

*REFINING THE EXPERIMENTAL ANALYSIS OF ACADEMIC SKILLS DEFICITS: PART II. USE OF BRIEF EXPERIMENTAL ANALYSIS TO EVALUATE READING FLUENCY TREATMENTS*

EDWARD J. DALY, III

UNIVERSITY OF NEBRASKA-LINCOLN

CHRISTINE M. BONFIGLIO

ELKHART COMMUNITY SCHOOLS

AND

TARA MATTSON, MICHAEL PERSAMPIERI, AND KRISTIN FOREMAN-YATES

UNIVERSITY OF NEBRASKA-LINCOLN

The technology of brief experimental analysis is just beginning to be used for identification of effective treatments for individual students who experience difficulty with oral reading fluency. In this study, the effect of a reading fluency treatment package was examined on easy and hard passages, and generalization was assessed on passages with high content overlap. The results suggest that the treatment package increased reading fluency for all 3 students. Effects were moderated by difficulty level for all 3 students. Results are discussed in terms of future refinements to the procedures, validation of the methods, and potential applications in clinical and school settings.

DESCRIPTORS: academic performance, experimental analysis, generalization, reading fluency

---

A number of studies have operationalized instructional and motivational variables and examined their effects on various forms of academic responding under the conceptual and methodological umbrella of functional or experimental analysis (Daly, Martens, Dool, & Hintze, 1998; Daly, Martens, Hamler, Dool, & Eckert, 1999; Daly, Murdoch, Lillenstein, Webber, & Lentz, 2002; Eckert, Ardoin, Daisey, & Scarola, 2000; Eckert, Ardoin, Daly, & Martens, 2002; Hendrickson, Gable, Novak, & Peck, 1996; Jones & Wickstrom, 2002; McComas et al., 1996; Noell, Freeland, Witt, & Gansle, 2001; Noell et al., 1998; VanAuken, Chafouleas, Bradley, & Martens, 2002). Characterized by single-case experimental designs, these experimental analyses often involve “test-

ing treatments” to identify a treatment that has potential for improving a student’s academic responding.

Most of these studies have been conducted with oral reading fluency as the target for change and the outcome measure. Oral reading fluency is a legitimate instructional target in its own right (National Reading Panel, 2000; National Research Council, 1998) because it serves as a necessary prerequisite to independent reading comprehension. Indeed, oral reading fluency is the single best measure of reading competence for students learning to read (Shinn, Good, Knutson, & Tilly, 1992). Most learners for whom English is a second language receive reading instruction in English; for these students, effective principles of reading instruction are just as necessary as for native speakers (Gersten & Baker, 2003), and oral reading fluency should figure prominently as a sensitive indicator of reading performance and instructional outcomes.

---

Address correspondence to Edward J. Daly, III, University of Nebraska–Lincoln, Educational Psychology, 33 Teachers College, Lincoln, Nebraska 68588 (e-mail: edaly2@unl.edu).

doi: 10.1901/jaba.2006.13-05

Experimental analyses will be of limited value to the field, however, if their use fails to help students ultimately attain the criterion of generalized responding. In the area of oral reading fluency, it is important that treatments that are effective during reading instruction generalize to other similar reading tasks. To date, only a few reports of experimental analyses in the literature have included direct measures of generalization across texts (Daly, Bonfiglio, Mattson, Persampieri, & Foreman-Yates, 2005; Daly, Martens, Kilmer, & Massie, 1996; Daly *et al.*, 1998, 1999; Jones & Wickstrom, 2002). In these studies, generalization has been operationalized as word overlap between instructional and subsequent probe passages.

The purpose of this investigation was to examine the effects of a treatment package designed to improve oral reading fluency. The package was initially tested in a brief experimental analysis on easy and hard passages. Students' responses to passages of differing difficulties were examined in the brief experimental analysis to determine whether differential effects would be obtained as a function of the difficulty level of passages chosen for assessment. Following the brief experimental analysis, the generalized effects were examined in passages with high content overlap during an extended analysis.

