 1993; Taylor & Romanczyk, 1994). To date, experimental analysis literature has focused largely on behavioral excesses (Ervin et al., 2001). The variability in behavior that is associated with behavioral excesses may be appealing to researchers because it is more readily amenable to manipulations via experimental analyses with single-case research designs. The lack of variability in behavior that is associated with academic skills (e.g., an inability to read) may be more difficult to analyze with single-case designs.

In spite of the challenges associated with studying academic behavior, an increasing number of studies have applied principles of experimental analysis to problematic academic behaviors (e.g., Daly, Martens, Dool, & Hintz, 1998; Daly, Martens, Hamler, Dool, & Eckert, 1999; Duhon et al., 2004; Eckert, Ardoin, Daly, & Martens, 2002; Hendrickson, Gable, Novak, & Peck, 1996; Jones & Wickstrom, 2002; McComas et al., 1996; Noell et al., 1998; VanAuken, Chafouleas, Bradley, & Martens, 2002). These studies have used abridged single-case design elements to identify key instructional components that accelerate academic responding. For example, Lentz (1988) made a distinction between skill-based versus performance-based deficits. A student is said to have a skill deficit when he or she does not possess adequate skills to be successful with the current instructional task. A performance deficit manifests itself when the student has the skills to perform the task but the contingencies fail to support occurrence of the behavior. The distinction of skill versus performance deficits has proven useful in a number of studies that have discriminated students' performance-based and skill-based instructional needs (Duhon et al., 2004; Eckert, Ardoin, Daisey, & Scarola, 2000; Eckert et al., 2002; Noell et al., 1998).

The Instructional Hierarchy (IH; Haring & Eaton, 1978) is another conceptual framework which has allowed investigators to refine the conceptualization of skill deficits and which has been used in several studies to identify combinations of effective treatment components through experimental analysis (Daly et al., 1998; Daly et al., 1999; VanAuken et al., 2002). Haring and Eaton (1978) describe learning as consisting of four stages: (a) acquisition, (b) fluency, (c) generalization, and (d) adaption (i.e., the modification of the learned skill in the face of novel environmental demands). That is, when a new skill is being taught, the learner must first acquire it and then become fluent in the use of this skill. If the learner is fluent and instruction is appropriate, he or she is more likely to generalize its use to novel contexts and finally adapt the use of the skill to modify the response in order to accommodate its use to novel demands. This framework encompasses different levels of skill development in a learner's progression toward skill mastery and each level has corresponding instructional procedures that efficiently improve student performance. Thus, the use of the IH provides researchers and practitioners with a conceptual framework to identify functional variables to remediate skill deficits.

To date, experimental analyses of academic performance based on these conceptualizations of academic responding have been conducted with individual students in sessions in which interventions were individually delivered. Although learning is an individual

phenomenon, academic interventions can seldom be delivered individually in schools. The natural context for most reading instruction is the small group instructional format (Foorman & Torgesen, 2001). Therefore, a logical next step in the progression of research in this area is to examine the use of experimental analysis under different instructional delivery formats like those most commonly used in classrooms.

The purpose of this study was to extend experimental analysis methods by examining the validity of a method for providing instructional trials through small group sessions. A treatment package consisting of empirically validated reading fluency instructional methods and utilizing instructional and motivational variables consistent with prior conceptualizations of academic responding was delivered within a multiple-probe design. The treatment package was dismantled until the most efficient, yet effective package was identified. The goal was to determine the simplest treatment package (i.e., the one with the fewest number of treatment components), albeit a package that would produce increased oral reading fluency for all students. An experimenter conducted instructional sessions until the effective package was determined. A special education teacher then conducted instructional sessions to determine whether the same results could be obtained and to provide a stronger basis for evaluating the social validity of the empirically identified treatment package.

Method

Participants and setting

One reading group, consisting of four 4th grade students from the same elementary school classroom, served as participants in this study. Three of the students were male (Blake, Cody, and Devon). Karla was the lone female participant. The ethnic background of the group was diverse, consisting of two Caucasian students (Blake and Cody), one African-American student (Devon), and one Hispanic student (Karla). The students were identified as poor readers by their elementary teacher; however, none were receiving special education services.

The experimental sessions were carried out in a classroom as a part of small reading group instruction. An experimenter implemented the reading group four days per week. Students were assessed individually four days a week. Assessments were conducted at a small table in the school psychologist's office.

Materials

Instructional reading passages

Reading passages of narrative and expository texts were obtained from the Houghton Mifflin Reading Series. Eight passages were identified and assigned to a specific week in a random fashion (six initially and two additional passages halfway through the study, as it appeared that more would be needed to complete the experimental analysis). These passages were used for instructing small group reading. The Spache readability formula (Spache, 1953) was used to identify the difficulty level of passages. The average readability of the passages was 4.38 ($SD = 0.23$; range, 4.0 to 4.7).

Experimental design and dependent variables

A multiple-probe design across tasks (reading passages) was used to examine changes in correctly read words and errors per min (Wolery, Bailey, & Sugai, 1988).

Oral reading fluency

Correctly read words (CRW) and errors per min were used to assess reading fluency in the instructional passages. A CRW was defined as a word that was pronounced correctly within 3-s. Errors included hesitations of 3-s or more, omissions, substitutions, transpositions, and mispronunciations. The experimenter scored CRW and errors while the student read the passage for 1 min. All sessions were taped using an audiocassette recorder to assess interscorer agreement (IOA).

Student and teacher behaviors

Student and teacher behaviors were recorded by observers using a 10-s partial-interval sampling scheme during instructional sessions. Specifically, observers measured student academic engagement and teacher instruction. Student academic engagement was recorded for a 10-s interval if the student was observed producing a verbal or written response, attending to instruction (i.e., looking in the direction of the teacher or textbook), and/or reading aloud or following along while others were reading. Teacher instruction was recorded if the experimenter or teacher was observed to be asking questions, telling/explaining, modeling correct responding, correcting an incorrect verbal or written response, and/or describing contingencies for completion. Each student's behavior was recorded every 4th interval. Participants were randomly assigned to each of the four seats on a daily basis in an attempt to control for patterns that might occur through behavioral cycles or seating arrangements.

