Psychology in the Schools, Vol. 37(5), 2000

EMPIRICALLY EVALUATING THE EFFECTIVENESS OF READING INTERVENTIONS: THE USE OF BRIEF EXPERIMENTAL ANALYSIS AND SINGLE CASE DESIGNS

TANYA L. ECKERT, SCOTT P. ARDOIN, DONNA M. DAISEY, AND MARK D. SCAROLA

Syracuse University

Recent advances in area of brief experimental analysis have demonstrated that comparative single case designs can be used to meet criteria for empirically supported treatments. The purpose of the present study is to provide an overview of how single case design methodology can be used to demonstrate the effectiveness of school-based interventions in the area of reading. We examined the effects of combining skill-based and performance-based reading interventions on the oral reading fluency of four elementary-aged students identified with mild reading problems. A brief experimental analysis using an adapted alternating treatment design was employed to examine the effects of the intervention conditions across a 12-week period. For a majority of the participants, the effects of combining these two types of interventions resulted in greater oral reading rates than when a skill-based intervention was presented in isolation. The results of this study demonstrate how single case design elements, combined with brief experimental analysis, can be used to evaluate the effectiveness of treatments as well as treatment components. © 2000 John Wiley & Sons, Inc.

Recent publications have advocated the use of empirically supported treatments for children and adolescents suffering from clinical and educational disorders (Chamblass et al., 1996; Task Force on Promotion and Dissemination of Psychological Procedures, 1995). The basis for this argument is that practice guidelines based on scientific evidence ensures the reliability and validity of behavioral health care, and is consistent with the scientist-practitioner model of professional practice (Belar & Perry, 1992). However, two specific developments in the field of psychology, the rise of increasingly more effective psychosocial treatments and the growth of sophisticated statistical methods to document effective research outcomes, have directly influenced practice guidelines (Nathan, 1998). The impact of these developments can be directly observed in the specific criteria for establishing empirically supported treatments, delineated in the practice guidelines. These criteria include strong evidentiary support using methodologically rigorous experimental procedures, well-developed treatment manuals, and clear guidelines regarding sample characteristics (Chamblass et al., 1996).

A number of authors have attempted to incorporate these guidelines to identify empirically supported interventions for a number of adult clinical disorders, including marital distress (Baucom, Shoham, Mueser, Daiuto, & Stickle, 1998), chronic illness (Compas, Haaga, Keefe, Leitenberg, & Williams, 1998), and mental health disorders (DeRubeis & Crits-Christoph, 1998). Efforts have also been made to identify a number of empirically supported treatments for a number of childhood clinical disorders, including depression (Kaslow & Thompson, 1998), phobic and anxiety disorders (Ollendick & King, 1998), autism (Rogers, 1998), conduct disorder (Brestan & Eyberg, 1998), and Attention Deficit/Hyperactivity Disorder (Pelham, Wheeler, & Chronis, 1998). Furthermore, the benefits of developing effective and clinically useful treatments for adult and child clinical populations have been noted (Kendall, 1998; Weisz & Hawley, 1998).

Despite emerging support for the identification of empirically supported treatments, debate has ensued regarding the validity of this approach to treatment, as well as the validity of the criteria for documenting empirically supported treatments (Hughes, 1999; Lonigan, Elbert, & Johnson, 1998).

The authors express their sincere appreciation to Victoria Slaughter and Robin Shulsky, whose efforts were vital to completion of the study.

Correspondence concerning this article may be addressed to Tanya L. Eckert, Syracuse University, Department of Psychology, 430 Huntington Hall, Syracuse, NY 13244. Electronic mail may be sent via Internet to MACROBUTTON HtmlResAnchor *tleckert@PSYCH.SYR.EDU*.

First, no data supports the argument that applying empirically supported treatments will improve clinical practice outcomes (Kazdin, 1997). Second, a scientist-practitioner model of professional practice cannot be simplified to matching treatments to clinical symptoms (Martens & Eckert, in press; Nathan, 1998). Third, the extent to which laboratory research generalizes to clinical settings and practice is often unclear due to differences in sample characteristics, treatment procedures, and clinical practice needs (Kazdin, 1997; Kazdin, Siegel, & Bass, 1990). Fourth, the role of theory in conceptualizing and treating disorders cannot be ignored (Martens & Eckert, in press; Kazdin, 1997; Kazdin & Kendall, 1998). It has been argued that establishing empirically supported treatments, without developing a more effective model for guiding the progression of theory, research, and practice, may not improve the nexus between clinical research and clinical practice (Kazdin, 1997; Kazdin & Kendall, 1998; Sobell, 1996).