## METHOD

### *Participants and Setting*

Participants in this study were 3 students who attended the same elementary school and who had been referred by their teachers for reading problems. At the beginning of the study, Jana was 9 years 4 months old, Jorge was 9 years 10 months old, and Maria was 9 years 11 months old. Although all were in the third grade, their teachers reported that each of them was reading at a second-grade level. Jana was Kurdish, and Jorge and Maria were Hispanic. All 3 students were receiving 2 hr of service per day as students with limited English proficien-

cy. Parental consent and the child's assent were obtained prior to screening and experimental procedures. Screening and experimental sessions were conducted in a small classroom or at a table in the hallway, whichever was available.

### *Materials*

*Instructional passages.* Second-grade reading passages were chosen randomly from the Silver, Burdett, and Ginn basal reading series (Pearson *et al.*, 1989). Only narrative and expository texts were used. Passages contained approximately 170 words. These passages were used for instruction in the treatment condition. Readability scores for the second-grade passages were computed using the Spache formula (Spache, 1953). The average readability was 2.2 (range, 2.0 to 2.6).

*Passages with high content overlap.* Thirteen passages with high content overlap (HCO) were created by rewriting the original passages using the majority of words from the original instructional story. These HCO passages contained a high percentage of the same words as corresponding instructional passages (Daly *et al.*, 1996) and were used to assess generalization. The percentage of word overlap was calculated by dividing the number of words that appeared in both passages by the total number of words in the HCO passage. The mean word overlap was 85% (range, 80% to 94%). The average readability of the HCO passages was 2.94 (range, 2.3 to 3.9).

### *Dependent Variable*

Correctly read words (CRW) per minute in the HCO passages was the dependent variable. A CRW was defined as a word that is pronounced correctly in 3 s (Shinn, 1989). If the student mispronounced a word, omitted a word, substituted another word, or did not read a word within 3 s, the examiner did not record it as a CRW. Students were instructed to read aloud for 1 min to the examiner, who followed along on a separate copy of the story. As the student read a passage aloud, the

examiner recorded CRW. The number of CRW was calculated for each passage. An audiocassette recorder was used during all sessions to assess interobserver agreement and treatment integrity.

### *Sequence of Analyses*

First, students read passages during a prescreening to determine which passages were easy and which were hard. Next, a brief experimental analysis was conducted to evaluate the effects of a treatment package on oral reading fluency on easy and hard passages. Finally, the generalized effects were evaluated on HCO passages during an extended analysis.

### *Preexperimental Screening*

Students read all 13 HCO passages aloud for 1 min in several sessions across 3 days. Based on performance during the screening, passages were ranked from easiest to hardest in terms of CRW and errors per minute for each individual (i.e., the passage with the highest CRW was identified as the easiest, and the passage with the lowest CRW was identified as the hardest). The two easiest and the two hardest passages were first selected and randomly assigned either to the control condition or to the experimental condition for the brief experimental analysis. Eight of the remaining nine passages were assigned to treatment and control conditions in a semirandom fashion for the extended analysis, such that as passages were selected for the analysis (without replacement), the next two closest passages were randomly assigned to either treatment or control. This procedure was followed until there were no more close matches for difficulty level.

Across all matched pairs of passages (i.e., control and treatment passages that were matched for difficulty level), Jorge read four more CRW per minute in the control passages than in the treatment passages, Maria read six more CRW per minute in the control passages than in the treatment passages, and Jana read four more CRW per minute in the treatment

passages than in the control passages. In the generalization analysis passages only, Jorge read an average of 71 CRW per minute in control passages (range, 66 to 74) and 72 CRW per minute in HCO passages (range, 70 to 74); Maria read an average of 42 CRW per minute in control passages (range, 37 to 49) and 41 CRW per minute in HCO passages (range, 35 to 48); and Jana read an average of 55 CRW per minute in control passages (range, 51 to 59) and 54 CRW per minute in HCO passages (range, 51 to 59). Therefore, equivalence of passage difficulty across conditions was very strong.