Independent variables and treatment conditions

A variety of intervention components was used throughout the experimental analysis. Strategies included passage previewing and modeling, practice, error correction, praise for correct reading, and contingent reward. The full treatment package included all of these components. Individual components were selectively withdrawn through the course of the study, which is described in the Procedures section. What follows is merely a description of the individual components that were used at one point or another in the study.

Taped preview (TP)

For this condition, the passage was pre-recorded on an audiocassette player by the experimenter. At the beginning of reading group instruction, the entire group and the experimenter first listened to the story on tape while following along in the text.

Choral reading (CR)

During CR, the reading group read the passage aloud in unison with the experimenter, which provided opportunities to respond.

Error correction (EC)

A Word Drill error correction technique was used when a student read a word incorrectly. Upon mispronunciation, the experimenter stopped the student from reading, read the word

aloud correctly, and required all four students to read the error word correctly three times. This procedure provided a model for correct responding and three additional opportunities to respond for all group members.

Reward

Reward (R) involved the presentation of preferred stimuli contingent upon meeting an oral reading fluency criterion. The experimenter first asked the student if he or she wished to attempt to exceed the goal. If so, the experimenter told the student that if he or she read the passage at a rate that matched or exceeded the previous number of CRW while matching or decreasing the number of errors, he or she could choose an item from the “goodie bag,” which included small tangible items like pencils, erasers, hair clips, and small plastic toys. Goals were developed on an individual basis in which each participant’s best score on the specific passage being used served as the goal (i.e., the highest number of CRW and the lowest number of errors for the passage). Goals were revised daily throughout the study when R was a part of the intervention package.

Procedures

Direct observations

Graduate students in a school psychology training program were trained in direct observation using the format created for this study. Trainees practiced observation using the observation code with videotapes first and then in a classroom environment during academic instruction. Interobserver agreement was computed for all practice sessions. Prior to participation in the study, observers were expected to demonstrate interobserver reliability of at least 80%.

Direct observations were conducted daily during small reading groups. At least one observer conducted observations for each session. Approximately every third session, two observers conducted observations concurrently to compare results for the purpose of obtaining interobserver agreement. Observers were positioned off to the side of the reading group so that they could see both the teacher and the four students. Sessions were between 6 minutes and 20 minutes in length, depending on the condition in effect. An audiotaped signal indicated the correct observational interval.

Measurement of oral reading fluency

On a daily basis, students were individually assessed using instructional passages. CRW and errors were recorded as the student read the passage for 1 min. Assessment occurred in the story taught the previous day and in at least two other instructional passages, some of which were still in baseline and some of which were in the maintenance phase. In this way, each assessment session sampled performance for the passage being instructed for the week and other passages to meet the requirements of a multiple-probe design. Baseline and maintenance passages were sampled without replacement until all passages were assessed at least once. Then the process was repeated to assure ongoing measurement across all passages and phases.

Experimental sessions

Following the establishment of baselines in the instructional passages, treatment conditions were carried out. Instruction with the complete treatment package (i.e., TPP, CR, EC, R) was carried out in the first and second passages, respectively. The treatment package was administered for two consecutive weeks (with two separate passages) to determine whether the treatment effects were consistent across passages. As new passages were introduced, intervention components were sequentially withdrawn according to whether they were acquisition components, fluency components, or reward components. The acquisition components were first withdrawn, leaving fluency and reinforcement components in the third week of treatment. The fluency and reinforcement components were then withdrawn in the same fashion during weeks four and five of treatment, respectively. During all instructional sessions, the experimenter encouraged the students by using praise.

Visual inspection of results was used to identify the treatment combination that was associated with the highest levels of responding. If fluency decreased when a component was withdrawn, it was placed back in the package during subsequent instructional sessions and another component was removed. This process continued until the most efficient and effective treatment package was identified.

In the second to last passage (i.e., week 7), the identified efficient, effective treatment package was administered. During this week, the special education teacher observed the experimenter lead the reading group while following along on a scripted protocol. The following week (i.e., the final week of the study), the teacher implemented the treatment package with the reading group as the experimenter observed. The protocol was utilized throughout the reading group so that the teacher could rely on it during implementation.

Teacher training

Once the most effective package was identified for the group as a whole, the teacher was trained to implement the package for the following week. In the training session, the experimenter used explanation, modeling, practice, and feedback with the teacher to assure that she would be able to do the intervention. A scripted protocol detailing the procedures in a step-by-step fashion was provided to the teacher. Furthermore, as noted earlier, the teacher had observed the experimenter implement the package during the reading group for two days prior to her implementation.

Interscorer and interobserver agreement

An independent observer listened to the audiotape recorded sessions and scored the passages for CRW and errors. To compute interscorer agreement, the total number of agreements for CRW and errors was divided by the total number of words in the passage, which represents all possible agreements plus disagreements. A total of 34.4% of all sessions (i.e., 11 of the 32 sessions) was assessed for interscorer agreement. The mean agreement was 99.2% (range, 95.1–100%) across all participants.

To compute interobserver agreement, the total number of agreements for occurrences was divided by the total number of agreements plus disagreements. Agreement was assessed on an interval-by-interval basis. A total of 34.4% of the direct observation sessions (11 of the 32 sessions) was assessed for interobserver agreement. The mean agreement was 91.5% (range, 79–100%).

Treatment integrity

Experimenter integrity

Observers were provided with a protocol that outlined the full treatment package (i.e., TP, CR, EC, and R). While observing the reading group, observers recorded whether each step was completed in the specified order. The total number of steps completed was divided by the total number of steps in the condition to yield the percentage of steps completed for each session. Treatment integrity was assessed for 34.4% of all sessions (i.e., 11 of the 32 sessions). The mean percentage of correctly implemented steps was 99.3% (range, 88–100%).

Teacher integrity

Treatment integrity was assessed during teacher led instructional sessions. Seventy-five percent of the week's sessions (i.e., 3 of 4) were observed by two independent observers. The percentage of correctly implemented steps was 100% for all sessions.

