

Outcomes of a Teacher-Led Reading Intervention for Elementary Students at Risk for Behavioral Disorders

KATHLEEN LYNNE LANE

M. ANNETTE LITTLE

JENNY REDDING-RHODES

ANDREA PHILLIPS

MEGAN T. WELSH

Peabody College, Vanderbilt University

ABSTRACT: *To date, reports of reading interventions for students at risk for emotional/behavioral disorders (E/BD) that have been published in refereed journals have involved sustained support by university or school-site personnel. This study examined the efficacy and feasibility of a reading intervention that 2 general education teachers implemented in inclusive settings to support 7 first-grade students at risk for E/BD and reading difficulties. Results of a multiple baseline design revealed lasting improvements in reading fluency for all students, accompanied by decreases in variability of academic engagement for 4 students. Although intervention goals, procedures, and outcomes exceeded teacher expectations, social validity ratings for some students declined between the onset and the conclusion of the intervention. This article presents limitations, future directions, and educational implications.*

Educators have long recognized that students with emotional/behavioral disorders (E/BD) and those who are at risk for such problems have social and behavioral patterns that impede instruction. For example, these students often misinterpret neutral social cues as hostile, exhibit verbal and physical aggression, and demonstrate high levels of non-compliance (Walker, Ramsey, & Gresham, 2004).

In recent years, educators have also recognized the academic characteristics of these youngsters, such as low levels of academic engagement (Van Acker & Talbott, 1999) and below-average performance in core academic areas (e.g., reading, writing, and math; Mattison, Spitznagel, & Felix, 1998; Nelson, Benner, Lane, & Smith, 2004).

The increased attention to the academic needs of this population is encouraging, because their academic deficits, like their antisocial behav-

ior patterns (Kazdin, 1987), tend to broaden over time (Anderson, Kutash, & Duchnowski, 2001; Nelson, Benner et al., 2004). In the absence of effective interventions, students with E/BD are more likely than general education students and students in other high-incidence disability categories (e.g., learning disabilities) to experience academic failure, be retained in grade, and leave school before graduation (Wagner & Davis, 2006; Wagner, Kutash, Duchnowski, & Epstein, 2005). These students struggle beyond the school setting, as evidenced by underemployment and unemployment, impaired social relationships, and higher than average need for mental health services (Walker et al., 2004; Zigmond, 2006).

Thus, it is imperative for the research and teaching communities to continue to explore the most effective and feasible methods for meeting the academic needs of students with and at risk for E/BD—particularly in the area of reading, because reading is a keystone skill that allows students to access all subsequent learning (Foorman, Francis, Shaywitz, Shaywitz, & Fletcher, 1997; O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2002). Of related interest is the extent to which improved reading performance also produces collateral effects on behavior (Hinshaw, 1992; Petras et al., 2004). In other words, do problem behaviors diminish when students have the reading skills necessary to participate in teacher-assigned instructional activities? Do other variables such as cognitive ability or inattention mediate this relationship? Is this pattern of responding the same for students with E/BD as for those at risk for E/BD?

READING INTERVENTIONS: EFFECTS ON READING AND BEHAVIORAL PERFORMANCE

Although extensive research has described the relationship between problem behaviors and academic underachievement in reading (Hinshaw, 1992; Petras et al., 2004), only a limited number of school-based, treatment-outcome studies have attempted to determine the immediate and collateral effects of academic interventions for students with or at risk for E/BD who also have limited reading skills (Lane, 2004). A recent, systematic review of the

literature of reading interventions conducted in the past 10 years with students with and at risk for E/BD identified five reading intervention studies with elementary age students that examined reading and behavioral outcomes (Lane, Barton-Arwood, Rogers, & Robertson, in press). The subjects for two of these studies were identified students who were receiving special education services, and the subjects of the other three studies were students identified as at risk for E/BD.

INTERVENTIONS FOR STUDENTS WITH E/BD

Wehby, Falk, Barton-Arwood, Lane, and Cooley (2003) examined the effects of an intensive reading program that combined a modified version of Open Court Reading (OC; Adams et al., 2000) and Peer Assisted Learning Strategies (PALS; Fuchs, Fuchs, Mathes, & Simmons, 1997) with eight students receiving special education services for behavior and learning problems. The intervention occurred in two self-contained classrooms; research assistants (RAs) conducted the OC instruction, and the classroom teachers conducted the PALS instruction. The researchers used a multiple baseline design to examine the associated effects on early literacy skills, attending, and inappropriate behavior. Results revealed moderate improvements in sound naming, blending, and nonsense word fluency. There were slight increases in attending; however, the researchers noted no decreases in inappropriate behavior. They did not collect maintenance data to assess sustainability of effect nor did they assess social validity.

Do problem behaviors diminish when students have the reading skills necessary to participate in teacher-assigned instructional activities?

In a similar study, Barton-Arwood, Wehby, and Falk (2005) assessed the effects of a comprehensive reading program that included Horizons Fast Track A-B Reading Program (Engelmann, Engelmann, & Davis, 1997) and PALS (Fuchs et al., 1997) on the reading and behavioral performance of 6 third-grade students with E/BD. The

researchers conducted this intervention in a self-contained public school for special education students with severe behavioral concerns. RAs provided instruction in Horizons, and noncertificated classroom personnel (two paraprofessionals and one mental health counselor) conducted PALS with student dyads, both of which were introduced simultaneously during the intervention phase. Results of a multiple baseline across participants design revealed improvements in blending and segmentation for most students and in word reading for all students. The students did not exhibit improved levels of engagement; however, students in the second leg showed decreases in inappropriate behavior. Again, the researchers did not report generalization and social validity data.

Collectively, these two studies support modest improvements in reading skills with limited collateral effects on behavior for students receiving special education services for behavioral and learning concerns. However, because maintenance data were absent, ascertaining the extent to which the intervention produced meaningful, *lasting* changes—the goal of all intervention work—was impossible (Baer, Wolf, & Risley, 1968).

Further, the behavior patterns may have been resistant to intervention because they were severe and great enough that they ultimately necessitated special education services. Perhaps students with less severe behavioral patterns would have experienced improved levels of engagement and decreased disruption after they acquired the requisite reading skills.

INTERVENTIONS FOR STUDENTS AT RISK FOR E/BD

Of the three studies conducted with students at risk for E/BD, all involved supplemental reading programs; and either RAs (Lane, O'Shaughnessy, Lambros, Gresham, & Beebe-Frankenberger, 2001; Nelson, Stage, Epstein, & Pierce, 2005) or school-site support staff other than teachers (e.g., a literacy coach; Lane et al., 2002) conducted the interventions. For example, Lane et al. (2001) examined the efficacy of a supplemental program, Phonological Awareness Training for Reading (PATR; Torgesen & Bryant, 1994) on the early literacy skills, disruptive classroom behavior, and social interactions of seven general education stu-

dents identified as at risk for E/BD. The intervention took place in empty classrooms or conference rooms on a general education campus. RAs conducted the PATR intervention during the intervention phase. Results of a multiple baseline across intervention groups design revealed improvements in nonsense word fluency and oral reading fluency that continued into the follow-up phase. Further, total disruptive behavior decreased from baseline to intervention phases. Six students also showed decreases in negative interactions on the playground setting, with two students maintaining these changes into the follow-up phase. Both teachers and students rated the intervention as socially valid.

Nelson et al. (2005) used group design methodology to examine the impact of a supplemental prereading intervention, Stepping Stones to Literacy (Nelson, Cooper, & Gonzalez, 2004), on the reading and behavioral performance of 63 kindergarten students at risk for E/BD and reading problems. RAs who were paraprofessional-level tutors conducted the one-to-one intervention ($n = 47$). Students in the control condition ($n = 16$) received regular school practices. The researchers did not specify the location of the intervention. Results of multivariate analyses revealed significant improvements in phonological awareness, word identification, word-attack skills, and letter-naming fluency for students in intervention relative to controls. Further, no significant differences occurred between the treatment and control groups on the behavior rating scale. Unfortunately, the study reported neither social validity nor maintenance data to examine consumer perspectives or sustainability of effects.

Only one study involved a school site literacy coach who conducted an intervention. Lane et al. (2002) examined the effects of a supplemental early literacy program, Shefelbine's *Phonics Chapter Books* (1998), with seven first-grade students who were nonresponsive to a schoolwide primary prevention program. The literacy coach implemented the intervention in a small-group format in the general education classroom during the traditional school day. The study used a multiple baseline across intervention groups design to evaluate the impact of the program on early literacy skills, disruptive classroom behavior, and negative social interactions on the playground. Results

revealed strong growth in nonsense word fluency, as well as improvements in oral reading fluency. In addition, the students exhibited lasting decreases in disruptive classroom behavior and negative social interactions on the playground. The study reported social validity for teachers and students, with the ratings generally favorable.

In sum, the results of studies conducted with students at risk for E/BD using single-case methodology suggested that improved reading performance was associated with collateral effects on direct observations of behavior. However, the same was not true for the group design study (Nelson et al., 2005) that used a behavior rating scale to measure behavioral performance.

As a whole, this body of literature suggests the efficacy of small group (Barton-Arwood et al., 2005; Lane et al., 2001; Lane et al., 2002; Wehby et al., 2003) and one-to-one (Nelson et al., 2005) interventions in improving the early reading skills of elementary students with or at risk for E/BD. In each study, students showed improvements in such reading skills as decoding and oral reading fluency. However, all the interventions required the sustained support of additional adults such as RAs (e.g., Lane et al., 2001; Nelson et al., 2005) and/or additional school personnel (e.g., paraprofessionals or reading specialists; Barton-Arwood et al.; Lane et al., 2002). In no instances did the students' teachers assume the role of sole primary interventionist.

In addition, the results of these studies do not offer definitive conclusions about the extent to which improved reading skills are associated with collateral effects on behavior for students with or at risk for E/BD. Partial evidence supports the existence of collateral effects on disruptive behavior exhibited by students at risk for E/BD when researchers assess behavior by using direct observations of student behavior (Lane et al., 2001; Lane et al., 2002). The lack of treatment outcomes for behavioral change, as measured by behavior rating scales, is consistent with other studies using similar outcome measures (e.g., Lane, 1999). Schneider (1991) indicates that teacher ratings in intervention studies tend to be stable over time and do not always align with direct observation findings. Thus, it is important to include additional measures such as direct

observations that are more sensitive to change (Lane, 1999) to accompany behavior rating scales.