### *Brief Experimental Analysis*

The brief experimental analysis was conducted with the easiest and hardest passages to determine the degree to which difficulty level (within the range of passages sampled) might influence conclusions resulting from brief experimental analyses. The brief experimental analyses were conducted across two sessions that examined performance in four conditions: control and treatment in easier passages and control and treatment in harder passages, as identified during the screening. During each session, the experimenter tested control and treatment in random order. The experimental comparison was the degree to which control and treatment passages were similar in difficulty level but differed in word overlap with the treatment session.

*Control.* The participant read a passage aloud for 1 min.

*Treatment.* A package of motivational and instructional variables (Daly et al., 2005), designed and demonstrated to improve oral reading fluency, was implemented with all 3 participants. Rewards were chosen by participants from a variety of tangible items (e.g., pocket games, pencils, and baseball cards). Participants were told that they could earn a reward if they met a predetermined criterion for fluency (expressed as CRW per minute) and accuracy (expressed as number of errors). The

criterion was set based on preexperimental screening results. To earn a reward, the student had to read at least one more word correctly and make three or fewer errors. Students were told that they “could earn the reward if they beat their last score while making no more than three errors.” This criterion was based on Shapiro’s (2004) standards for instructional placement and previous use of this decision rule (Daly *et al.*, 2002).

Treatment consisted of listening passage preview (Daly & Martens, 1994), repeated readings (Rashotte & Torgesen, 1985), phrase drill error correction (O’Shea, Munson, & O’Shea, 1984), and a syllable segmentation and blending lesson (when errors were repeated after the second student reading of the passage; Daly *et al.*, 2005). Listening passage preview involved having the experimenter read the passage aloud to the student while he or she followed along with his or her finger. For repeated readings students repeatedly read the passage twice. To correct participants’ errors, the experimenter had the student reread phrases containing error words three times after the experimenter modeled correct reading of incorrect words. Syllable segmentation was applied when the student read a word incorrectly again during the student reading of the passage (after having been corrected during the first student reading). Words were broken into individual syllables; the experimenter and then the student read the syllables in order and together as a word.

#### *Generalization Analysis*

Control and treatment conditions were evaluated using the remaining passages in an alternating fashion. All control and treatment conditions were counterbalanced semirandomly.

#### *Interobserver Agreement*

Experimenters audiotaped all sessions to score student performance. An independent observer scored 56% of sessions to assess interobserver agreement. Interobserver agree-

ment was computed by dividing the number of agreements (on words read correctly and words read incorrectly) on a word-by-word basis by the total number of words. The mean interobserver agreement was 98% (range, 91% to 100%).

#### *Treatment Integrity*

Experimenters followed a protocol that was in the form of a checklist (available from the first author) for each condition to prompt them to follow each step in the correct order. Verbal instructions and feedback to students were scripted. All sessions were audiotaped. An independent observer assessed treatment integrity on 53% of the experimental sessions and 65% of the control sessions. Using the same checklist, the independent observer scored the checklist while listening to the audiotaped session. The mean treatment integrity for the experimental sessions was 99.6% (range, 96% to 100%) across all participants. The mean treatment integrity for the control sessions was 98% (range, 80% to 100%) across all participants.

## RESULTS

### *Brief Experimental Analysis*

Table 1 displays the results of the brief experimental analysis for Jorge, Maria, and Jana (including screening results for the brief experimental analysis passages). CRW per minute for easier and harder passages are shown for both control and treatment conditions. For all 3 participants, performance in treatment passages was superior to performance in control passages for both difficulty levels, suggesting that treatment was effective at improving reading fluency. However, the differences between treatment and control were larger in the harder passages than in the easier passages for all 3 participants, with the difference being most dramatic for Jana. For Jorge and Maria, the highest absolute performance was achieved in the easier passage. For Jana, the highest

Table 1  
Screening and Brief Experimental Analysis (BEA) Results

Participant		Easier passages		Harder passages	
		Control	Treatment	Control	Treatment
Jorge	Screening CRW/min	82	81	64	57
	BEA CRW/min	84	88	71	83
	Change in CRW/min <sup>a</sup>	+2	+7	+7	+16
Maria	Screening CRW/min	52	53	30	27
	BEA CRW/min	33	73	29	50
	Change in CRW/min <sup>a</sup>	-19	+20	-1	+23
Jana	Screening CRW/min	60	64	46	46
	BEA CRW/min	54	65	58	74
	Change in CRW/min <sup>a</sup>	-6	+1	+12	+28

<sup>a</sup> Change in CRW per minute relative to initial screening score on the same passage prior to the brief experimental analysis.

absolute performance was achieved in the harder passage.