Treatment acceptability

The teacher completed the Intervention Rating Profile-15 (IRP; Martens, Witt, Elliott, & Darveaux, 1985) by responding to the 15-item questionnaire. The IRP-15 has a 6-point Likert Scale, with 1 indicating “strongly disagree” and 6 indicating “strongly agree.” Also, she was asked to offer any other feedback on the intervention procedures.

Results

Results for oral reading are displayed in Figs. 1 through 4 and in Table 1. Although there were initial increases in performance for some of the baselines (e.g., Blake, passages 4 & 5; Corey, passage 7), the baselines were generally stable before the introduction of intervention conditions for any given passage. In general, all participants increased CRW and decreased errors per min in almost all treatment conditions (relative to baselines; exceptions are noted below). Effect sizes (using the no assumptions method; Busk & Serlin, 1992) were calculated as an adjunct to visual inspection. Effect sizes are large and range from 1.37 to 6.99 across all treatments for all participants. Visual inspection of maintenance data suggests that participants maintained fluency rates (i.e., CRW per min) after instruction was withdrawn from a passage. Therefore, results support the efficacy of all of the combinations of treatment components.

The full treatment package (TP + EC + CR + R) was delivered in the first two passages. Acquisition components (TP + EC) were then withdrawn in the third passage. Acquisition components were subsequently added back to the treatment while the practice component (CR) was withdrawn in the fourth passage. Finally, CR was added back to the treatment while the motivational variable (R) was withdrawn in the fifth passage. For reasons elaborated below, this last condition (TP + EC + CR) was deemed the most parsimonious, effective treatment and was implemented a second time by the experimenter (passage 6) and then by the teacher (passage 7). Hence, it is referred to as the efficient treatment in the individual analysis of results that follows.

Table 1 Average CRW per min, errors per min, and effect sizes for oral reading

Condition	Blake		Cody		Devon		Karla	
	CRW	Errors	CRW	Errors	CRW	Errors	CRW	Errors
TP + EC + CR + R ₁ ^a								
Baseline	108	1.75	91	4.25	98.75	8.75	98.25	6.50
Treatment	173	1.50	112.25	1.50	161.75	4.25	170	3.50
Effect Size	3.19		2.21		4.08		2.95	
TP + EC + CR + R ₂								
Baseline	93.60	4.80	77	4.25	89.40	7.40	74.20	10.60
Treatment	157.50	1.25	111.25	.25	138	1.50	150	7
Effect Size	3.62		2.35		3.38		3.56	
CR + R ₃								
Baseline	114.33	5	80.80	4	74	6.80	85.40	7.80
Treatment	162.75	2.25	105.75	1.50	122.25	6	123.75	5.50
Effect Size	2.67		2.22		6.99		6.69	
TP + EC + R ₄								
Baseline	109.90	5	89.43	4	101.30	5.10	91.70	13.10
Treatment	164.75	1.25	141.25	0.75	156.75	2.25	142.00	5.25
Effect Size	3.34		4.66		5.08		2.82	
TP + EC + CR ₅								
Baseline	118.75	2.10	100	3.25	101.80	4.80	104.40	8.10
Treatment	164.75	1.25	128	2	146.50	1.75	144.50	4
Effect Size	2.51		1.37		4.68		4.10	
TP + EC + CR ₆								
Baseline	120.50	4	91	1	90.50	1.20	104.60	4.40
Treatment	180.25	2.25	113	0.33	159.75	2	197.75	5
Effect Size	5.33		1.37		4.68		4.10	
TP + EC + CR ₇ (Teacher)								
Baseline	89.90	4.75	84	1.40	80.30	5.30	81.30	8.40
Treatment	134.75	1.25	127.50	0.25	132	2.50	126.80	5
Effect Size	5.2		2.42		2.43		4.28	

^aThe subscript refers to the passage number (1 through 7).

For Blake (Fig. 1), increasing trends in performance during instruction are most evident with TP + EC + CR + R (the full treatment package in passage 1), TP + EC + R (passage 4), and TP + EC + CR (the efficient treatment package administered in passages 6 & 7). Blake obtained the highest absolute increases for CRW per min from baseline with the full treatment package (65 and 63.9 CRW per min in passages 1 and 2, respectively). The next highest treatment effect was produced by TP + EC + CR in passage 6 (increase of 59.75 CRW per min), the efficient package. The highest absolute fluency level achieved during instruction was with the full treatment package (208 CRW per min in passage 1), followed by TP + EC + CR in passage 6 (206 CRW per min), which was followed by TP + EC + R (201 CRW per min). The largest effect size was obtained with the second administration of TP + EC + CR (passage 6; 5.33), followed by the teacher administration of the same package (5.2 in passage 7). Error rates were reduced substantially, indicating that accuracy also improved as a function of treatment regardless of the condition.

For Cody (Fig. 2), increasing trends in performance during instruction are evident in all passages. Cody obtained the highest absolute increases for CRW per min from baseline with TP + EC + R (51.85 CRW per min in passage 4). The next highest treatment effect was