THE PRESENT STUDY

The current study examined the effectiveness of a supplemental reading intervention, PALS, in improving the early literacy skills of first-grade students at risk for E/BD who also had co-occurring reading deficits. PALS has met with demonstrated success with students with E/BD (Barton-Arwood et al., 2005; Wehby et al., 2003) when RAs or paraprofessionals implement it in self-contained settings in conjunction with other evidence-based programs (e.g., OC, Fast Tracks). Further, PALS has been implemented successfully by general education teachers to support students with poor reading skills (Allor, Fuchs, & Mathes, 2001; Fuchs et al., 2001), diverse learners with low-, average-, and high-performance patterns (Fuchs et al., 1997; Mathes, Howard, Allen, & Fuchs, 1998), and learning disabilities (Fuchs, Fuchs, Mathes, & Martinez, 2002). However, this is the first attempt to have general education teachers serve as the sole primary interventionist with the goal of meeting the academic needs of students at risk for E/BD in a general education, inclusive setting with only limited initial assistance from other support staff.

This study is important for three key reasons. First, in addition to having academic deficits comparable to those of students with learning disabilities, students with and at risk for E/BD have behavior patterns that often impede the ability of even the most seasoned teachers to deliver instruction (Walker et al., 2004). These students often exhibit noncompliant, disruptive behaviors to escape instructional tasks and seek teacher attention (Umbreit, Lane, & Dejud, 2004). Such behavior makes it difficult—if not impossible—to implement evidence-based instructional practices such as PALS within the context of inclusive classrooms with sufficient treatment integrity to obtain the desired outcomes with these students. In short, students at risk for E/BD may exhibit behaviors so challenging that teachers may not be able to implement otherwise effective instructional interventions as designed. The literature suggests that general education teachers can im-

plement PALS in whole-class contexts quite successfully with students without E/BD and that university personnel in self-contained settings can implement PALS with students with E/BD. Yet, researchers have not determined whether general educators, with limited support, can implement these practices reasonably and effectively with students who have E/BD while still meeting the multiple needs of all students in inclusive settings.

Second, evidence suggests that improving the early reading skills of students with E/BD is possible, as previously discussed. However, all intervention efforts to date have required ongoing, high-magnitude support from other adults. Given the trend toward inclusive programming (Fuchs & Fuchs, 1994) and the call for academic excellence for all students (e.g., No Child Left Behind; Fournier, 2002), it is imperative that researchers and practitioners identify evidence-based practices that teachers in general education classrooms can implement with limited—or at least reasonable—support. One goal of most intervention research is to determine how to build teachers' capacity to implement with fidelity and sustain evidence-based practices in the general education setting with reasonable support while attending to the numerous demands of the classroom setting. With this study, we hope to obtain a sense of what is necessary to build capacity at the school site for implementing and sustaining evidence-based practices for this difficult-to-teach population.

Third, when educators address reading problems (and possibly behavior problems) during the early elementary years, students can participate more fully in academic instruction, thereby decreasing the likelihood that their skill and knowledge deficits will become greater (Bullis & Walker, 1994). This third reason is essential as students shift from "learning to read" to "reading to learn" in subsequent grades (Foorman et al., 1997).

In addition, this study extends this line of inquiry by examining the degree to which a general education teacher-led reading intervention is (a) able to produce meaningful, lasting changes in the reading skills of seven students at risk for E/BD and (b) whether these improved reading skills are associated with classroom behavior. In previous studies of students with E/BD, RAs (Lane et al., 2001; Nelson et al., 2005) or school-site personnel (Lane et al., 2002) have offered the sustained in-

tervention support. On the basis of the results of previous studies, we anticipated sustained improvements in students' decoding and oral reading skills. Further, we expected that students' level of academic engagement during literacy instruction would increase following the completion of the intervention, given that the students' negative behavior patterns were not as pronounced as those of students already identified with E/BD (Barton-Arwood et al., 2005; Wehby et al., 2003).

METHOD

PARTICIPANTS AND SETTING

The participants were 7 first-grade Caucasian students (3 girls and 4 boys) identified as at risk for behavioral and reading difficulties. Their ages ranged from 6.1 to 7.7 years, with a mean age of 6.96 ($SD = 0.66$). Two students, Haley and James, had intellectual functioning that was slightly below average (83), as measured by a short form of the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991), which included two subtests: block design and vocabulary subtests (split-half reliabilities of .82 for both). We used Satler's (1991) conversion formula ($r = 0.91$) to compute an estimate of intellectual functioning. The remaining students scored in the average intellectual range. None of the students were receiving special education services nor did they have any diagnosed disorders according to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR; American Psychiatric Association, 2000). However, six of the students received additional support from the reading specialists at their school during the entire study (see Table 1).

The students were attending school in two first-grade general education classrooms in two schools in rural middle Tennessee. The district subscribed to a full-inclusion model, with no self-contained classes at any elementary school. Three students were in one classroom, and four were in another. Both teachers were female and Caucasian, and both held general education credentials. The first teacher had 36 years of teaching experience, and the second had been teaching for

TABLE 1
Participant Characteristics

| Variable | Level | Students | | | | | | | |
|--------------------------------|-----------------------|---------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|-------------------------------|----------------|-------------------------------|--|
| | | Elmer | Jimmy | Rose | Haley | James | Alex | Kylie | |
| Demographics | Gender | M | M | F | F | M | M | F | |
| | Age | 7.1 | 7.0 | 6.11 | 6.1 | 7.0 | 7.7 | 7.7 | |
| | Ethnicity | Caucasian | Caucasian | Caucasian | Caucasian | Caucasian | Caucasian | Caucasian | |
| | IQ | 97.1 | 100 | 91 | 83 | 83 | 97 | 86 | |
| | Group | 1 | 1 | 1 | 2 | 2 | 2 | 2 | |
| | Previous intervention | Academic | Control | Academic | Control | Academic | Control | Control | |
| WJ-III SS (%tile) | Broad reading | 86 (18) | 86 (17) | 80 (9) | 84 (14) | 70 (2) | 82 (12) | 72 (3) | |
| | Word attack | 108 (71) | 99 (48) | 101 (52) | 106 (66) | 83 (13) | 102 (54) | — ^a | |
| SSRS-T (SS) | Problem behavior | 102 | 90 | 121 ^b | 84 ^c | 108 | 110 | 121 ^b | |
| | Social skills | 106 | 94 | 78 ^c | 106 | 92 | 98 | 82 ^c | |
| | Academic competence | 104 | 86 | 79 ^c | 87 | 84 ^c | 100 | 88 | |
| SSRS-P (SS) | Problem behavior | 94 | 84 ^c | 116 ^b | 93 | 100 | — ^a | 114 | |
| | Social skills | 118 ^b | 109 | 103 | 116 ^b | 109 | — ^a | 112 | |
| Sessions attended <i>n</i> (%) | Training | 8 (100) | 7 (88) | 8 (100) | 7 (88) | 7 (88) | 8 (100) | 8 (100) | |
| | Intervention | 28 (100) | 28 (100) | 28 (100) | 26 (93) | 28 (100) | 27 (96) | 28 (100) | |
| Supplemental Support | | Small group (30 min/ 3 days/wk) | Small group (30 min/ 3 days/wk) | Small group (30 min/ 3 days/wk) | Small group (30 min/ 3 days/wk) | 1:1 (30 min/ 5 days/wk) | none | 1:1 (30 min/ 5 days/wk) | |

Note. IQ = estimate of intellectual function; SS = standard score; %tile = percentile score; Supplemental support = supplemental instruction, separate from the project, which the school provided. WJ-III = Woodcock-Johnson III Tests of Achievement; SSRS-T = Social Skills Rating System—Teacher Version; SSRS-P = Social Skills Rating System—Parent Version.

^aNot received. ^bAbove-average. ^cBelow-average

15 years. Class sizes were comparable, with the teachers serving 22 and 21 students, respectively.

PROCEDURES

Participant Selection Criteria. All seven students in the present study were participants in a larger study, Project PREVENT (see Lane, Wehby, Phillips, Weisenbach, Little, & Merwin, 2006, for the outcomes of the overall study). The researchers selected students for participation in the overall study during the fall of their kindergarten year according to one of three systematic screening tools: the Systematic Screening for Behavior Disorders (SSBD; Walker & Severson, 1992); the Student Risk Screening Scale (SRSS; Drummond, 1994); or the Achenbach (1991) aggression subscale of the Teacher Report Form (TRF). In addition, the students had difficulties in reading, as measured by the Woodcock Johnson III (WJ; Woodcock, McGrew, & Mather, 2001). Specifically, the researchers invited students to participate if they (a) exceeded normative criteria on Stage 2 of the SSBD for either the externalizing or internalizing domains; (b) scored as moderate or high risk for antisocial behavior on the SRSS; or (c) scored in the borderline or clinical range on the TRF. Students also had academic deficits in reading, operationally defined as scoring at or below the 25th percentile on one or more of the Woodcock Johnson III (WJ-III) reading subtests.

Systematic Screening for Behavior Disorders. The SSBD (Walker & Severson, 1992) is a cost-effective, validated, multiple-gating procedure used to identify elementary-age students with behavior disorders. The SSBD evaluates students on externalizing and internalizing dimensions. In Stage 1, each teacher places in rank order his or her students from "most like" to "least like" on internalizing and externalizing dimensions to determine the students whose behavior patterns most closely parallel the behavioral profiles described. The three highest-ranked students on the internalizing and externalizing dimensions move through Gate 1 into Stage 2. In Stage 2, teachers complete two rating scales, the Critical Events Index (CEI) and the Combined Frequency Index (CFI), for the six students in Stage 2. The CEI is a 33-item checklist of "behavioral earthquakes,"

that is, low-frequency, high-intensity behaviors (e.g., setting fires, vomiting after eating). Teachers record the presence or absence of each behavior. The CFI assesses high-frequency, low-intensity behaviors on maladaptive domains (e.g., pouting or sulking) and adaptive domains (e.g., doing seatwork as directed). Students exceeding normative criteria move through Gate 2 into Stage 3. In Stage 3, a professional other than the teacher (e.g., a psychologist) conducts systematic observations of academic engaged time during seatwork and observes peer interactions during unstructured times. The SSBD distinguishes among students with externalizing behaviors, internalizing behaviors, emotional disturbances, and those with typical development (e.g., Walker et al. 1994), with coefficient alphas above .90 for the standardization sample (Walker et al., 1990).