In the area of applied behavior analysis, a technology for empirically identifying functional relations between problem behaviors and events surrounding its occurrence has emerged (Martens & Eckert, in press). The use of functional analysis provides an empirical conceptualization of problem behaviors and specifies individually-tailored treatment procedures that have conceptual relevance (Baer, Wolf, & Risley, 1968). This allows for the application of effective behavior techniques that are grounded in theory and science. A common approach to evaluating functional relations between behavior and events surrounding its occurrence is to employ single case experimental designs. Although a number of single case experimental designs have been reported in the literature, all of these designs share common elements including collecting repeated observations over time, comparing data series across treatment and baseline conditions, and replicating these comparisons (Hayes, 1981; McCormick, 1990). The use of single case designs is in keeping with practice standards for establishing empirically supported treatments. However, most of the literature reviewing empirically supported treatments has relied predominantly on group design research. Furthermore, often empirically supported treatments for children and adolescents have focused on interventions evaluated in clinical settings, such as inpatient hospital settings, outpatient treatment centers, or summer treatment programs. Few studies have reviewed empirically supported treatments for school-based applications.

Recently, investigators have adapted elements of single case experimental designs to systematically examine school-based intervention components (Harding, Wacker, Cooper, Millard, & Jensen-Kovalan, 1994; McComas et al., 1996). These adapted experimental designs permit a brief experimental analysis of behavior by comparing a number of treatment options over a relatively short period of time (Horner, 1994). In addition, these brief experimental analyses allow for the efficient comparison of multiple treatments prior to full scale implementation (Martens, Eckert, Bradley, & Ardoin, 1999). This type of experimental verification is in keeping with practices standards for establishing empirically supported treatments. For example, Chamblass et al. (1996) and the Task Force on Promotion and Dissemination of Psychological Procedures (1995) note that the criteria for establishing efficacious interventions can include a single-case designs that employ an adequate experimental design and that compare two or more interventions.

School-Based Interventions for Reading Problems

In the area of school-based interventions for reading problems, a few studies have combined brief experimental analysis with single case experimental designs to evaluate the effectiveness of reading intervention components (Daly, Martens, Hamler, Dool, & Eckert, 1999; Noel et al., 1998). These studies have attempted to examine the instructional contingencies affecting oral reading fluency (Daly, Witt, Martens, & Dool, 1997). While overall reading ability is a complex skill involving numerous components, including word decoding, semantic access, and sentence processing (Lombard, 1988), for the purposes of these studies, oral reading fluency has been identified as a meaningful measure of the construct of reading and defined as the number of words read correctly in one

minute (Shinn, 1998). Numerous research studies have documented the validity of using oral reading fluency as an indicator of basic reading skill (see Good & Jefferson, 1998, for a review).

The theoretical underpinnings of studies examining the effectiveness of school-based reading interventions has relied on the Instruction Hierarchy (Harding, Lovitt, Eaton, & Hansen, 1978) as a model for understanding the stages of reading. According to this model, students pass through a series of stages that correspond with different reading proficiency levels. These stages include: acquisition, fluency, generalization, and adaptation. Attempts are made to match reading interventions based on a student reading proficiency level. For example, a student at the acquisition level of the instructional hierarchy requires modeling, prompting, and corrective feedback in the area of reading. Interventions that focus on these skills should be provided. However, a student at the fluency level of the instructional hierarchy requires repeated opportunities to practice oral reading. Interventions that allow students to practice reading aloud rapidly and with proficiency should be selected. In most cases, students who are referred for skill deficits in reading are most likely experiencing difficulty mastering the first two stages of the instructional hierarchy.