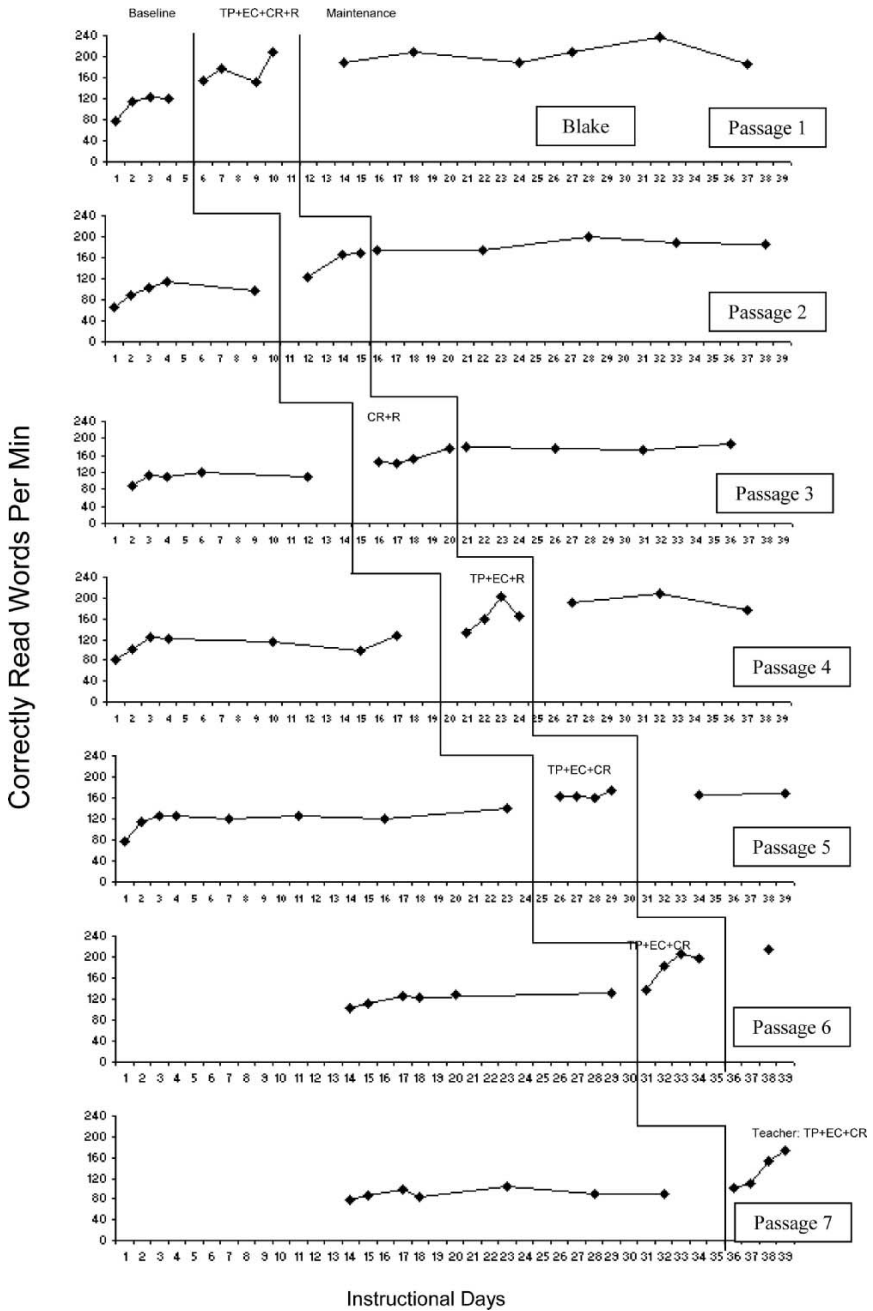


Fig. 1 Measure of oral reading fluency for Blake: CRW/min

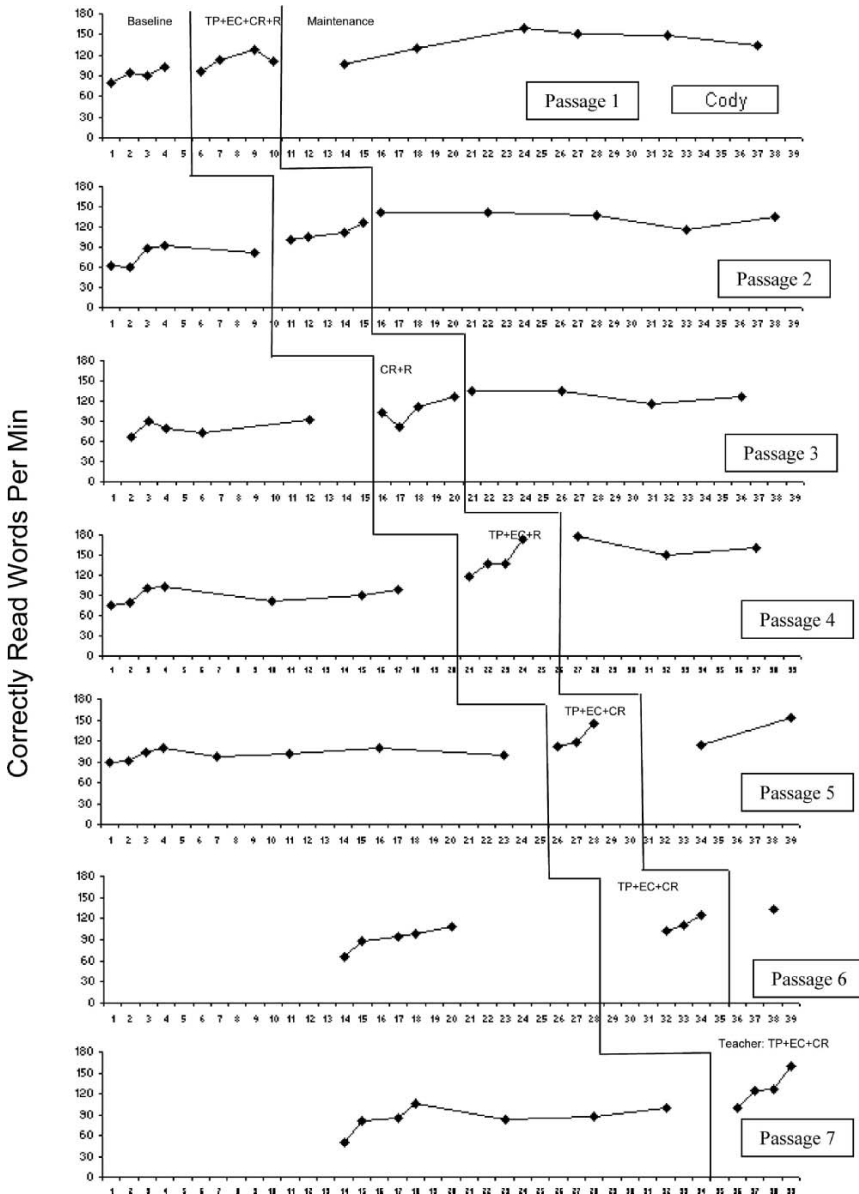


Fig. 2 Measure of oral reading fluency for Cody: CRW/min

produced by TP + EC + CR in passage 7 (teacher delivered instruction; increase of 43.5 CRW per min). The highest absolute fluency level achieved during instruction was with TP + EC + R (174 CRW per min in passage 4), followed by TP + EC + CR in passage 7 (teacher delivered instruction; 160 CRW per min), which was followed by TP + EC + CR in passage 6 (145 CRW per min). The largest effect size was obtained with TP + EC + R in passage 4 (4.66), followed by TP + EC + CR in passage 7 (teacher delivered instruction; 2.42). Error rates were reduced by a third or more across all treatment conditions.