Student Risk Screening Scale. The SRSS (Drummond, 1994) is a cost-effective, seven-item mass screening tool used to identify elementary-age students who are at risk for antisocial behavior patterns (Drummond, Eddy, & Reid, 1998a, 1998b). Teachers rate each student by using a 4-point Likert-type scale (*never* = 0, *occasionally* = 1, *sometimes* = 2, *frequently* = 3) on seven items: (a) steals; (b) lies, cheats, sneaks; (c) behavior problems; (d) peer rejection; (e) low achievement; (f) negative attitude; and (g) aggressive behavior. Total scores range from 0 to 21, with high scores suggesting high levels of antisocial behavior. Risk status ranges from low (0 to 3) to moderate (4 to 8) to high risk (9 to 21). The SRSS is able to differentiate between students who exhibit early behavior patterns indicative of antisocial behavior patterns and those who do not. The SRSS significantly correlates ($r = 0.79$) with the aggression subscale of the Child Behavior Checklist (CBCL; Walker, Ramsey, & Gresham (2004). Further, longitudinal studies of the predictive utility of the SRSS indicate that SRSS scores are predictive of negative behavioral and academic outcomes from 18 months to 10 years (Drummond, Eddy, Reid, & Bank, 1994).

Achenbach Screening. The Achenbach screening procedure (Achenbach, 1991) uses a combination of teacher nominations and teacher ratings on the aggression subscale of the Teacher Report Form (TRF; Achenbach; $r = 0.92$). After reading a description of antisocial behavior, teachers

nominate up to five students whose behavioral repertoires most closely parallel the antisocial profile (Walker et al., 1995). Next, teachers complete the aggression subscale items ($n = 15$; e.g., explosive unpredictable behavior) for these students, rating each item on a 3-point Likert-type scale ranging from 1 to 3 (1 = *not true*, 2 = *somewhat or sometimes true*, and 3 = *very true or often true*). Students with aggression scores that are two or more standard deviations above the mean are considered at heightened risk for antisocial behavior, thereby warranting early interventions to address their ideographic behavioral concerns. The TRF is a psychometrically sound instrument that has "set the standard for psychometric excellence in the study of child psychopathology and adjustment problems" (Severson & Walker, 2002, p. 48).

After we secured parental consent and student assent, we administered a battery of assessments, including the WJ-III. We then assigned students with co-occurring reading deficits, operationally defined as scoring at or below the 25th percentile on one or more of the reading subtests, to one of three conditions: academic, behavioral, or control. The students in the current study were either students in the control condition ($n = 4$) whose follow-up WJII Broad Reading Scores were at or below the 25th percentile on the test given following summer break or students who completed the full academic intervention ($n = 3$), which included small-group instruction in Ladders to Literacy (O'Connor, Notari-Syversen, & Vadasky, 1998), consisting of three 30-min sessions each week for 16 weeks but who were non-responsive ($n = 2$) or did not maintain their gains over the summer break ($n = 1$; Elmer). Therefore, intervention efforts implemented during the following academic year used single-case methodology. The researchers assigned the three students in the first class to the first intervention leg and assigned the four students in the second class to the second intervention leg.

INTERVENTION: PEER ASSISTED LEARNING STRATEGIES (PALS)

Description. First-grade PALS (Fuchs et al., 1997) is a peer-directed, structured, supplemental reading program used to teach decoding and reading fluency skills. In both classrooms, teach-

ers paired stronger readers (those in the top 50% of the class) with weaker readers (those in the bottom 50% of the class). The teachers paired student participants ($n = 7$) with teacher-selected, general education students who had typical or above-average reading skills. Pairs remained together during the entire intervention phase. Each student took turns being the coach (tutor) and the reader (tutee). The lessons included three key activities: sounds and words, a speed game, and partner reading. During sounds and words instruction, the teacher modeled how to segment and blend words. The lessons next focused on four partner activities: saying sounds, sounding out, sight words, and stories. During these activities, students focused on letter-sound correspondence, blending, sight-word recognition, decoding, and reading connected text. During the speed game, the student pairs alternated reading the same sight words or stories (three trials each) during a fixed period of time. The goal was for students to increase the number of words that they read in each successive reading; thus, students competed against themselves and not against their partner.

During the fifth week, students began partner reading. The students took turns reading aloud for 10 min from a level-appropriate book. The program contained a systematic reinforcement component, implemented across all activities, in which students explicitly reinforced each other with verbal praise and points. Students earned points for demonstrating their best effort and for adhering to program rules. Students used verbal praise only and did not use tangible reinforcers. Specifically, the student pages included pictures of stars embedded throughout; and when the pair of students reached a star, the coach praised the reader. In addition, the teacher awarded points randomly to pairs for completing an activity, increasing self-scores in the speed game, and following PALS rules. Students filled in their star charts when they received points and received incentives for each completed point sheet.

Training and Implementation Logistics. Both teachers attended a full-day training session held during the summer to learn the first-grade PALS curriculum. The primary investigator and the project director conducted the training. The con-

tent included (a) a brief review of the research on the efficacy of PALS with at-risk students, including benefits for the higher-functioning students; (b) review of the scope and sequence; and (c) opportunities to observe and practice lessons. Teachers received professional development hours for participating in the training.

RAs implemented the program with the whole class during the literacy block, which consisted of oral language activities, vocabulary development, reading comprehension, spelling, grammar, and writing. RAs who were graduate students in special education conducted the training lessons ($n = 8$) over a 2-week period to introduce the program to the class and model implementation for the teachers. The teachers then conducted all 30-min-long lessons 4 days per week over a 7-week period, for a total of 14 intervention hours. Including the training lessons, students received a total of 18 hr of supplemental instruction.

Treatment Integrity. Teachers and RAs assessed treatment integrity by using component checklists of key intervention components and procedures (e.g., they prompted the onset of each activity, completed each activity, and provided positive feedback). They rated each component on a 3-point Likert-type scale, where 0 = *not in place*, 1 = *partially in place*, and 2 = *in place*. The RA observed one lesson each week. At the end of the lesson, both the RA and the teacher completed the treatment integrity assessment independently. To compute the percentage of implementation from both the RA's perspective and the teacher's perspective for the given lesson, we summed the number of points earned, divided that quantity by the number of points possible, and multiplied the quantity by 100. We computed an overall session integrity score for each intervention leg from the teacher and RA perspectives by averaging the weekly percentages (range: 0%–100%).

DESCRIPTIVE MEASURES

To obtain information on student characteristics (e.g., cognitive ability; behavioral performance from teacher and parent perspectives; and attendance and discipline history) that might mediate treatment outcomes, the researchers completed

four descriptive measures at the beginning of the intervention.

Wechsler Intelligence Scale for Children, Third Edition. A trained RA administered the WISC-III (Wechsler, 1991), described previously, prior to intervention to obtain an estimate of intellectual functioning.

Woodcock-Johnson III Tests of Achievement. The WJ-III (Woodcock et al., 2001) assesses a range of academic competencies, including total achievement; oral language (oral language, oral expression, and listening comprehension); reading (broad reading, basic reading skills, and reading comprehension); mathematics (broad math, math calculation skills, and math reasoning); and written language (broad written language, basic writing skills, and written expression). This nationally normed, standardized achievement battery has excellent psychometric properties, with reliability estimates of .80 and higher. For this study, we reported broad reading cluster scores, which are a combination of letter-word identification, reading fluency, and passage comprehension tests, because this measure most closely approximates the skills covered in the academic interventions (median reliability is 0.93).

Social Skills Rating System—Teacher Version. Teachers completed the Social Skills Rating System—Teacher Version (SSRS-T; Gresham & Elliott, 1990) before the beginning of the intervention to assess students' social skills, problem behavior, and academic competence from the perspective of the teacher. The social-skills scale consists of 30 items equally distributed across three factor analytically derived subscales: cooperation, assertion, and self-control. Teachers rated the frequency and importance of each social-skills item on two 3-point Likert-type scales (frequency: *never* = 0, to *sometimes* = 1, to *very often* = 2; importance: *not important* = 0, to *important* = 1, to *critical* = 2). The problem behavior scale contains 18 items equally distributed across three factor analytically derived subscales: internalizing, externalizing, and hyperactivity. Teachers rate only the frequency of occurrence on these items. Teachers rated the 9 items constituting the academic competence scale on a 5-point Likert-type scale with each point corresponding to various clusters of behaviors of students in the classroom

(1 = lowest 10%, 5 = highest 10%). SSRS-T internal consistency reliabilities range from .82 to .94.

Social Skills Rating System—Parent Version. Parents completed the Social Skills Rating System—Parent Version (SSRS-P; Gresham & Elliott, 1990) before the beginning of the intervention to assess students' social skills and problem behavior from the parents' perspective. The SSRS-P social skills scale consists of 38 items distributed across four factor analytically derived subscales: cooperation, assertion, responsibility, and self-control. Parents rated the frequency and importance of each social-skills item on two 3-point, Likert-type scales, as previously described. The SSRS-P problem behavior scale contains 17 items distributed across three factor analytically derived subscales: internalizing, externalizing, and hyperactivity. Parents rate only the frequency of occurrence on these items. SSRS-P internal consistency reliabilities range from .82 to .94.

OUTCOME MEASURES

The researchers measured academic and behavioral performance twice a week during baseline and once a week during postintervention. Academic performance included a measure of decoding and oral reading fluency. Behavioral performance included academic engagement.

Nonsense Word Fluency and Oral Reading Fluency. We used two subtests of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Kaminski & Good, 1996), Nonsense Word Fluency (NWF) and Oral Reading Fluency (ORF), to monitor student growth on key indicators of early literacy skills. Both subtests contain a series of nationally normed, 1-min probes designed for frequent, repeated assessment of elementary age students. NWF assesses students' knowledge of the alphabetic principle, with an emphasis on letter-sound correspondence and blending letters. It presents the student with randomly ordered nonsense words (e.g., *sig, rav, ov*), and the student either says the individual letter sound or the entire nonsense word. For example, if the stimulus word is *vaj*, the student could say */v/ /a/ /j/* or the word */vaj/* for a correct response. The test yields a score for the number of letter sounds that the student correctly produces per minute.