Conducting a brief experimental analysis of skill-based reading interventions affords an empirical examination of interventions and increases the likelihood that an effective intervention can be selected (Daly et al., 1997). For example, Daly and colleagues (1999) evaluated the relative effects of five skill-based reading interventions (i.e., interventions that target specific reading skills) on the oral reading fluency of four elementary students experiencing reading difficulties. By presenting the intervention components in a sequential, hierarchical application, comparisons were made between the relative effect of intervention components in comparison to baseline conditions. The sequential application of intervention components was discontinued when visibly discernible improvements in performance were observed. In another investigation, Noel and colleagues (1998) examined the separate and combined effects of skill-based and performance-based (i.e., providing contingent reinforcement based on performance) reading interventions to improve the oral reading fluency of three elementary students experiencing reading difficulties. The relative effects of each type of intervention were examined in a sequential, hierarchical application under different stimulus conditions (i.e., grade level material). Although both studies represent empirical advances in the literature pertaining to the assessment and identification of efficacious reading interventions, the use of sequential, hierarchical applications limits the extent to which information can be gathered regarding the effect of independent intervention components. In addition, incorporating a sequential, hierarchical design does not allow for all possible treatment combinations to be examined.

Purpose of Present Study

The purpose of the present study is to demonstrate how single case design methodology can be used to examine the effectiveness of school-based interventions in the area of reading. Specifically, we examined the effects of seven reading interventions on the oral reading fluency of four elementary-aged students. We hypothesized that all participants would demonstrate greater improvements in oral reading fluency when combined derivations of the reading interventions were presented then when baseline or an isolated reading intervention was presented. In addition, we hypothesized that participants would display differences in responding across the combined intervention conditions, replicating previous research in this area (Daly et al., 1999; Noel et al., 1998).

Method

Participants and Setting

Four male students were referred by their classroom teachers due to academic difficulties in reading. Parental permission and student assent was obtained prior to the commencement of the study.

The chronological ages of three of the four participants are as follows: Abel was 7 years, 8 months; Grafton was 7 years, 3 months; and Brett was 8 years, 5 months. Chronological age was not available for Rubin. All four participants received educational services in general education classrooms, with no supplemental or special educational services provided. Brett and Grafton were both in third grade classrooms, while Abel was in a combined first and second grade classroom, and Rubin was in a combined third through sixth grade classroom.

The setting was two large, urban, elementary school serving students in grades Kindergarten through sixth. Experimental sessions were conducted in one of two rooms used specifically for testing (9m by 9m or 5m by 5m) or an open space adjacent to the hallway (8 m by 10 m). All rooms were equipped with a table and two chairs. For the purpose of assessing interscorer agreement and procedural integrity, audiotape equipment was present during a majority of the sessions.

Materials

Reading passages were selected from the pre-primer 1, pre-primer 2, pre-primer 3, and first grade levels of the Silver, Burdett, and Ginn reading series using standard procedures for developing passages. All passages were retyped onto a separate sheet of paper to minimize the effects of pictures upon students' oral reading performance. The average length of words for the pre-primer 1 passages was 50.33 (range, 45 to 56), 54.12 for the pre-primer 2 passages (range, 46 to 69), 59.14 for the pre-primer 3 passages (range, 54 to 69), and 57.18 for the first grade passages (range, 44 to 72).

Response Measurement and Data Collection

Oral reading fluency, which was calculated by computing the number of words read correctly per minute (WRCM), was used to assess the effects of the experimental conditions on students' reading ability. WRCM was determined by subtracting the number of errors from the total number of words read during one-minute. A word was scored as correct if the student read the word correctly in 3 seconds or self-corrected a mispronounced word within 5 seconds. A word was scored as an error if a student mispronounced a word, substituted another word, omitted a word, or did not read a word within 3 seconds. In addition, if an entire line of text was skipped, then the number of words contained within the line was subtracted from the total number of words read and one error was recorded.

Procedures

General procedures. For each session, trained experimenters collected the data. Each session was approximately 15 minutes in duration and sessions occurred approximately four times a week. This analysis was conducted over a 10-week period, with no more than two experimental conditions conducted per day.

Curriculum-based assessment (CBA). Prior to beginning this study, a curriculum-based assessment in reading was conducted to determine the appropriate level to implement the interventions (see Shapiro, 1996, for a description of procedures used). It was determined that the highest instructional reading level for Brett was first grade, pre-primer 3 for Grafton, and pre-primer 2 for Abel and Rubin.

Experimental Conditions

Baseline (BL). During this condition, the student was instructed to read a passage aloud. As soon as the student read the first word in the passage, the experimenter started the stopwatch. The experimenter recorded the student's performance on the experimenter's copy of the passage by marking errors with a slash. At the end of one minute, the experimenter placed a closed bracket after the

last word read and allowed the student to read the remainder of the passage. If the student did not read a word within 3 seconds, the experimenter provided him with assistance. No other instructional intervention was provided in this condition.