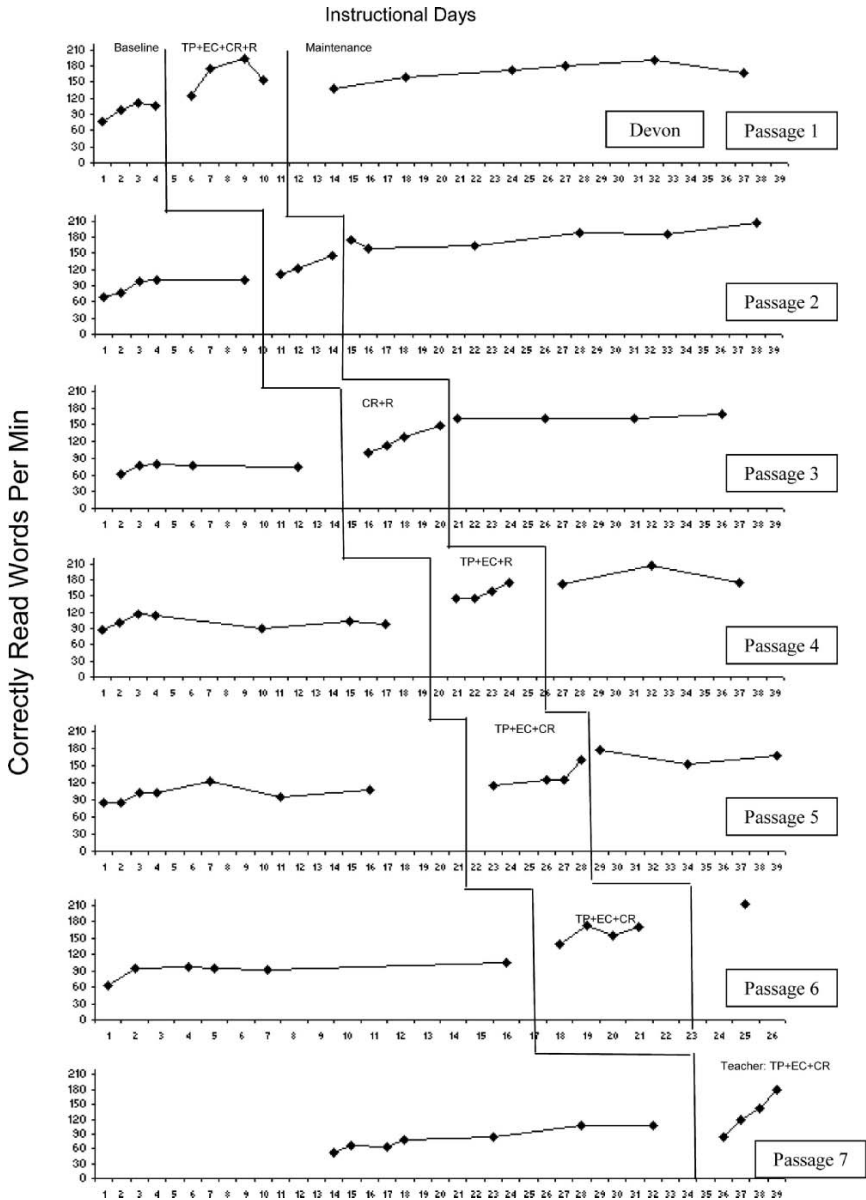


Fig. 3 Measure of oral reading fluency for Devon: CRW/min

For Devon (Fig. 3), increasing trends in performance during instruction are evident in all passages. Devon obtained the highest absolute increases for CRW per min from baseline with TP + EC + CR in passage 6 (69.25 CRW per min in passage 6). The next highest treatment effect was produced by TP + EC + CR + R (the full package) in passage 1 (increase of 63 CRW per min). The highest absolute fluency level achieved during instruction was with TP + EC + CR + R (193 CRW per min in passage 1), followed by TP + EC + CR

(teacher delivered instruction in passage 7; 180 CRW per min), which was followed by TP + EC + R (176 CRW per min). The largest effect size was obtained with CR + R in passage 3 (6.99), which was followed by TP + EC + R in passage 4 (5.08). Error rates were reduced substantially in all conditions except CR + R (passage 3) and TP + EC + CR (passage 6). With respect to the first condition, it is likely that Devon was trying to read as fast as he could and not attending to errors during assessments. Because there had been no prior modeling or error correction (i.e., TP + EC had not been delivered), his accuracy did not improve. Therefore, his faster rate (evident in the large effect size) came at a price of compromising accuracy. For the TP + EC + CR condition in passage 6, his error rate was already low (average of 1.20 during baseline). Although a slightly higher average error rate was obtained during treatment (2 errors per min), errors remained essentially unchanged.

For Karla (Fig. 4), increasing trends in performance during instruction are most evident with both TP + EC + CR + R conditions (passages 1 & 2), TP + EC + R (passage 4), and TP + EC + CR (passages 6 & 7). Karla obtained the highest absolute increases for CRW per min from baseline with TP + EC + CR (the efficient package; 93.15 CRW per min in passage 6). The next highest treatment effect was produced by TP + EC + CR + R (the full package) in passages 2 and 1 (increases of 75.8 and 71.75 CRW per min, respectively). The highest absolute fluency level achieved during instruction was with TP + EC + CR (222 CRW per min in passage 6), followed by TP + EC + CR + R in passage 1 (194 CRW per min), which was followed by the same treatment in passage 2 (176 CRW per min). The largest effect size was obtained with CR + R in passage 3 (6.69), followed by TP + EC + CR (teacher delivered instruction in passage 7; 4.28). Error rates were relatively high during most of the baselines. Errors were reduced during all treatment conditions, but still remained relatively high.