ORF uses grade-level passages to assess reading fluency in first through sixth grade. It presents the student with a grade-level reading passage, which he or she reads aloud for 1 min. The RA records errors (e.g., omissions, substitutions, or hesitations that exceed 3 s) and computes the number of words that the student reads correctly each minute.

RAs participated in a 4-hr-long training session that involved an introduction to curriculum-based measurement, review of DIBELS materials, and practice with actual students. RAs did not administer the probes until they reached the training criterion of 95% accuracy for administration and scoring over three consecutive probes.

After they reached the training criterion, RAs administered the baseline probes individually to the students in an empty classroom, an office, or an empty hallway to minimize distractions. They gave the students a sticker at the completion of each assessment. We collected interobserver agreement data (IOA) on 25% of the observation sessions, with a mean IOA of 97.76% (range: 91.89–100).

Academic Engagement. Academic engaged time (AET) refers to the length of time that a student spends engaged in the teacher-assigned instructional activity during the 10-min-long observation session. Examples of AET include attending to the material and the task, making appropriate motor responses (e.g., writing, following the rules of a game, and looking at the teacher or the student speaking), asking for assistance in an acceptable manner, and waiting appropriately for teacher instruction. Nonengaged behavior refers to participating in an alternative task. Examples include looking around the room, leaving the desk or walking around during instructional activity, disrupting others, and sleeping. The researchers scored the student as being nonengaged after he or she was not attending for 5 s. The researchers considered the student to be engaged again after 5 s of engaged behavior.

Observers collected data during each 10-min session by using the Multiple Option Observation System for Experimental Studies (MOOSES; Tapp, Wehby, & Ellis, 1995). MOOSES is computer-based software that simultaneously collects frequency and duration of behavior in real time during behavioral observations. The observers

used handheld Dell Axim computers to run the MOOSE software for the observations and measured the percentage of AET by using duration recording procedures. To obtain a percentage, we divided the amount of AET by the total observation length and then multiplied by 100 to obtain a percentage. A second observer coded engaged and nonengaged behaviors simultaneously but independently of the primary observer. We collected IOA data on 25% of the sessions, $M = 97.03\%$ (range: 76.2–100).

Before collecting direct observation data, RAs completed a 5-hr-long formal training program that involved learning the codes, learning how to operate the handheld computer and use the MOOSE program, and practicing in actual classrooms. Practice continued until the RAs attained 90% accuracy across three consecutive sessions with the project director in classrooms.

SOCIAL VALIDITY

We assessed social validity from teacher and student views before the beginning of the intervention and following its completion. Teachers completed the Intervention Rating Profile (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985) for each student. The IRP-15 is a 15-item measure used to assess teachers' perceptions of treatment acceptability. Teachers rated each item (e.g., "I liked the procedures used in this intervention") on a 6-point Likert-type scale ranging from *strongly disagree* = 1 to *strongly agree* = 6. Internal consistency estimates range from .88 to .98.

Students completed the Children's Intervention Rating Profile (CIRP; Martens et al., 1985). The CIRP is a seven-item measure used to assess students' perceptions of treatment acceptability. Students rated each item on a 6-point Likert-type scale ranging from *I do not agree* = 1 to *I agree* = 6. We modified the wording of the CIRP, as in previous studies (e.g., Lane, 1999; Lane et al., 2001) to increase readability for the first-grade students. Internal consistency reliabilities range from .75 to .89.

EXPERIMENTAL DESIGN AND STATISTICAL ANALYSIS

We used a multiple baseline across classrooms to examine intervention outcomes (Kennedy, 2005).

Three students were in the first classroom, and four were in the second classroom. The baseline consisted of regular literacy practices for first-grade students; these practices included preparing students to become independent readers through phonics instruction, oral language activities, shared reading, guided reading, independent reading, daily grammar instruction, story writing, and journaling. Teachers used the Macmillan/McGraw-Hill language arts curriculum. During training, RAs conducted the eight training lessons in each class to model implementation for the teacher and students. The teachers then conducted all lessons during the intervention phase. We collected data only for student participants, with baseline probes collected twice weekly. We administered weekly probes thereafter, with maintenance probes administered 4, 5, and 6 weeks following the completion of the intervention. The nonsense word fluency variable guided phase change decisions. Leg 1 remained in baseline for 3 weeks, and Leg 2 for 10 weeks.

We analyzed the data by using traditional single-case design techniques that included visual inspection techniques to determine stability, level, and trend. Further, we used mean and slope comparisons by phase to examine more subtle changes in student performance. Finally, we computed the percentage of nonoverlapping data points (PND). We examined social validity ratings by using basic descriptive procedures (e.g., total scores).

RESULTS

TREATMENT INTEGRITY

Treatment integrity was slightly higher for Leg 1 from both the teacher and RA perspectives, with respective mean scores of 100% and 88.73%. The level of implementation was slightly lower for Leg 2, with teacher and RA mean scores of 83.04% and 83.63%, respectively. Teacher and RA scores were more closely aligned for Leg 2 than for Leg 1.

ACADEMIC OUTCOMES

Baseline. During the baseline period, students' NWF scores were low and stable (see Table 2 and Figure 1). NWF mean baseline scores ranged between 24.20 (Elmer) and 45.20 (Rose). ORF

TABLE 2

Student Outcomes Over Time: Means, Slopes, and PND

| Student | Phase | NWF | | | ORF | | | AET | | | TI RA | TI Teacher |
|----------------|--------------|------------------|------------------------------|------|------------------|------------------------------|------|------------------|------------------------------|-----|----------|---------------|
| | | M(SD) | Slope (SE _{XY}) | PND | M(SD) | Slope (SE _{XY}) | PND | M(SD) | Slope (SE _{XY}) | PND | | |
| Leg 1 Elmer | Baseline | 24.20 (10.96) | -6.20 (5.67) | | 16.40 (6.80) | -4.20 (1.71) | | 65.96 (10.19) | 4.63 (8.19) | | 88.73% | 100% |
| | Training | 36.50 (19.09) | 27.00 (a) | | 23.50 (3.54) | -5.00 (a) | | 65.35 (0.49) | 0.70 (a) | | | |
| | Intervention | 37.71 (7.67) | 2.54 (5.89) | | 28.67 (8.59) | 1.86 (8.32) | | 63.67 (20.81) | -2.77 (21.84) | | | |
| | Maintenance | 46.00 (29.51) | 24.50 (23.27) | 33% | 52.67 (10.60) | -6.50 (11.84) | 100% | 69.70 (19.22) | -4.75 (26.33) | 33% | | |
| Rose | Baseline | 45.20 (7.29) | 2.60 (6.96) | | 12.60 (2.30) | -0.10 (2.65) | | 83.62 (19.57) | -6.75 (18.94) | | 88.73% | 100% |
| | Training | 55.50 (0.71) | 1.00 (a) | | 14.00 (4.24) | -6.00 (a) | | 93.00 (9.90) | -14.00 (a) | | | |
| | Intervention | 66.57 (9.95) | 3.68 (6.55) | | 32.67 (6.11) | 0.11 (6.68) | | 77.36 (13.73) | 1.19 (14.78) | | | |
| | Maintenance | 82.00 (13.11) | 13.00 (2.45) | 100% | 56.33 (13.65) | -8.50 (15.11) | 100% | 80.17 (18.05) | 12.10 (18.94) | 33% | | |
| Jimmy | Baseline | 31.40 (6.27) | 0.60 (7.16) | | 12.40 (5.46) | -1.80 (5.38) | | 79.06 (17.34) | -1.17 (19.91) | | 88.73% | 100% |
| | Training | 44.50 (4.95) | 7.00 (a) | | 14.00 (4.24) | -6.00 (a) | | 43.10 (30.97) | 43.80 (a) | | | |
| | Intervention | 52.14 (6.15) | 1.50 (5.72) | | 25.67 (10.18) | 3.43 (7.65) | | 83.89 (6.44) | 0.68 (6.86) | | | |
| | Maintenance | 70.67 (8.50) | 8.00 (4.08) | 100% | 47.67 (12.34) | -3.5 (16.74) | 100% | 65.8 (31.80) | 19.50 (35.52) | 0% | | |

| | | | | | | | | | |
|----------------|--------------|------------------|-------------------|------------------|------------------|------------------|------------------|--------|--------|
| Leg 2 James | Baseline | 27.05 (6.44) | 0.66 (5.42) | 13.05 (7.25) | 0.96 (4.95) | 65.69 (19.77) | 1.75 (17.63) | 83.04% | |
| | Training | 38.50 (3.54) | 5.00 (a) | 21.00 (9.90) | -14.00 (a) | 67.75 (25.81) | 36.50 (a) | | 83.63% |
| | Intervention | 46.57 (11.33) | 3.21 (9.80) | 21.33 (8.28) | -0.79 (8.88) | 86.93 (5.98) | -0.42 (6.48) | | |
| | Maintenance | 61.67 (1.53) | 1.50 (0.41) | 32.67 (2.52) | -1.00 (3.27) | 70.70 (9.90) | -7.40 (9.31) | | |
| Alex | Baseline | 31.11 (7.77) | 0.09 (7.98) | 31.63 (9.49) | 0.53 (9.27) | 86.61 (13.07) | 0.60 (13.00) | 0% | |
| | Training | 27.00 (4.24) | 6.00 (a) | 46.00 (2.83) | 4.00 (a) | 76.95 (8.84) | 12.50 (a) | | |
| | Intervention | 39.33 (9.44) | 0.81 (10.37) | 50.80 (3.76) | 0.57 (3.97) | 72.87 (30.29) | 1.74 (33.60) | | |
| | Maintenance | 35.67 (7.09) | -4.5 (7.76) | 61.67 (3.06) | -3.00 (0.82) | 75.50 (33.37) | -5.50 (46.54) | | 33% |
| Haley | Baseline | 29.58 (5.17) | 0.40 (4.78) | 16.11 (8.35) | 0.85 (7.05) | 85.80 (10.26) | 0.94 (9.03) | 33% | |
| | Training | 38.50 (2.12) | -3.00 (a) | 32.00 (9.90) | 14.00 (a) | 91.65 (11.81) | 16.70 (a) | | |
| | Intervention | 40.57 (8.75) | -0.07 (9.59) | 40.83 (8.08) | -0.07 (8.85) | 87.16 (14.62) | -1.30 (15.71) | | |
| | Maintenance | 71.33 (20.53) | 6.00 (27.76) | 53.00 (2.65) | -2.00 (2.45) | 93.00 (6.76) | 2.60 (8.82) | | 33% |
| Kylie | Baseline | 36.58 (6.30) | 0.56 (5.61) | 10.32 (7.02) | 0.21 (7.13) | 73.66 (19.20) | -0.09 (19.75) | 0% | |
| | Training | 60.50 (4.95) | -7.00 (a) | 19.00 (4.24) | 6.00 (a) | 76.50 (25.17) | 35.60 (a) | | |
| | Intervention | 58.43 (3.78) | -0.71 (3.78) | 22.00 (9.22) | -2.21 (8.63) | 89.07 (7.51) | -0.53 (8.13) | | |
| | Maintenance | 45.00 (14.93) | -11.50 (13.47) | 33.00 (13.11) | -5.00 (17.15) | 78.73 (0.06) | -0.05 (0.04) | | 0% |

Note. NWF = Nonsense Word Fluency; ORF = Oral Reading Fluency; AET = Academic Engaged Time; TI = treatment integrity.