Skill-based intervention. The skill-based intervention (SI) incorporated listening passage preview and repeated readings interventions. That is, students first listened to the experimenter read the passage aloud. Then each student read the passage aloud for three successive trials, while the experimenter tracked the student's performance using procedures identical to those described in baseline. Following each trial, the experimenter told the student the time in which he read the story. Results are based on the student's performance during the first minute of the last reading of the text.

Performance-based interventions. Three performance-based interventions were implemented: (a) goal setting plus performance feedback (PF); (b) contingent reinforcement (CRf); and (c) the combination of goal setting plus performance feedback and contingent reinforcement (PF+CRf).

During PF sessions, the student set goals for the amount of time it would take him to read the passage aloud and the number of errors he would make. After reading the passage, the student was informed of his reading time and number of reading errors and asked to record this information on corresponding bar graphs. During each session, the experimenter tracked the student's performance using procedures identical to those described in baseline.

In the CRf condition, the student was asked to select two educationally relevant reinforcers (e.g., pencil, ruler, eraser) from a pool of approximately 10 items. The first item chosen was considered the student's most highly preferred reinforcer of the two items. After selecting the items, the student was told that if he read the story in less than 3 minutes and made less than 3 errors he would receive the first item chosen and if he read the passage in less than 4 minutes with less than 8 errors he would receive the second item chosen. After reading the passage, the student was told his reading time and the number of errors he made and then given the reinforcer earned.

During the PF+CRf condition, the student was first asked to set goals (reading time and number of misread words) and record the goals on the corresponding bar graphs. Then the student was asked to select two items from the educationally relevant reinforcers in a fashion similar to the CRf condition. After selecting the items, the student was given the same instructions concerning how he could attain the reinforcers. After completing the passage, the student recorded his reading time and misread words on the appropriate graphs and was given the reinforcer earned.

Combined skill-based and performance-based interventions. Three combined skill-based and performance-based interventions were employed: (a) skill-based intervention combined with goal setting and performance feedback (SI+PF); (b) skill-based intervention combined with contingent reinforcement (SI+CRf); and (c) skill-based intervention combined with goal setting and performance feedback and contingent reinforcement (SI+PF+CRf). During all conditions, the experimenter tracked the student's performance using procedures identical to those described in baseline for each of the three successive readings. Results for all conditions are based upon the student's performance during the first minute of the last reading of the text.

During the SI+PF condition, the student set his two reading goals (i.e., time and errors) and marked these goals on separate graphs prior to the experimenter reading the passage to him. After each of the three successive readings of the passage by the student, the experimenter told the student his reading time and number of errors and then the student recorded this information on the appropriate graphs.

During the SI+CRf condition, the experimenter first read the story to the student and then the student was requested to select two reinforcers. After each reading trial, the experimenter told the student the amount of time he took to read the story. Prior to the third reading of the story, the student was given the same instructions regarding how he could attain the two. After the final reading

of the passage, the student was told his reading time and number of errors and then given the reinforcer earned.

When all interventions were combined in the SI+PF+CRf condition, the student selected reading goals, the experimenter read the passage, and the student then read the passage to the experimenter for three successive trials. Following each trial, the student was told his reading time and number of errors and recorded these on appropriate charts. Prior to the third trial, the student chose two reinforcers and was told the contingencies for receiving the reinforcers. If the student read the passage in less time than his second reading of the passage and with fewer errors than his second reading of the passage he received the reinforcer he chose first. If he read the passage in less time than his first reading of the passages with fewer errors than he made during the first reading, he received the less preferred reinforcer. After the last reading of the passage by the student, he was told his reading time and number of errors, recorded this information on the appropriate graphs, and was then given the reinforcer earned.

Experimental Design

To examine the effects of the experimental reading conditions on oral reading fluency, baseline and experimental conditions were presented in an adapted alternating treatment design . Phases were represented by the following notation: BL = baseline, SI = listening passage preview + repeated readings, PF = goal setting and performance feedback, CRf = contingent reinforcement, PF+CRf = goal setting and performance feedback + contingent reinforcement, SI+PF = listening passage preview + repeated readings + goal setting and performance feedback, SI+CRf = listening passage preview + repeated readings + contingent reinforcement, and SI+PF+CRf = listening passage preview + repeated readings + goal setting and performance feedback + contingent reinforcement.