Conclusions regarding relative treatment effectiveness were the same across participants who were reading at different initial fluency levels (based on screening results). Differential conclusions about magnitude of treatment effect as a function of difficulty level were also the same across all 3 participants. A general pattern of greater proportional increases relative to preexperimental screening results for harder passages than for easier passages was found. For Jana, only the slightest improvement (1 CRW per minute) relative to screening results was observed during the single treatment trial conducted in the easier passage. For Jorge and Maria, greater improvement was found for the easier passage, with Jorge improving by 8% and Maria improving by 38% over screening. These results suggest that conclusions based on brief experimental analyses may vary as a function of initial fluency rates across passages of varying difficulty. An extended analysis was conducted to determine whether and to what degree differences between treatment and control observed during the brief experimental analysis would generalize to HCO stimulus materials.

#### *Extended Analysis*

Figure 1 displays the results for Jorge, Maria, and Jana during the extended analysis. Results

are ordered according to the screening results from the easiest passage to the hardest passage. (During the experimental analysis, conditions were randomized across matched pairs of passages.) Jorge consistently displayed higher rates of CRW per minute in the treatment passages relative to screening results and control passages of equal difficulty. His performance in the treatment condition exceeded the control condition on all four comparisons. Three comparisons were conducted for Maria. Her performance in the treatment condition exceeded her performance in the control condition for two of the three passages. It is noteworthy that these increases were in the two harder passages, and the effects were rather large. In the one case in which the performance increase in the control passage exceeded the increase in the instructional passage (first session), the differences between screening and experimental analysis results were much smaller. Jana's performance was (a) superior for two of the treatment passages (the first and the third instructional passages) to the corresponding control passages, (b) equal across treatment and control passages for one set of passages (see the second session), and (c) higher for the control passage than for the treatment passage for one set of passages (see the last instructional and control passages). Therefore, less consistent treatment effects were observed for Jana than for Jorge and Maria.