^aSE_{XY} not provided since there were only two data points during this phase.

FIGURE 1

Student Performance Over Time: Nonsense Word Fluency

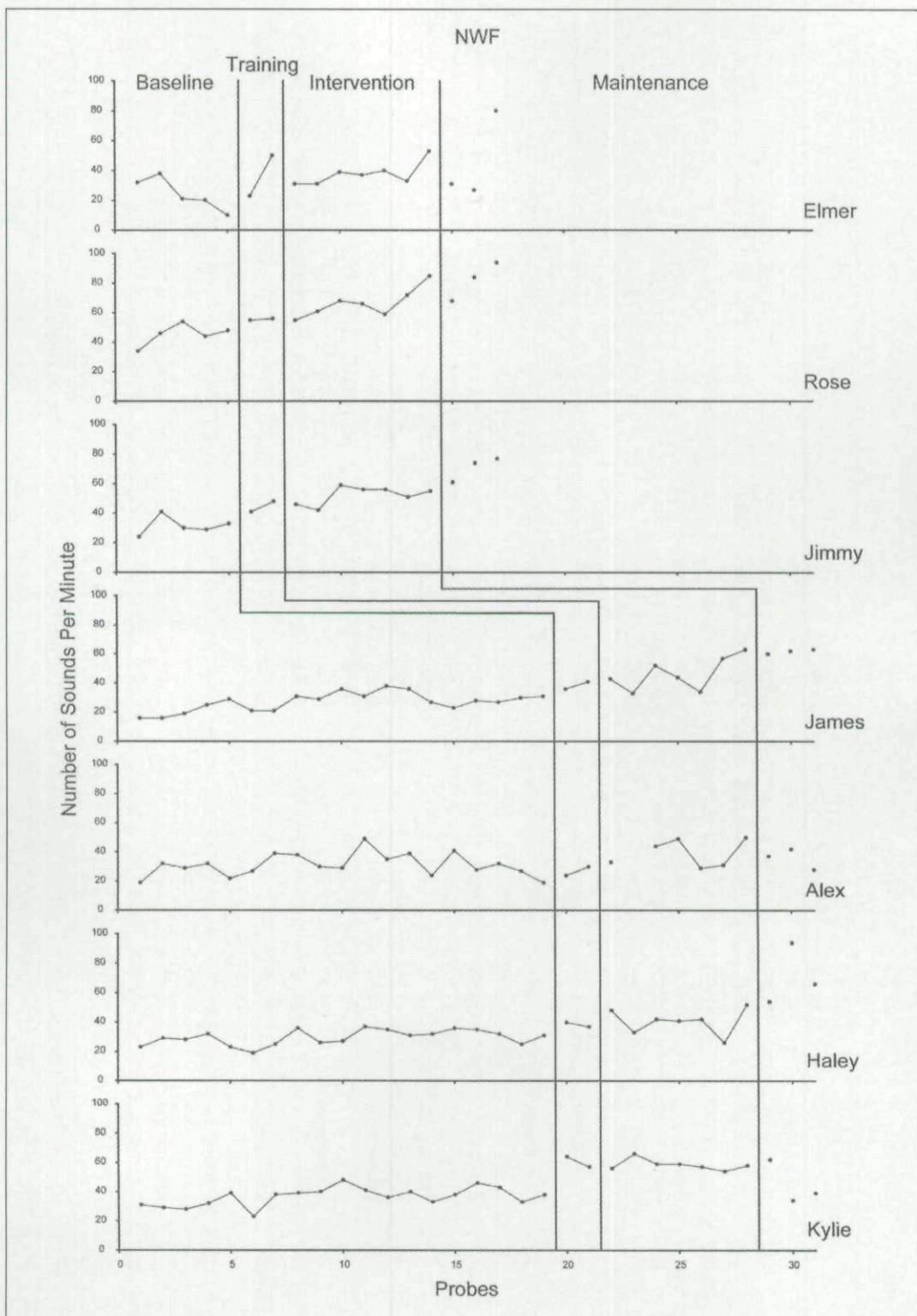


FIGURE 2

Student Performance Over Time: Oral Reading Fluency

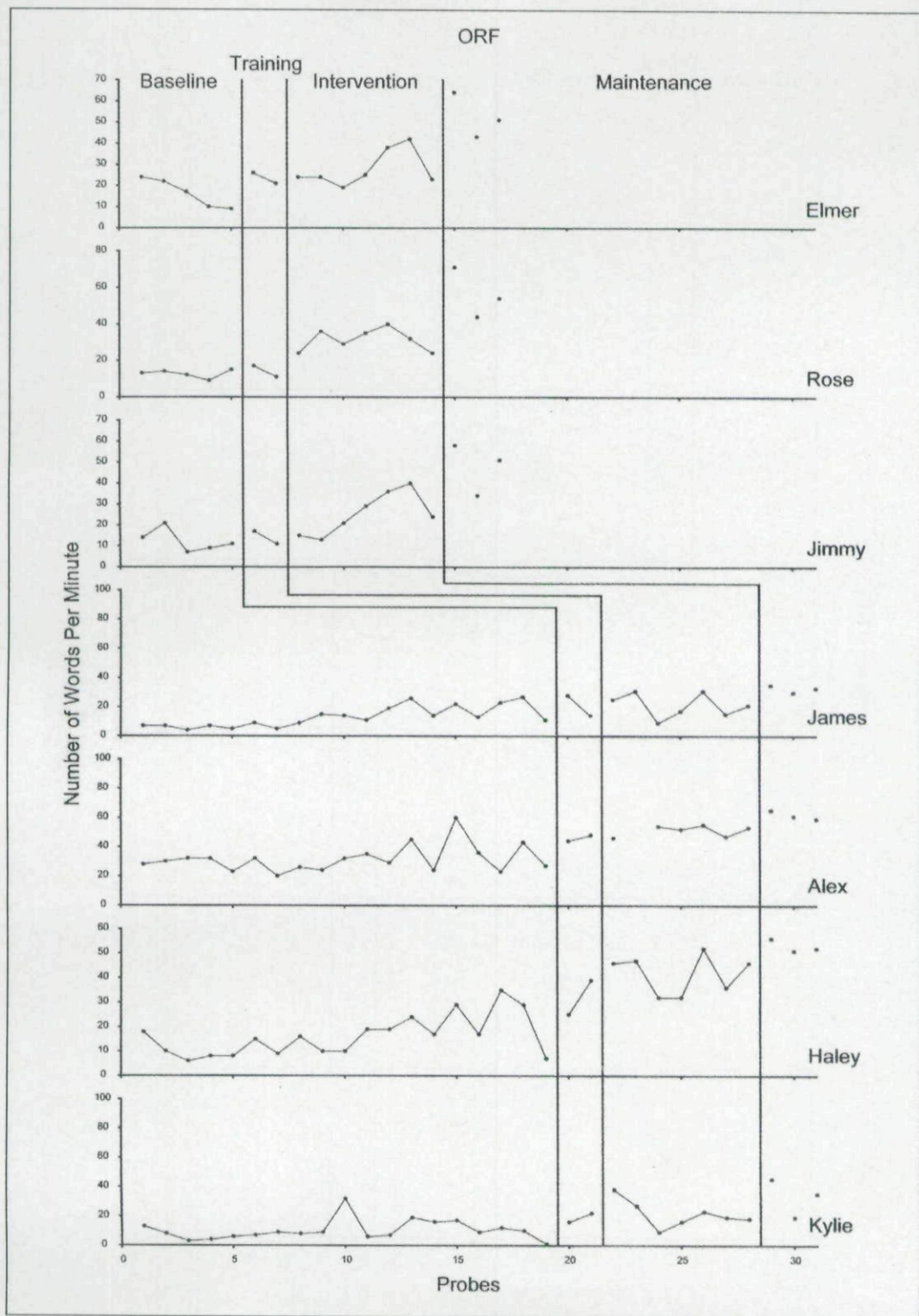
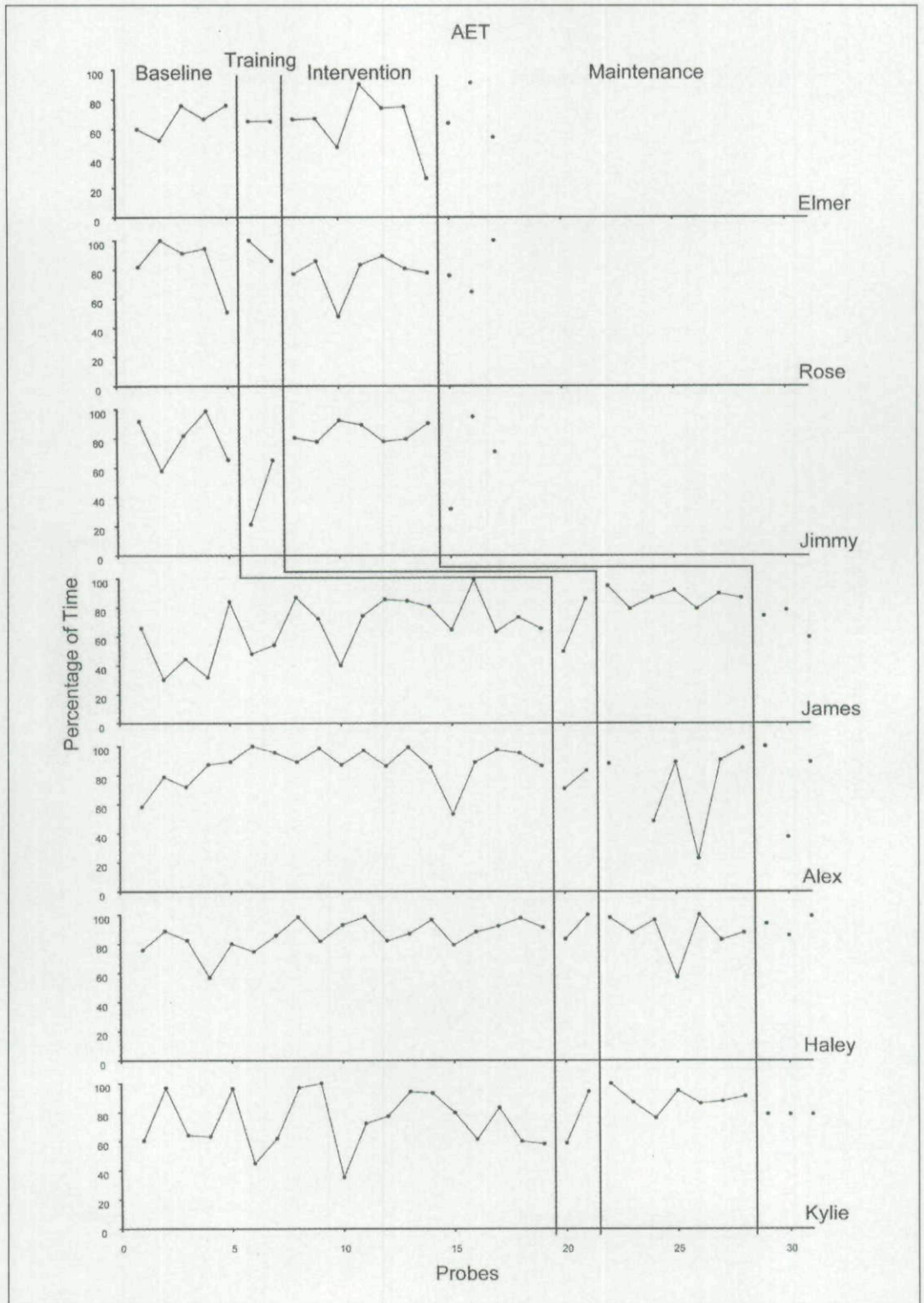