Overall, some variation in participants' relative increases in CRW per min from baseline to treatment and absolute levels of responding was observed across participants, perhaps partially as a function of baseline levels of performance. There was also some variation in responsiveness to the different treatment packages. When examined in a variety of ways (i.e., visual inspection, relative increases, absolute levels of responding achieved, effect sizes), the general pattern favored the full treatment package across participants (TP + EC + CR + R) which was followed by the efficient package (TP + EC + CR). This conclusion is not an absolute, however, for all the participants. For instance, in some cases (e.g., Blake and Cody), TP + EC + R produced strong effects as well. Results for maintenance were relatively uniform across participants, with participants maintaining and sometimes exceeding levels achieved during instruction following the withdrawal of instruction.

Direct observational data indicated that the greatest average levels of student engagement during instructional sessions was for the teacher delivered instructional trials (TP + EC + CR in passage 7; see Table 2). It is also in this condition that the least amount of variability was found. The next highest levels of engagement were for the full treatment package (TP + EC + CR + R) and the efficient treatment package (TP + EC + CR). At a minimum, therefore, the combination of TP + EC + CR seems optimal for promoting levels of student engagement. The contingency for performance (R) did not appear to affect performance in any appreciable way. Of course, the contingency was delivered during a later assessment session. The condition that was associated with the highest percentage of time devoted to teacher instruction was TP + EC + CR, with the classroom teacher attaining slightly higher levels of intervals of teacher instruction (and less variability) than the experimenter. On the IRP-15, the mean rating across all items was 4.86. The teacher did express concern about how a teacher would implement this package with an entire class. Thus, a lower score on

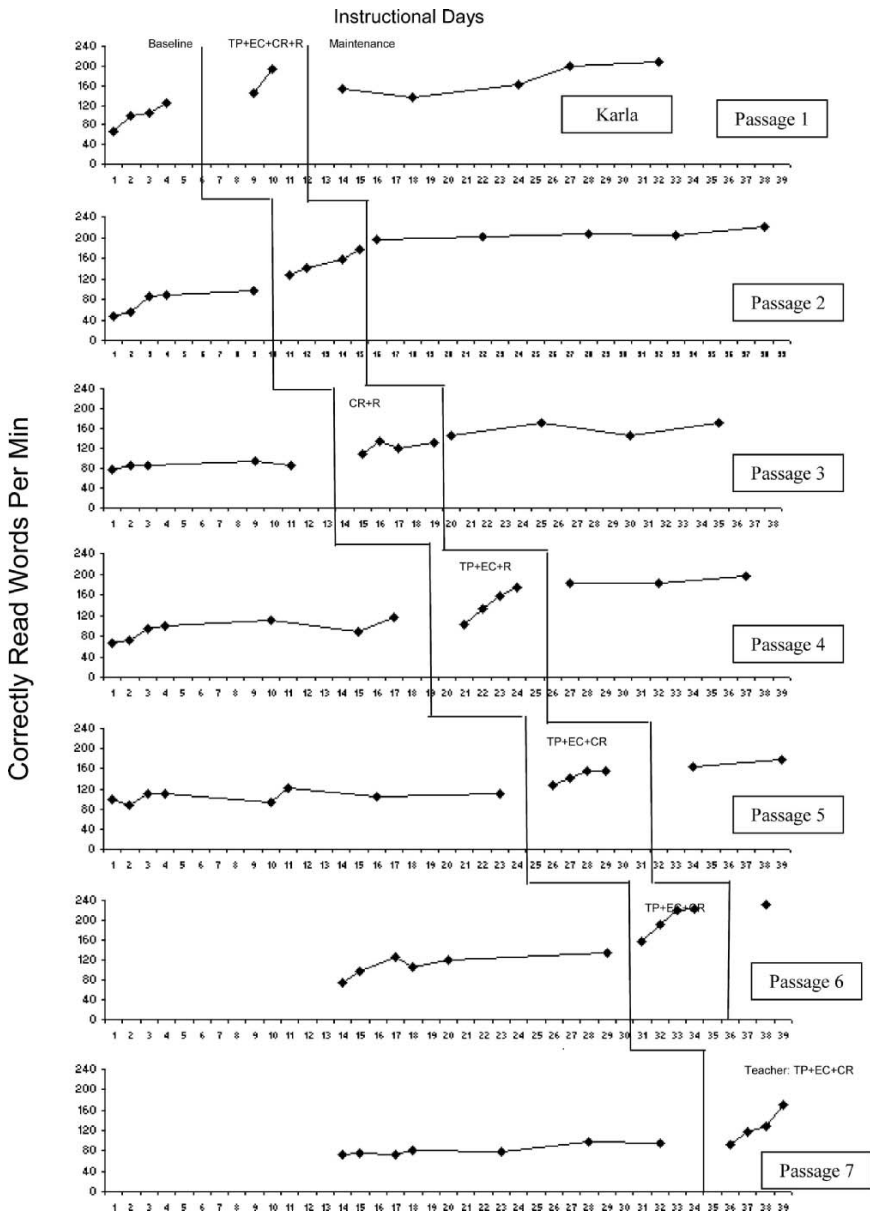


Fig. 4 Measure of oral reading fluency for Karla: CRW/min

“Would you be willing to implement this in the classroom setting” was obtained (i.e., “4”). It was explained that the procedures were designed for small groups and she then agreed that this intervention could and would be productive if implemented with a small reading group. The teacher further added verbally that the intervention procedures were valuable and that she would begin implementing them in the classroom with two small reading groups.