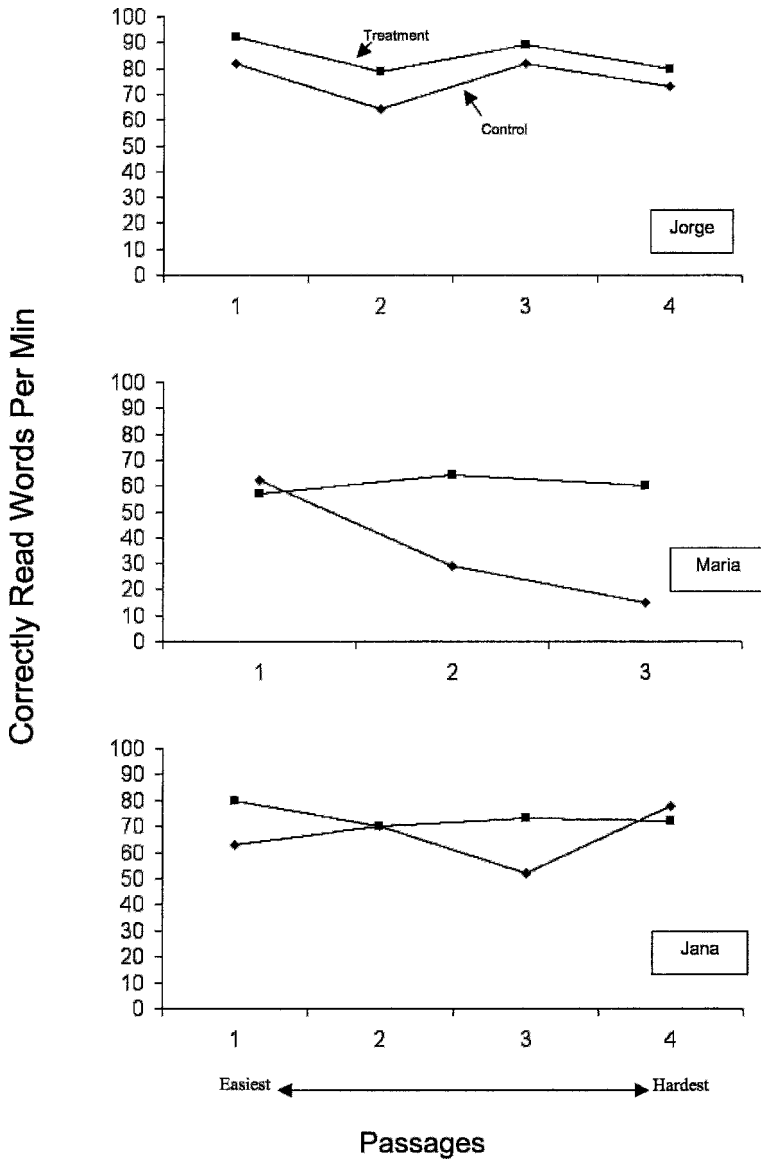


Figure 1. The number of CRW per minute for Jorge, Maria, and Jana during the extended analysis of reading performance. Note that the passages used were those that remained after the brief analysis that examined the easiest and hardest of the ranked passages. In terms of difficulty level, they were between the easy and hard passages of the brief analysis.

### DISCUSSION

The purpose of brief experimental analysis of reading fluency treatments is to efficiently identify a potentially effective treatment to improve student responding. Sensitive to time constraints of educators, investigators have

attempted to minimize the number of data points gathered to increase the efficiency of the process (Martens, Eckert, Bradley, & Ardoin, 1999). Although efforts have been made to control for difficulty level by consulting readability analyses and conducting preexperimental analyses in many studies, few investiga-

tions have systematically examined difficulty level of passages and potential impact of this variable on conclusions. In the current study, difficulty level was manipulated as a part of the brief experimental analysis. Had only a single instructional trial been conducted with one or the other of the passages, different conclusions regarding the presence (Jana) and magnitude (Jorge and Jana) of treatment effects might have been reached. What is especially interesting is that results were stronger in harder passages than in easier passages for all participants, suggesting that instruction in easier materials may not always produce the most beneficial effect, as is often assumed.

In a general sense, results from the extended experimental analyses support conclusions of overall treatment effectiveness inferred from results of brief experimental analyses. For instance, positive effects were observed for all participants in both brief and extended analyses. There were more subtle commonalities as well. Jana displayed the most variable results across both types of analyses. Also, absolute performance across students was the same in both analyses. However, the treatment interaction effect with difficulty level found during the brief experimental analysis was only observed with Maria during the extended analysis. This finding may not be surprising in light of two facts. First, the passages used in the extended analysis were closer in difficulty level than those used in the brief experimental analyses and may fall within a range in which the treatment interaction is not observable. Second, Maria was the least fluent reader of the group and may have had the greatest range for improvement; Jorge and Jana may have been approaching an upper limit of a range within which results vary within a single treatment session. Jorge displayed consistent but smaller differences between treatment and control conditions, and Jana's treatment effects were less consistent across passages (not varying at all by difficulty level).

In terms of addressing appropriate instructional difficulty level, the results of this study must be considered preliminary. First, it is not possible to determine whether the optimal ranges were chosen or whether maximum treatment effects were achieved based on this study alone. Designating passages as "easy" or "hard" was done on an individual basis (i.e., based on preexperimental screening results) and therefore varied by participant. Thus, what is easier for one participant may have been harder for another participant. Future parametric research could perhaps establish quantitative ranges for difficulty level that are generalizable across learners. Second, the results do not speak to differential trends that may result over time for treatment at one difficulty level as opposed to another difficulty level. Finally, readers should keep in mind that differences in difficulty level across passages in this study were relatively small. Greater differences across passages or at other levels of responding may yield different results. Nonetheless, the results are consistent with those of Daly et al. (2005), who used the same procedures and found that all participants showed greater relative increases in harder materials and, for 1 participant, greater absolute levels of responding in harder materials. Therefore, future investigations will contribute to our understanding of the broader notion of instructional match by considering a wider range of performance levels and examining relative as well as absolute performance increases.