FIGURE 3

Student Performance Over Time: Academic Engagement



scores were also low and relatively stable during baseline, with variability increasing for Alex prior to intervention onset (see Figures 2 and 3).

ORF scores ranged from 10.32 (Kylie) to 31.63 (Alex). ORF slope scores for all students in Leg 1 were slightly decreasing. In contrast, slope scores were slightly increasing for all students in Leg 2, since the students were showing some responsiveness to the general education instruction provided during the 10-week baseline period.

Training. During the 2-week training period, all students except Alex showed increases in NWF mean levels relative to the baseline. Alex did, however, show an increasing trend during the intervention phase. Similarly, all students showed increases in ORF mean levels compared with the baseline.

Intervention. During the intervention phase, all students except Haley and Kylie demonstrated increasing trends in NWF. Rose and James showed the most pronounced trends, with slopes of 3.68 and 3.21, respectively. Haley and Kylie instead showed increased NWF mean levels, with decreased variability for Kylie.

For all seven students, ORF mean levels increased from the baseline phase to the intervention phase. Students' mean levels continued to increase through the maintenance phase. Although some students exhibited variable performance patterns during trainings (e.g., Rose, James, Haley, and Kylie), others showed positive trends (e.g., Elmer and Jimmy). Alex demonstrated an increase in level, with diminished variability relative to the baseline level.

Thus, experimental control was established for all seven students in terms of NWF, as evidenced by changes in either slope (Elmer, Rose, Jimmy, James, and Alex) or level (Haley and Kylie) when the intervention began. Similarly, all students demonstrated increased mean levels during the intervention phase relative to the baseline. With the exception of James and Kylie, experimental control was established between the introduction of the intervention and changes in the students' mean level of performance.

Maintenance. All students showed improvements from baseline levels in NWF and ORF following intervention. Five students showed clear changes in NWF levels during maintenance, with mean scores increasing from a low 1.1-fold for

Alex to a high of 2.28-fold for James. PND scores were 100% for Rose, Jimmy, James, and Haley and 33% for Elmer and Kylie. In addition, a positive trend occurred during the maintenance phase for all students in Leg 1, as well as for James and Haley in Leg 2.

Improvements in ORF were even more pronounced. All seven students showed clear changes in level, with maintenance scores increasing from a low of 1.97-fold for Alex to a high of 4.47-fold for Rose. PND scores comparing baseline and maintenance phases were 100% for all students except Alex and Kylie, whose PND was 33%. Collectively, the findings support improvement in decoding and reading fluency skills.

ACADEMIC ENGAGEMENT

Baseline. With the exception of Elmer (65.96) and James (65.69), all students showed average AET levels (see Figure 3), with scores ranging from 73.66% to 86.61%. Further, standard deviations showed relatively comparable levels of variability, ranging from 10.19 (Elmer) to 19.57 (Rose).

Training. During the 2 weeks of training, all students except Rose showed increasing trends for AET. Elmer, James, Haley, and Kylie exhibited similar mean levels of performance during the training phase.

Intervention. When the researchers introduced the intervention, Jimmy, Rose, James, and Kylie showed decreased variability relative to the baseline phase. Although no functional relationship existed between the introduction of the intervention and changes in level or slope for AET, experimental control was established between the introduction of the intervention and decreases in variability for Jimmy, Rose, James, and Kylie. Experimental control was not established for the other three students.

Maintenance. Students Elmer and Rose in Leg 1 showed virtually no changes in level (PND = 33%), trend, or variability relative to the baseline. Jimmy showed an increase in variability, with one particularly low maintenance score. Three students in Leg 2 (James, Haley, and Kylie) showed increases in mean AET scores, with substantial decreases in variability for all three stu-

TABLE 3

Teacher and Student Social Validity Scores Over Time

| Student | Phase | Social Validity Measure | |
|---------|------------------------|-------------------------|-----------------|
| | | IRP-15 (Total) | CIRP (Total) |
| Leg 1 | Baseline <i>M (SD)</i> | 37.67 (21.08) | 42.00 (0.00) |
| | Post <i>M (SD)</i> | 71.00 (8.54) | 33.67 (6.43) |
| Elmer | Baseline | 25 | 42 |
| | Post | 70 | 29 |
| Rose | Baseline | 26 | 42 |
| | Post | 80 | 41 |
| Jimmy | Baseline | 62 | 42 |
| | Post | 63 | 31 |
| Leg 2 | Baseline <i>M (SD)</i> | 75.00 (0.00) | 42.00 (0.00) |
| | Post <i>M (SD)</i> | 76.00 (0.00) | 36.75 (4.50) |
| James | Baseline | 75 | 42 |
| | Post | 76 | 37 |
| Alex | Baseline | 75 | 42 |
| | Post | 76 | 31 |
| Haley | Baseline | 75 | 42 |
| | Post | 76 | 42 |
| Kylie | Baseline | 75 | 42 |
| | Post | 76 | 37 |

dents. However, PND scores were only 33% for Alex and Haley and 0% for James and Kylie.

SOCIAL VALIDITY

Students rated the intervention as highly socially valid prior to intervention onset, with total scores of 42 for each student (see Table 3). However, with the exception of Haley, students' scores declined following the completion of the intervention, with mean scores of 33.67 ($SD = 6.43$) and 36.75 ($SD = 4.50$), for Legs 1 and 2, respectively.

In contrast, the teachers' social validity rating increased following the completion of the intervention for each student, with dramatic increases for Elmer and Rose, in particular. Teacher IRP-15 scores showed the most dramatic increases for students in Leg 1, with mean scores increasing from 37.67 ($SD = 21.08$) before the beginning of the intervention to 71.00 ($SD = 8.54$) following its completion.

In sum, postintervention CIRP scores were still favorable; however, the intervention goals,

procedures, and outcomes did not meet students' initially high expectations. Quite the opposite was true for teachers. IRP-15 scores indicated that the intervention goals, procedures, and outcomes exceeded teacher expectations to varying degrees for all seven students.

DISCUSSION

Students with E/BD, although most often noted for their social and behavioral problems, have below-average academic performance levels—particularly in reading (Lane, 2004; Nelson, Benner, et al., 2004). Collectively, such deficits place these students at heightened risk for unfavorable outcomes within and beyond the school setting (Wagner & Davis, 2006). Most treatment-outcome studies have focused on social and behavioral concerns (e.g., Mathur, Kavale, Quinn, Forness, & Rutherford, 1996). Only in the last decade have educators placed increased emphasis

on learning how to meet the academic needs of these students (Lane, 2004).

Reading interventions conducted since 1997 suggest that reading interventions, when implemented by university personnel or with the support of school-site staff, have been effective in improving early reading skills of students with (Barton-Arwood et al., 2005; Wehby et al., 2003) and at risk (Lane et al., 2001; Lane et al., 2002; Nelson et al., 2005) for E/BD. Yet, with the call for inclusive programming (Fuchs & Fuchs, 1994; MacMillan, Gresham, & Forness 1996) and the expectation of academic excellence for all students (e.g., No Child Left Behind), it is imperative that researchers and practitioners identify evidence-based practices that general education teachers can implement with limited—or at least reasonable—support. Specifically, is it possible to provide supplemental reading instruction for students at risk for E/BD in the general education setting in the absence of sustained support from either university personnel or school site support staff?

Is it possible to provide supplemental reading instruction for students at risk for E/BD in the general education setting in the absence of sustained support from either university personnel or school site support staff?

This study represents the first attempt to answer this question. The investigation sought to address the question by examining the effectiveness of a supplemental reading intervention, PALS, in improving the early literacy skills of first-grade students at risk for E/BD who also had co-occurring reading deficits. Consistent with a direct-collaborative consultation model (Watson & Robinson, 1996; Witt, Gresham, & Noell, 1996), research staff initially modeled the intervention in the general education classroom by conducting eight training lessons over a 2-week period. The general education teacher then assumed full teaching responsibilities.