Table 2 Observational data

Condition	Student academic engagement			Teacher instruction		
	Mean	SD	Range	Mean	SD	Range
TP + EC + CR + R ₁₊₂ ^a	86.75	8.72	84–97	70	4.66	62–74
CR + R ₃	78.25	6.23	65–85	57.5	6.61	49–65
TP + EC + R ₄	82.5	6.61	75–91	64.5	9.16	53–75
TP + EC + CR ₅₊₆	83	13.22	64–97	79.88	6.06	69–87
TP + EC + CR ₇ ^b	93.25	3.77	90–97	80.75	2.5	78–84

^aThe subscript refers to the passage number (1 through 7).

^bTeacher delivered instructional sessions.

Discussion

In this study, treatment packages were analyzed and an “effective package” was identified utilizing a dismantling procedure (Barnett, Daly, Jones, & Lentz, 2004) in an attempt to create an equally effective package that was more efficient (i.e., easier to do). Through an experimental analysis utilizing reading intervention components that (a) have been empirically validated [e.g., TPP, Rose & Beattie, 1986; CR, Eckert et al., 2002; WD, Rosenberg, 1986; Cossairt, Hall, & Hopkins, 1973; and CTR, Billingsley, 1977], (b) correspond to the conceptual framework of academic responding within a learning hierarchy [i.e., the Instructional Hierarchy (Haring & Eaton, 1978)] and (c) address the issues of skill-based versus performance-based deficits (Lentz, 1988), an effective package was identified. The results of this study indicate that all treatments were effective at increasing responding for all four participants, with the identified “effective package” being most successful in increasing responding for three of the four participants.

In most cases, students nearly doubled their reading fluency rates by the end of the study. Immediate effects were observed in most instances and there was clear evidence that students maintained these effects once treatment was withdrawn. That is, CRW/min did not decrease substantially and/or ER/min did not increase substantially with the withdrawal of treatment. Active student engagement increased from typical instruction to the effective package by 6.25%. This finding is consistent with increases in CRW/min, as both are overlapping response classes. That is, if a student is reading aloud, (s)he is actively engaged. Total percentage of teacher time spent in interaction with students during small group instruction actually decreased by 5.12% with the effective package. Therefore, performance and academic engagement increased, while teacher effort decreased; thus, identifying an effective, yet more efficient reading intervention package.

Several elements of the method may have contributed to the favorable rating of social validity. Although a formal survey of acceptability was administered (i.e., the IRP-15) following the study, the teacher had the luxury of implementing the procedures herself prior to evaluating them. Therefore, verbal consultation occurred throughout the study with the teacher, teacher training was delivered, and the teacher experienced the intervention package firsthand. The acceptability ratings were positive, as she was pleased with student outcomes and with the efficiency of the small group reading instruction. Since that time, she has implemented the procedures within her classroom for two small reading groups, providing even stronger evidence of the social validity of the intervention (Gresham & Lopez, 1996).

These data are encouraging, as they speak to the efficacy of the procedures and the effectiveness of the treatment package selected in promoting maintenance following instruction.

They lend further empirical support to the robustness of the antecedent and consequent strategies employed, and furthermore, increase our understanding of the likelihood of achieving maintenance when instruction occurs repeatedly, as is typically the case during small group reading instruction. As such, the results replicate and extend a similar finding of Bonfiglio, Daly, Martens, Lin, and Corsaut (2004) that used similar procedures with individualized interventions.

Although the “efficient package” (i.e., TPP + EC + CR) was only the most effective treatment for three of four participants (Blake, Devon, and Karla), it was selected as the treatment for the group. Cody obtained slightly higher increases in CRW/min in another treatment condition; however, one goal of the study was to determine the most effective, yet efficient package for the small reading group as a whole. Therefore, the combination of TPP + EC + CR was identified and implemented as the most efficient and effective package for the final phase of the study. The decision appeared to be the correct decision, as Cody’s CRW’s/min increased by 43.5 words. When the teacher implemented this package, it became the most effective for Cody as well. Based on these data, it appears reasonable to assume that the correct “effective package” was selected and implemented by the classroom teacher.

It might be that different results would have been obtained with students who were reading at different fluency levels (e.g., frustrational level at the fourth grade or instructional at a lower grade level, first or second). The advantage of the method employed in this study is that although the treatment package selected for these students in this small reading group was effective, other combinations of variables might be more effective for other groups just as a function of prior baseline and grade levels. Thus, a package can be custom-made to meet the needs of different students with varying characteristics using a method for experimental analysis.

There are several limitations to the study that should lead the reader to exercise caution in interpreting the results. First, participants were frequently and repeatedly instructed and probed on the same passages, which increased opportunities to respond and might have augmented treatment effects above and beyond what would have been obtained if such repeated probing not been done. Second, the passages utilized in this study were shorter in length, averaging 212 words (range = 162–237), than an average story obtained from a curricular reading book, again producing increased opportunities to respond. Third, because there were only four students in the small group for reading instruction, these students might have had higher rates of opportunities to respond than occur in a classroom where small groups are large. As the opportunities to respond increase, the rate of oral reading fluency also increases (Eckert et al., 2002; Levy, Nicholls, & Kohen, 1993; Skinner, Ford, & Yunker, 1991). In a typical classroom environment, the number of opportunities to respond may be limited based on the number of students in a reading group. Fourth, efficiency was presumed based only on the smallest number of strategies used and not on a direct measure of efficiency (e.g., measuring teacher effort or time).

Fifth, the multiple probe design does not control well for sequence effects and therefore results may suffer from the cumulative effects of treatments over time. In the future, investigators may confirm results with other design elements (e.g., multielement design) before proceeding to final validation. Sixth, the social validity rating was based on only one respondent. Other teachers may not have rated the intervention as favorably. Finally, the methods employed in this study may not yet be readily adapted to the school setting. The analysis was time consuming and employed a plethora of data collection procedures, not easily amenable to practice. However, this study was an initial effort in the area of analysis of academic responding within a small reading group context and provides valuable data that can be applied to future investigations. The replication of these results is vital to continued evaluation of the

procedures, as well as combinations of other variables and modified designs that may enable more efficient application to applied settings. Furthermore, analysis implemented initially by the teacher may alleviate the time restrictions consultants typically face.