Interobserver Agreement and Procedural Integrity

Interscorer agreement data were collected during 33% of the sessions for each participant, respectively. A second, independent observer recorded the students' audiotaped oral reading responses. Comparisons were conducted between the second observer's responses and the experimenter's responses on a word-by-word basis. Interscorer agreement was calculated as the total number of agreements divided by agreements plus disagreements multiplied by 100% (House, House, & Campbell, 1981). The mean agreement coefficient for WRCM was 98% (range, 96% to 100%).

Procedural integrity was conducted during 33% of the sessions for each participant. An independent observer evaluated the experimenter's performance using a procedural integrity checklist. Procedural integrity was monitored for the following areas: (a) correct presentation of materials; (b) correct delivery of the experimental instructions; and (c) accurate recording of time. Across all participants, procedural integrity was 100%.

Results

Table 1 displays the participants' mean WRCM and standard deviation across the experimental conditions. The graphic results for all four participants are presented in Figure 1. For all participants, increases in the number of WRCM were observed following presentation of the SI condition. Strong improvements in the number of WRCM were displayed for Brett from BL (range, 15 to 22) to SI (range, 44 to 56) as well as for Grafton from BL (range, 23 to 27) to SI (range, 42 to 68). Moderately strong improvements in the number of WRCM were evident for Rubin from BL (range, 25 to 36) to SI (range, 46 to 59) and mild improvements were noted for Abel from BL (range, 37 to 65) to SI (range, 41 to 76).

Table 1.
Participants' Mean Words Read Correctly and Standard Deviations across Conditions

	Words Read Correctly Per Minute							
	BL	SI	SI+CRf	SI+PF	SI+PF+CRf	CR	PF	CR+PF
Participants	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Abel	51.33 (14.01)	57.33 (17.62)	69.00 (11.00)	63.33 (10.41)	45.00 (13.00)	46.33 (7.23)	39.33 (11.59)	48.33 (4.51)
Brett	19.33 (3.79)	51.33 (6.43)	46.00 (16.70)	52.00 (6.56)	52.00 (13.75)	27.00 (8.72)	40.33 (22.75)	38.33 (21.59)
Grafton	24.33 (2.31)	57.00 (13.45)	47.00 (9.17)	37.33 (12.34)	46.00 (13.08)	32.33 (8.39)	14.00 (5.00)	20.67 (6.51)
Rubin	31.67 (5.86)	50.67 (7.23)	52.00 (1.00)	66.00 (6.56)	62.67 (5.51)	26.67 (1.53)	29.00 (13.89)	28.67 (4.51)

Notes. Underlined text refers to greatest increase in words read correctly for each participant. BL = baseline; SI = skill-based intervention; SI+CRf = skill-based intervention plus contingent reinforcement; SI+PF = skill-based intervention plus performance feedback; SI+PF+CRf = skill-based intervention plus performance feedback plus contingent reinforcement; CR = contingent reinforcement; PF = performace feedback; CR+PF = contingent reinforcement plus performance feedback.