Treatment fidelity levels, as measured by RA and teacher perspectives, suggested that teachers

were able to implement the reading intervention with a high level of fidelity in the absence of sustained university or paraprofessional support. Although initial social validity ratings indicated skepticism on the part of the teachers for some students (e.g., Elmer and Rose), postintervention scores suggest that the intervention goals, procedures, and outcomes exceeded teachers' expectations. However, the intervention fell short of meeting students' expectations, with the primary dissatisfaction stemming from having to work with the same partner throughout the process.

Student outcomes were consistent with previous single-case studies of students at risk for E/BD (Lane et al., 2001; Lane et al., 2002). Experimental control was established between the introduction of the intervention and increases in either the level or trend of students' NWF. Further, experimental control was also established with respect to the mean level of oral reading fluency for all students except James and Kylie. Students exhibited lasting increases in early reading skills, thereby confirming our initial hypothesis. However, questions remain regarding the extent to which reading skills were associated with improved classroom behavior—the second objective of this study.

IMPROVED READING SKILLS: ARE THEY ASSOCIATED WITH IMPROVED BEHAVIOR?

The relation between academic intervention and problem behavior continues to be a focal point of interest among researchers studying E/BD (Hinshaw, 1992; Petras et al., 2004). Despite this interest, the body of existing literature does not offer definitive conclusions about the extent to which improved reading skills are associated with collateral effects on behavior for students with or at risk for E/BD. Studies that have used single-case methodology offer partial evidence to suggest that improved reading skills are associated with decreased levels of disruptive behavior exhibited by students at risk for E/BD (Lane et al., 2001; Lane et al., 2002). These studies did not explore whether students then became more engaged in instruction. Specifically, although these students were less disruptive, were they more engaged? Studies have examined both engagement and disruption with students receiving special education

services for E/BD. The studies suggest limited improvement (Wehby et al., 2003) to no improvement (Barton-Arwood et al., 2005) in engagement. Reading studies with at-risk populations have not explored effects on engagement. We hypothesized that if students experienced sustained improvements in their decoding and oral reading skills, students at risk for E/BD would demonstrate higher levels of engagement during literacy instruction, given that their behavior patterns are more amenable to intervention efforts than those of students with E/BD (Kazdin, 1987; Walker et al., 2004).

Yet, consistent with the findings of the Barton-Arwood et al. (2005) study, although student participants in this study gained increased reading fluency, these improvements were not associated with clear changes in AET level or trend. Experimental control was established between the introduction of the intervention and changes in variability for four students. Although educators might hope for changes in level and trend, changes in variability are also beneficial because teachers can more easily predict and manage behavior that students manifest consistently than they can predict and manage inconsistent behavior patterns (Walker et al., 2004). It is also important to recognize that baseline AET levels were not drastically low, ranging from 65.69 (James) to 86.61 (Alex). Only two students (Elmer and James) were below normative levels at intervention onset (Walker & Severson, 1992). Thus, because of a ceiling effect, the lack of formidable increases in AET is not surprising.

LIMITATIONS AND FUTURE DIRECTIONS

Researchers should consider five limitations of this study when interpreting current findings and conducting future studies.

The most serious limitation of this study is that it includes only two legs. Horner et al. (2005) recommend that the design provide a minimum of three demonstrations of experimental effect at three different time points to demonstrate experimental control. Although the present study offers seven demonstrations of experimental effect (three in Leg 1 and four in Leg 2), they do not occur at three different points in time. To more firmly establish internal validity, future

studies should include a multiple baseline study across three legs. Including the third leg was not possible in this study because only two teachers had students who met the inclusion criteria.

The second limitation, which is related to the first limitation, is that only seven students participated in the study. Additional replications are necessary to support the generalizability of the findings for students at risk for E/BD and early reading difficulties. In addition, because these students were all of early elementary age, future studies should establish the efficacy of interventions led by general education teachers with limited support across the grade span. Outcomes may be less positive as the content becomes more differentiated and as students' deficits become more pronounced over time (Lane, 2004; Lane, Pierson, & Givner, 2004).

This study occurred in an inclusive school district in rural middle Tennessee. All general and special education teachers employed by the district were fully certified and dedicated to serving students with exceptionalities in a full-inclusion model. Consequently, it is possible that the teachers in the study were more accustomed to accommodating individual differences in the general education setting than teachers working in noninclusive districts. To establish the external validity of this study, future investigations should establish the efficacy of these procedures in other settings (e.g., urban settings), in districts with less inclusive philosophies, and with older students (Lane, Weisenbach, Little, Phillips, & Wehby, 2006).

This study focused on academic engagement rather than on disruption. The goal was to determine whether enhanced academic performance was associated with higher levels of engagement during literacy instruction. Future studies should perhaps continue to include data on disruption, as did previous studies (e.g., Barton-Arwood et al., 2005; Lane et al., 2001, Lane et al., 2002; Wehby et al., 2003). Decreases in disruption precede improvements in engagement.

Finally, some students registered discontent with having the same partner throughout the study. This discontent may have been one reason for the decline in students' social validity scores. Future investigators may wish to explore variations in procedures, such as altering partners to

enhance procedural acceptability from the student's perspective (Lane & Beebe-Frankenberger, 2004).

CONCLUSIONS AND EDUCATIONAL IMPLICATIONS

As schools seek to serve students in inclusive settings (Fuchs & Fuchs, 1994), meet the call for academic excellence for all students (Fournier, 2002), and meet the needs of students who do not respond to more global intervention efforts such as schoolwide models of prevention (Horner & Sugai, 2000), general education teachers are apt to participate more fully in secondary and tertiary levels of prevention. Although behavioral specialists, reading specialists, and school psychologists often initiate such interventions, these individuals do not have the resources to provide sustained support with intervention implementation (Watson & Robinson, 1996). Thus, it is important to build capacity at the school site so that general education teachers can play a more active—yet feasible—role in intervention implementation. One goal of this study was to help determine the efficacy and feasibility of meeting this challenge.

We contend that this study was both methodologically and pragmatically successful, as defined by a recent request for proposals released by the Institute for Educational Sciences (IES; 2006), which defined "methodologically successful" as a study that

rigorously evaluated the impact of a clearly specified intervention on relevant student outcomes under clearly described conditions using a research design that meets the Institute's What Works Clearinghouse standards . . . whether or not the intervention is found to improve students outcomes relative to the comparison condition. (p. 43)

IES defined "pragmatically successful" as "the rigorous evaluation determined that the intervention has a net positive impact on student outcomes in relation to the program or practice to which it is being compared" (p. 43). Moreover, this study was methodologically successful in that it (a) described participants, selection criteria, and settings in detail that affords replication; (b) defined dependent variables operationally, measured

them repeatedly over time, and collected them with reliability; (c) described the independent variable to allow replication, including treatment fidelity data, and manipulated the independent variable systematically with experimenter control; (d) described baseline conditions in detail; (e) established experimental control of reading outcomes for 7 participants with a design that controls for common threats to internal validity; and (f) assessed social validity from multiple perspectives (Horner et al., 2005). Further, the study was pragmatically successful as previously defined, in that a clear functional relationship existed between the introduction of the intervention and lasting changes in oral reading and nonsense-word reading fluency and in that partial evidence supported a functional relationship between the introduction of the intervention and changes in the variability of academic engagement.

In sum, the results of this study support the effectiveness and feasibility of having general education teachers serve as the primary interventionists to meet the academic needs of students at risk for E/BD in the general education setting without sustained support from other support staff. In addition, the findings also indicated that there may be collateral effects on behavior in terms of academic engagement consistency. As schools attempt to meet the multiple needs of students with and at risk for E/BD, we hope that this study will offer a cost-effective, yet scientifically rigorous, method for supporting elementary students with behavior and learning concerns.

REFERENCES

- Achenbach, T. M. (1991). *The child behavior checklist: Manual for the teacher's report form*. Burlington: University of Vermont, Department of Psychiatry.
- Adams, M. J., Bereiter, C., Carruthers, I., Case, R., Hirshberg, J., McKeough, A., et al. (2000). *Open court reading*. Columbus, OH: SRA/McGraw-Hill.
- Allor, J. H., Fuchs, D., & Mathes, P. (2001). Do students with and without lexical retrieval weaknesses respond differently to instruction? *Journal of Learning Disabilities, 34*, 264-275.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-TR; 5th ed.)*. Washington, DC: Author.