In spite of the above limitations, the results of this study should encourage future study of experimental analyses of academic performance. The positive treatment effects and the appropriation of elements of the natural environment (i.e., small reading group instruction) suggest their potential utility in the classroom. The use of a single-case experimental design allowed the investigators to detect changes in performance for each child within the small group. Moreover, this study approximated the natural environment, yielding greater external or ecological validity than previous research conducted (e.g., Daly et al., 1998; Daly et al., 1999).

There are several implications for educators as a result of this study. First, because of the nature of this study (i.e., the identification of an effective, efficient reading intervention package for small groups), implementation to the classroom setting is viable. Within a small reading group context, it is valuable to identify a package that may not be the most effective package for every child, but for most, in that positive effects may be obtained for every child (i.e., an increase in oral reading fluency rates). Second, the utilization of the Instructional Hierarchy and grouping intervention components accordingly for experimental analysis can prove to be fruitful. The dismantling procedure can be utilized in an experimental analysis, making the process more efficient with perhaps the same results. Finally, directly applying the treatment to the target behaviors (i.e., oral reading fluency) in a classroom environment will be beneficial to all students whether difficulties with reading are present in all students.

References

- Barnett, D. W., Daly, E. J., III, Jones, K. M., & Lentz, F. E., Jr. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *Journal of Special Education, 38*, 66–79.
- Billingsley, F. F. (1977). The effects of self- and externally-imposed schedules of reinforcement on oral reading performance. *Journal of Learning Disabilities, 10*, 549–559.
- Bonfiglio, C. M., Daly, E. J., III, Martens, B. K., Lin, L. R., & Corsaut, S. (2004). An experimental analysis of reading interventions: Generalization across instructional strategies, passages, and time. *Journal of Applied Behavior Analysis, 37*, 111–114.
- Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: Applications in psychology and education* (pp. 187–212). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cossairt, A., Hall, R. V., & Hopkins, B. L. (1973). The effects of experimenter's instructions, feedback, and praise on teacher praise and student attending behavior. *Journal of Applied Behavior Analysis, 6*, 89–100.
- Daly, E. J., Martens, B. K., Dool, E. J., & Hintze, J. M. (1998). Using brief functional analysis to select interventions for oral reading. *Journal of Behavioral Education, 8*, 203–218.
- Daly, E. J., III, Martens, B. K., Hamler, K. R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis, 32*, 83–94.
- Duhon, G. J., Noell, G. H., Witt, J. C., Freeland, J. T., Dufrene, B. A., & Gilbertson, D. N. (2004). Identifying academic skills and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review, 33*, 429–443.
- Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single-case designs. *Psychology in the Schools, 37*, 463–474.
- 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.
- Ervin, R. A., Radford, P. M., Bertsch, K., Piper, A. L., Ehrhardt, K. E., & Poling, A. (2001). A descriptive analysis and critique of the empirical literature on school-based functional assessment. *School Psychology Review, 30*, 193–210.

- Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16*, 203–212.
- Gresham, F. M., & Lopez, M. F. (1996). Social validation: A unifying concept for school-based consultation research and practice. *School Psychology Quarterly, 11*, 204–227.
- Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. In N. G. Haring, T. C. Lovitt, M. D. Eaton, & C. L. Hansen (Eds.), *The fourth R: Research in the classroom* (pp. 23–40). Columbus, OH: Merrill.
- Hendrickson, J. M., Gable, R. A., Novak, C., & Peck, S. (1996). Functional assessment as strategy assessment for teaching academics. *Education and Treatment of Children, 19*, 257–271.
- Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: Further analysis of the brief assessment model for academic problems. *School Psychology Review, 31*, 554–568.
- Lentz, F. E., Jr. (1988). Effective reading interventions in the regular classroom. In J. L. Graden, J. E. Zins, & M. L. Curtis (Eds.) *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 351–370). Washington, DC: National Association of School Psychologists.
- Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analyses of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis, 26*, 293–319.
- Levy, B. A., Nicholls, A., & Kohen, D. (1993). Repeated readings: Process benefits for good and poor readers. *Journal of Experimental Child Psychology, 56*, 303–327.
- Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research & Practice, 16*, 191–198.
- McComas, J. J., Wacker, D. P., Cooper, L. J., Asmus, J. M., Richman, D., & Stoner, B. (1996). Brief experimental analysis of stimulus prompts for accurate responding on academic tasks in an outpatient clinic. *Journal of Applied Behavior Analysis, 29*, 397–401.
- The National Center for Education Statistics (2001). The nation's report card: Fourth grade reading 2000. *The National Center for Education Statistics*. Retrieved November 1, 2001, from <http://nces.ed.gov/nationsreportcard/>.
- Noell, G. H., Gansle, K. A., Witt, J. D., Whitmarsh, E. L., Freeland, J. T., LaFleur, L. H., Gilbertson, D. N., & Northup, J. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficulty. *Journal of Applied Behavior Analysis, 31*, 659–664.
- Rose, T. L., & Beattie, J. R. (1986). Relative effects of teacher-directed and taped previewing on oral reading. *Learning Disability Quarterly, 9*, 193–199.
- Rosenberg, M. S. (1986). Error-correction during oral reading: A comparison of three techniques. *Learning Disability Quarterly, 9*, 182–192.
- Skinner, C. H., Ford, J. M., & Yunker, B. D. (1991). An analysis of instructional response requirements on the multiplication performance of behavior disordered students. *Behavior Disorders, 17*, 56–65.
- Spache, G. (1953). A new readability formula for primary grade materials. *Elementary English, 53*, 410–413.
- Taylor, J. C., & Romanczyk, R. G. (1994). Generating hypotheses about the function of student problem behavior by observing teacher behavior. *Journal of Applied Behavior Analysis, 27*, 251–265.
- VanAuken, T., Chafouleas, S. M., Bradley, T. A., & Martens, B. K. (2002). Using brief experimental analysis to select oral reading interventions: An investigation of treatment utility. *Journal of Behavioral Education, 11*, 165–181.
- Wolery, M., Bailey, D. B., Jr., & Sugai, G. M. (1988). *Effective teaching: Principles and procedures of applied behavior analysis with exceptional children*. Boston: Allyn & Bacon, Inc.