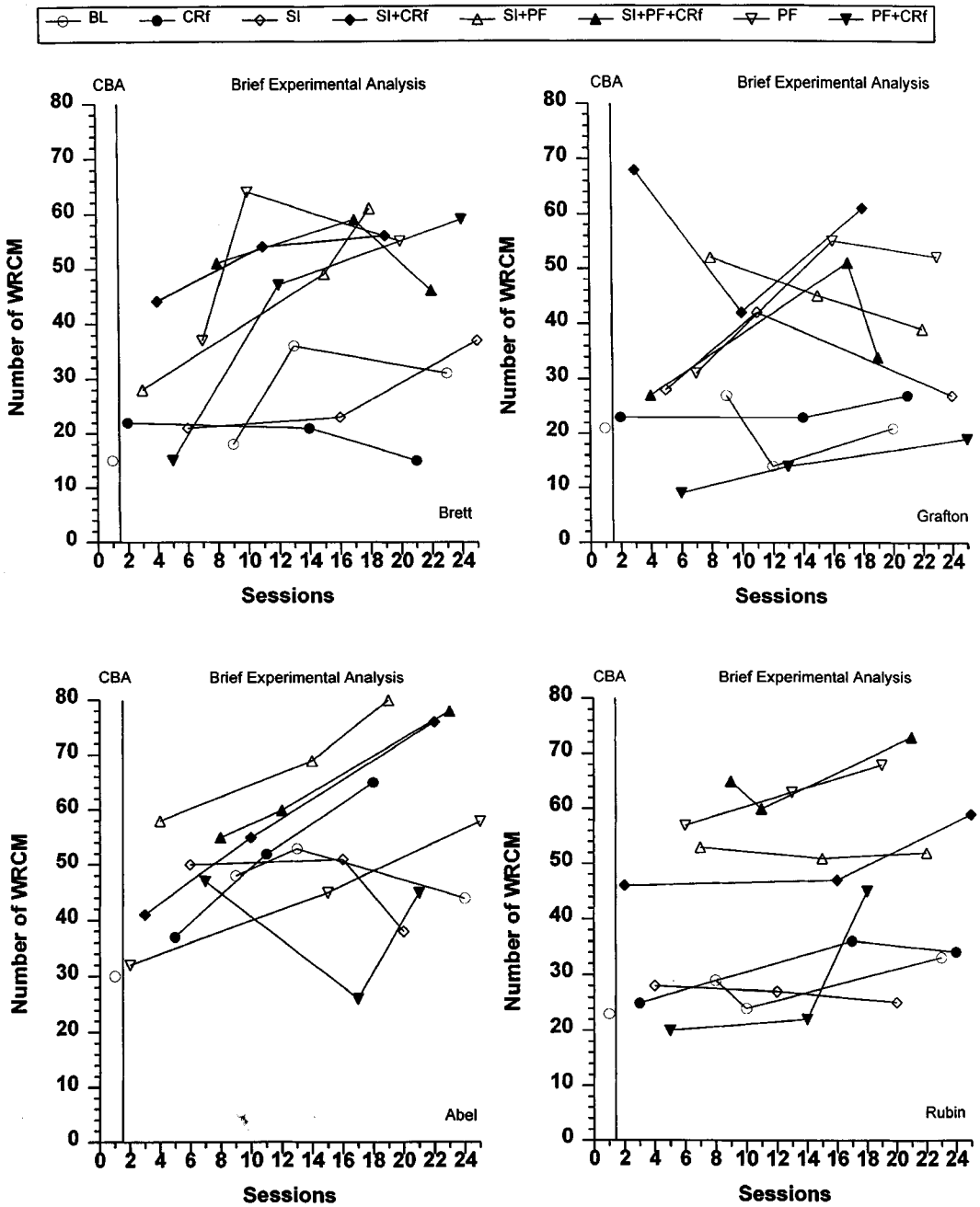

FIGURE 1. Words read correctly per minute (WRCM) across curriculum-based assessment (CBA) and brief experimental analysis for each participant.

For three of the four participants, the effectiveness of the SI was improved by combining one or more performance-based intervention. In comparison to the SI condition, consistently greater improvements in the number of WRCM were noted for Abel when either the SI+CRf (range, 55 to 75) or SI+PF (range, 32 to 58) conditions were presented. For Brett, greater improvements in the number of WRCM, in comparison to the SI condition, were observed for the SI+PF (range, 46 to 59) and SI+PF+CRf conditions (range, 37 to 64). A similar result was found for Rubin, with greater improvements in the number of WRCM noted during the SI+PF (range, 60 to 73) and SI+PF+CRf (range, 57 to 68) conditions. Although Grafton demonstrated the greatest increases in the number of WRCM during the SI condition, moderate improvements were noted during the SI+CRf (range, 39 to 57) and SI+PF+CRf (range, 31 to 55) conditions.

Discussion

The results of this study demonstrate that a brief experimental analysis of reading interventions can be used to examine the utility of one or more reading interventions for students on an individualized basis. In addition, individual differences in responding were noted, wherein, the most beneficial intervention varied across participants. The brief experimental analysis data indicated that for three of the four students, a combination of intervention conditions was associated with the greatest improvements in oral reading fluency. Thus, these results are consistent with a growing literature demonstrating the utility of incorporating brief experimental analysis methodology for identifying efficacious interventions in the area of reading (Daly et al., 1999; Noel et al., 1998).

This study suggests that examining the separate and combined effectiveness of reading interventions may afford a more comprehensive examination of relative effects. This type of comparative analysis procedure is in keeping with standards established for documenting empirically supported treatments (Chamblass et al., 1996; Task Force on Promotion and Dissemination of Psychological Procedures, 1995). Furthermore, this study demonstrates the importance of conducting exhaustive analyses of potential intervention conditions. By systematically evaluating all possible reading intervention conditions, it was possible to examine the relative effects of the reading interventions. Future research is necessary to determine the long-term effects of selecting and implementing a reading intervention based on this methodology.