- Anderson, J. A., Kutash, K., & Duchnowski, A. J. (2001). A comparison of the academic progress of students with EBD and students with LD. *Journal of Emotional and Behavioral Disorders, 9*, 106–115.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavioral Analysis, 1*, 91–97.
- Barton-Arwood, S. M., Wehby, J. H., & Falk, K. B. (2005). Reading instruction for elementary-age students with emotional and behavioral disorders: Academic and behavioral outcomes. *Exceptional Children, 72*, 7–27.
- Bullis M., & Walker, H. M. (1994). *Comprehensive school-based systems for troubled youth*. Eugene: University of Oregon, Center on Human Development.
- Drummond, T. (1994). *The Student Risk Screening Scale (SRSS)*. Grants Pass, OR: Josephine County Mental Health Program.
- Drummond, T., Eddy, J. M., & Reid, J. B. (1998a). *Follow-up study #3; risk screening scale: Prediction of negative outcomes by 10th grade from 2nd grade screening*. Unpublished technical report. Eugene, OR: Oregon Social Learning Center.
- Drummond, T., Eddy, J. M., & Reid, J. B. (1998b). *Follow-up study #4; risk screening scale: Prediction of negative outcomes in two longitudinal samples*. Unpublished technical report. Eugene, OR: Oregon Social Learning Center.
- Drummond, T., Eddy, J. M., Reid, J. B., & Bank, L. (1994, November). *The student risk screening scale: A brief teacher screening instrument for conduct disorder*. Paper presented at the Fourth Annual Prevention Conference, Washington, DC.
- Engelmann, S., Engelmann, O., & Davis, K. L. (1997). *Horizons learning to read fast track A-B*. Columbus, OH: SRA/McGraw-Hill.
- Foorman, B. R., Francis, D. J., Shaywitz, S. E., Shaywitz, B. A., & Fletcher, J. M. (1997). The case for early reading intervention. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 243–264). Mahwah, NJ: Erlbaum.
- Fournier, R. (2002, January 9). Education overhaul signed. *Riverside Press Enterprise*, pp. A1, A9.
- Fuchs, D., & Fuchs, L. S. (1994). Inclusive schools movement and the radicalization of special education reform. *Exceptional Children, 60*, 294–309.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Martinez, E. (2002). Preliminary evidence on the social standing of students with learning disabilities in PALS and No-PALS classrooms. *Learning Disabilities Research and Practice, 17*, 205–215.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*, 174–206.
- Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., et al. (2001). Is reading important in reading readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology, 93*, 251–267.
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.
- Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111*, 127–155.
- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S. L., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practices in special education. *Exceptional Children, 71*, 165–179.
- Horner, R. H., & Sugai, G. (2000). School-wide behavior support: An emerging initiative. *Journal of Positive Behavior Interventions, 2*, 231–232.
- Institute for Educational Sciences (IES). (2006). *Request for proposals: Serious behavioral competition*. Retrieved from <http://lies.ed.gov/>
- Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25*, 215–227.
- Kazdin, A. (1987). Treatment of antisocial behavior in children: Current status and future directions. *Psychological Bulletin, 102*, 187–203.
- Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston: Allyn & Bacon.
- Lane, K. L. (1999). Young students at risk for antisocial behavior: The utility of academic and social skills interventions. *Journal of Emotional and Behavioral Disorders, 7*, 211–223.
- Lane, K. L. (2004). Academic instruction and tutoring interventions for students with emotional/behavioral disorders: 1990 to present. In R. B. Rutherford, M. M. Quinn, & S. R. Mathur (Eds.), *Handbook of research in emotional and behavioral disorders* (pp. 462–486). New York: Guilford Press.
- Lane, K. L., Barton-Arwood, S. M., Rogers, L. A., & Robertson, E. J. (in press). Literacy interventions for students with and at-risk for emotional or behavioral disorders: 1997 to present. In J. C., Crockett, M. M. Gerber, & T. J. Landrum (Eds.), *Achieving the radical*

reform of special education: Essays in honor of James M. Kauffman. Mahwah, NJ: Lawrence Erlbaum.

Lane, K. L., & Beebe-Frankenberger, M. E. (2004). *School-based interventions: The tools you need to succeed.* Boston: Allyn & Bacon.

Lane, K. L., O'Shaughnessy, T., Lambros, K. M., Gresham, F. M., & Beebe-Frankenberger, M. E. (2001). The efficacy of phonological awareness training with first-grade students who have behavior problems and reading difficulties. *Journal of Emotional and Behavioral Disorders, 9*, 219-231.

Lane, K. L., Pierson, M., & Givner, C. C. (2004). Secondary teachers' views on social competence: Skills essential for success. *Journal of Special Education, 38*, 174-186.

Lane, K. L., Wehby, J. H., Menzies, H. M., Gregg, R. M., Doukas, G. L., & Munton, S. M. (2002). Early literacy instruction for first-grade students at-risk for antisocial behavior. *Education and Treatment of Children, 25*, 438-458.

Lane, K. L., Wehby, J. H., Phillips, A., Weisenbach, J., Little, M. A., & Merwin, M. (2006). *The utility of academic and behavioral interventions for early elementary students with reading and behavior problems.* Manuscript submitted for publication.

Lane, K. L., Weisenbach, J. L., Little, M. A., Phillips, A., & Wehby, J. (2006). Illustrations of function-based interventions implemented by general education teachers: Building capacity at the school site. *Education and Treatment of Children, 29*, 549-671.

MacMillan, D., Gresham, F., & Forness, S. (1996). Full inclusion: An empirical perspective. *Behavioral Disorders, 21*, 145-159.

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 and Practice, 16*, 191-198.

Mathes, P. G., Howard, J. K., Allen, S. H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly, 33*, 62-94.

Mathur, S. R., Kavale, K. A., Quinn, M. M., Forness, S. R., & Rutherford, R. B. (1996). Social skills interventions with students with emotional and behavioral problems: A quantitative synthesis of single-subject research. *Behavioral Disorders, 23*, 193-201.

Mattison, R. E., Spitznagel, E. L., & Felix, B. C. (1998). Enrollment predictors of the special education outcome for students with SED. *Behavioral Disorders, 23*, 243-256.

Nelson, J. R., Benner, G. J., Lane, K. L., & Smith, B. W. (2004). An investigation of the academic achievement of K-12 students with emotional and behavioral disorders in public school settings. *Exceptional Children, 71*, 59-73.

Nelson, J. R., Cooper, C., & Gonzalez, J. (2004). *Stepping stones to literacy.* Longmont, CO: Sopris West.

Nelson, J. R., Stage, S. A., Epstein, M. H., & Pierce, C. D. (2005). Effects of a pre-reading intervention on the literacy and social skills of children. *Exceptional Children, 72*, 29-45.

O'Connor, R. E., Notari-Syversen, A., & Vadasky, P. F. (1998). Ladders to literacy: The effects of teacher-led phonological activities for kindergarten children with and without disabilities. *Exceptional Children, 63*, 117-130.

O'Shaughnessy, T., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. (2002). Students with or at risk for learning and emotional-behavioral difficulties: An integrated system of prevention and intervention. In K. L. Lane, F. M. Gresham, & T. E. O'Shaughnessy (Eds.), *Intervention for children with or at risk for emotional and behavioral disorder* (pp. 3-17). Boston: Allyn and Bacon.

Petras, H., Schaeffer, C. M., Ialongo, N., Hubbard, S., Muthen, B., Lambert, S. F., et al. (2004). When the course of aggressive behavior in childhood does not predict antisocial outcomes in adolescence and young adulthood: An examination of potential explanatory variables. *Development and Psychopathology, 16*, 919-941.

Sattler, J. (1991). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.

Schneider, B. H. (1991). A comparison of skill-building and desensitization strategies for intervention with aggressive children. *Aggressive Behavior, 17*, 301-311.

Severson, H. H., & Walker, H. M. (2002). Pro-active approaches for identifying children at-risk for socio-behavioral problems. In K. L. Lane, F. M. Gresham, and T. E. O'Shaughnessy (Eds.), *Interventions for children with or at risk for emotional and behavioral disorders* (pp. 33-54). Boston: Allyn & Bacon.

Shefelbine, J. (1998). *Phonics chapter books 1-6: Teachers guide.* New York: Scholastic.

Tapp, J., Wehby, J., & Ellis, D. (1995). A multiple option observation system for experimental studies: MOOSES. *Behavior Research Methods, Instruments and Computers, 27*, 25-31.

Torgesen, J. K., & Bryant, B. R. (1994). *Phonological awareness training for reading.* Austin, TX: PRO-ED.

- Umbreit, J., Lane, K. L., & Dejud, C. (2004). Improving classroom behavior by modifying task difficulty: Effects of increasing the difficulty of too-easy tasks. *Journal of Positive Behavior Interventions, 6*, 13–20.
- Van Acker, R., & Talbott, E. (1999). The school context and risk for aggression: Implications for school-based prevention and intervention efforts. *Preventing School Failure, 44*, 12–20.
- Wagner, M., & Davis, M. (2006). How are we preparing students with emotional disturbances for the transition to young adulthood? Findings from the National Longitudinal Transition Study-2. *Journal of Emotional and Behavioral Disorders, 14*, 86–98.
- Wagner, M., Kutash, K., Duchnowski, A. J., & Epstein, M. H. (2005). The Special Education Elementary Longitudinal Study (SEELS) and the National Longitudinal Transition Study (NLTSS2): Study designs and implications for children and youth with emotional disturbance. *Journal of Emotional and Behavioral Disorders, 13*, 25–41.
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2004). *Antisocial behavior in schools: Evidence-based practices* (2nd ed.). Belmont, CA: Wadsworth.
- Walker, H. M., & Severson, H. (1992). *Systematic screening for behavior disorders: Technical manual*. Longmont, CO: Sopris West.
- Walker, H. M., Severson, H., Nicholson, F., Kehle, T., Jenson, W. R., & Clark, E. (1994). Replication of the Systematic Screening for Behavior Disorders (SSBD) procedure for the identification of at-risk children. *Journal of Emotional and Behavioral Disorders, 2*, 66–77.
- Walker, H. M., Severson, H., Todis, B. J., Block-Pedego, A. E., Williams, G. J., Haring, N. G., et al. (1990). Systematic Screening for Behavior Disorders (SSBD): Further validation, replication, and normative data. *RASE: Remedial and Special Education, 11*, 32–46.
- Watson, T. S., & Robinson, S. (1996). Direct behavioral consultation: An alternative to traditional behavioral consultation. *School Psychology Quarterly, 11*, 267–278.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—3rd Ed.* San Antonio, TX: Harcourt Brace Jovanovich.
- Wehby, J. H., Falk, K. B., Barton-Arwood, S., Lane, K. L., & Cooley, C. (2003). The impact of comprehensive reading instruction on the academic and social behavior of students with emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 11*, 225–238.
- Witt, J., Gresham, F. M., & Noell, G. (1996). What's behavioral about behavior consultation? *Journal of Educational and Psychological Consultation, 7*, 327–344.
- Woodcock, R. W., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.
- Zigmond, N. (2006). Twenty-four months after high school: Paths taken by youth diagnosed with severe emotional and behavioral disorders. *Journal of Emotional and Behavioral Disorders, 14*, 99–107.

ABOUT THE AUTHORS

KATHLEEN LYNNE LANE (CEC TN Federation), Assistant Professor; **M. ANNETTE LITTLE**, Project Director; **JENNY REDDING-RHODES**, Research Assistant; **ANDREA PHILLIPS**, Research Assistant; and **MEGAN T. WELSH** (CEC TN Federation), Research Assistant, Department of Special Education, Peabody College, Vanderbilt University, Nashville, Tennessee.

Address correspondence to Kathleen Lane, Department of Special Education, Peabody College, Box 328, Vanderbilt University, Nashville, TN 37203-5721 (e-mail: Kathleen.lane@vanderbilt.edu).

This research was supported by OSEP Field Initiated Grant H324C030044B and NICHD Grant P30HD15052.

Manuscript received October 2006; accepted February 2007.

Copyright of Exceptional Children is the property of Council for Exceptional Children and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.