The current study was conducted with the purpose of examining whether the outcomes of a brief experimental analysis might be influenced by difficulty level of the reading passages. Future studies of brief experimental analysis may identify more effective ways to examine difficulty level systematically so as to make stronger treatment recommendations to educators. If this line of investigation is fruitful, those treatment recommendations may one day be accompanied by expected effects for passages of

differing difficulty. Improving treatment recommendations in this way may lead teachers to more carefully control difficulty level prior to treatment.

A limitation of the current study is that the strong treatment that was used to produce generalized responding was never dismantled to identify which treatment variables contributed to observed effects. This was not an explicit purpose of the study. However, clinicians or investigators may consider using decreasing intensity designs (Barnett, Daly, Jones, & Lentz, 2004) by starting with strong multicomponent treatments like the one used in this study to bring the problem behavior under control more quickly. Subsequent portions of the analysis might involve withdrawing components or degrading the treatment in some way until the optimal treatment package is identified. Because generalization across texts has proven difficult and elusive, decreasing intensity designs that probe for appropriate forms of generalization may lead to effective treatments more quickly.

## REFERENCES

- Barnett, D. W., Daly, E. J., III, Jones, K. M., & Lentz, F. E., Jr. (2004). Empirically based special service decisions from increasing and decreasing intensity single case designs. *Journal of Special Education, 38*, 66–79.
- Daly, E. J., III, Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2005). Refining the experimental analysis of academic skills deficits: Part I. An investigation of variables that affect generalized oral reading performance. *Journal of Applied Behavior Analysis, 38*, 485–498.
- Daly, E. J., III, & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis, 27*, 459–469.
- Daly, E. J., III, 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.
- Daly, E. J., III, Martens, B. K., Kilmer, A., & Massie, D. (1996). The effects of instructional match and content overlap on generalized reading performance. *Journal of Applied Behavior Analysis, 29*, 507–518.
- Daly, E. J., III, Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E. (2002). An examination of methods for testing treatments: Conducting experimental analyses of the effects of instructional components on oral reading fluency. *Education and Treatment of Children, 25*, 288–316.
- 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.
- Gersten, R., & Baker, S. (2003). English-language learners with learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 94–109). New York: Guilford.
- 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.
- 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.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Available on-line: <http://www.nichd.nih.gov/publications/nrp/smallbook.htm>
- National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Noell, G. H., Freeland, J. T., Witt, J. C., & Gansle, K. A. (2001). Using brief assessments to identify effective interventions for individual students. *Journal of School Psychology, 39*, 335–355.
- Noell, G. H., Gansle, K. A., Witt, J. D., Whitmarsh, E. L., Freeland, J. T., & LaFleur, L. H., et al. (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.
- O'Shea, L. J., Munson, S. M., & O'Shea, D. J. (1984). Error correction in oral reading: Evaluating the effectiveness of three procedures. *Education and Treatment of Children*, *7*, 203–214.
- Pearson, P. D., Johnson, D. D., Clymer, T., Indirsano, R., Venezky, R. L., & Baumann, J. F., et al. (1989). *Silver, Burdett, and Ginn*. Needham, MA: Silver, Burdett, and Ginn, Inc.
- Rashotte, C. A., & Torgesen, J. K. (1985). Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly*, *20*, 180–188.
- Shapiro, E. S. (2004). *Academic skills problems: Direct assessment and intervention*. (3rd ed.). New York: Guilford.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.
- Shinn, M. R., Good, R. H., III, Knutson, N., & Tilly, W. D., III (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*, 459–479.
- Spache, G. (1953). A new readability formula for primary grade materials. *Elementary English*, *53*, 410–413.
- 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.

Received January 7, 2005

Final acceptance February 21, 2006

Action Editor, Jennifer J. McComas