The results of this study have important implications for the selection of intervention procedures to address reading difficulties. Although the most common interventions for improving reading difficulties have focused on skill-based interventions (see Sindelar & Stoddard, 1993, for a review), it appears important for clinicians to consider performance-based interventions as well as combined configurations of both intervention types. Further, reading interventions may need to be individually tailored to provide optimal treatment outcomes. In this study, we demonstrated that combining intervention components may produce greater gains in oral reading fluency then when presented in isolation. Therefore, school-based practitioners may want to consider this methodology for developing intervention programs in reading.

This study has a number of limitations that require consideration. Due to the brief nature of the experimental analysis, it is possible that an unreliable estimate of the students' oral reading fluency was obtained. For example, passages with varying degrees of difficulty may influence students' oral reading fluency. In addition, previous exposure to material presented in the passages could result in idiosyncratic performance. An extended analysis of the intervention conditions would have increased the reliability of our findings and decreased the likelihood that an erroneous conclusion was drawn. Second, within-subject and within-treatment variance limits the conclusions that can be drawn regarding the most efficacious treatment. Third, the extent to which the results were confounded by carry over effects cannot be completely dismissed. It is possible that the presentation of one intervention condition influenced the students' performance in the succeeding condition. Fourth, generalized improvements in reading performance were not observed. That is, minimal improvements

were observed in students' baseline reading data. This suggests that while the combined interventions improved oral reading fluency, they had little impact on oral reading fluency in other conditions. Finally, it is difficult to ascertain whether the observed effects will improve and maintain students' oral reading fluency over time.

On the basis of these findings, it would seem appropriate to recommend that future research examine the specific limitations of this methodology within the context of reading interventions. For example, it would be useful to compare the brief experimental analysis procedure with more extended analytic techniques. In addition, more information is needed regarding the long-term benefits of this approach selecting interventions for addressing basic deficits in reading. Using the aforementioned methodology, this could easily be established by examining the effects of the most efficacious intervention over an extended period of time.

In summary, this study offers a single case design application of empirically determining efficacious treatments in reading. The results of this study demonstrate that brief experimental analysis methodology can be used on a comparative basis to examine the relative effects of combined reading interventions for students experiencing reading difficulties. The results of this study also demonstrate that single case designs can be employed in school-based settings to assist in treatment-decision making. By employing a brief experimental analysis, school-based practitioners can empirically evaluate the relative effects of treatment options, and the results of this analysis can be directly linked to intervention.

References

Baer, D.M., Wolf, M.M., & Risley, T.R. (1968). Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1, 91–97.

Baucom, D.H., Shoham, V., Mueser, K.T., Daiuto, A.D., & Stickle, T.R. (1998). Empirically supported couple and family interventions for marital distress and adult mental health problems. Journal of Consulting and Clinical Psychology, 66, 53–88.

Belar, C.D., & Perry, N.W. (1992). National conference on scientist-practitioner education and training for the professional practice of psychology. American Psychologist, 47, 71–75.

Brestan, E.V., & Eyberg, S.M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. Journal of Clinical Child Psychology, 27, 180–189.

Chambless, D., Sanderson, W., Shoham, V., Johnson, S., Pope, K., Crits-Cristoph, P., Baker, M., Johnson, B., Woody, S., Sue, S., Beutler, L., Williams, D., & McCurry, S. (1996). An update on empirically validated therapies. The Clinical Psychologist, 49, 5–18.

Compas, B.E., Haaga, D.A.F., Keefe, F.J., Leitenberg, H., & Williams, D.A. (1998). Sampling of empirically supported psychological treatments from health psychology: Smoking, chronic pain, cancer, and bulimia nervosa. Journal of Consulting and Clinical Psychology, 66, 89–112.

Daly, E.J., 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.

Daly, E.J., Witt, J.C., Martens, B.K., & Dool, E.J. (1997). A model for conducting a functional analysis of academic performance problems. School Psychology Review, 26, 554–574.

DeRubeis, R.J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. Journal of Consulting and Clinical Psychology, 66, 163–167.

Good, R.H., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M.R. Shinn (Ed.), Advanced applications of curriculum-based measurement (pp. 61–88). New York: Guilford.

Harding, N.G., Lovitt, T.C., Eaton, M.D., & Hansen, C.L. (1978). The fourth R: Research in the classroom. Columbus, OH: Merrill.

Harding, J., Wacker, D.P., Cooper, L.J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. Journal of Applied Behavior Analysis, 27, 291–300.

Hayes, S.C. (1981). Single case experimental design and empirical clinical practice. Journal of Consulting and Clinical Psychology, 49, 193–211.

Horner, R.H. (1994). Functional assessment: Contributions and future directions. Journal of Applied Behavior Analysis, 28, 401–404.

House, A.E., House, B.G., & Campbell, M.B. (1981). Measures of interobserver agreement: Calculation formula and distribution effect. Journal of Behavioral Assessment, 3, 37–57.

Hughes, J.N. (1999). Best practices in school psychology: Chimera or coxswain. Paper presented at the annual convention of the American Psychological Association.

Kaslow, N.J., & Thompson, M.P. (1988). Applying the criteria for empirically supported treatments to studies of psychosocial interventions for child and adolescent depression. Journal of Clinical Child Psychology, 27, 146–155.

Kazdin, A.E. (1997). A model for developing effective treatments: Progression and interplay of theory, research and practice. Journal of Clinical Child Psychology, 26, 114–129.

Kazdin, A.E., & Kendall, P.C. (1998). Current progress and future plans for developing effective treatments: Comments and perspectives. Journal of Clinical Child Psychology, 27, 217–226.

Kazdin, A.E., Siegel, T.C., & Bass, D. (1990). Drawing upon clinical practice to inform research on child and adolescent psychotherapy: A survey of practitioners. Professional Psychology: Research and Practice, 21, 189–198.

Kendall, P.C. (1998). Empirically supported psychological therapies. Journal of Consulting and Clinical Psychology, 66, 3–6.

Lombard, T. (1988). Caution urged in embracing CBM assessment. NASP Communique, 16, 20.

Lonigan, C.J., Elbert, J.C., & Johnson, S.B. (1998). Empirically supported psychosocial interventions for children: An overview. Journal of Clinical Child Psychology, 27, 138–145.

Martens, B.K., & Eckert, T.L. (In press). The essential role of data in psychologic theory. Journal of School Psychology.

Martens, B.K., Eckert, T.L., Bradley, T.A., & Ardoin, S. P. (1999). Identifying effective treatments from a brief experimental analysis: Using single-case design elements to aid decision-making. School Psychology Quarterly, 14, 163–181.

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.

McCormick, S. (1990). A case for the use of single-subject methodology in reading research. Journal of Research in Reading, 13, 69–81.

Nathan, P.E. (1998). Practice guidelines: Not yet ideal. American Psychologist, 53, 290–299.

Noel, G.H., Gansle, K.A., Witt, J.C., Whitmarsh, E.L., Freeland, J.T., LaFleur, L.H., Gilbertons, D.N., & Northup, J. (1998). Effects of contingent reward and instruction on oral reading performance at differing levels of passage difficult. Journal of Applied Behavior Analysis, 31, 659–663.

Ollendick, T.E., & King, N.J. (1998). Empirically supported treatment for children with phobic and anxiety disorders: Current status. Journal of Clinical Child Psychology, 27, 156–167.

Pelham, W.E., Wheeler, T., Chronis, A., (1998). Empirically supported psychosocial treatments for Attention Deficit/Hyperactivity Disorder. Journal of Clinical Child Psychology, 27, 190–205.

Rogers, S.J. (1998). Empirically supported comprehensive treatments for young children with autism. Journal of Clinical Child Psychology, 27, 168–179.

Shapiro, E.S. (1996). Academic skills problems (2nd ed.). New York, NY: Guilford.

Silver, Burdett, & Ginn. (1991). Word of Reading. Morristown, NJ: Author.

Shinn, M.R. (1998). Advanced applications of curriculum-based measurement. New York: Guilford.

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

Sobell, L.C. (1996). Bridging the gap between scientists and practitioners: The challenge before us. Behavior Therapy, 27, 297–320.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychosocial treatments: Report and recommendations. The Clinical Psychologist, 48, 3–23.

Weisz, J.R., & Hawley, K.M. (1998). Finding, evaluating, refining, and applying empirically supported treatments for children and adolescents. Journal of Clinical Child Psychology, 27, 206–216.