The purpose of this study was to examine the effects of an intervention involving reciprocal peer tutoring and self-graphing of reading data on the disruptive behavior, active responding, and reading fluency of students with emotional or behavioral disorders (EBD). Four middle school students with EBD participated in this research. Results indicated that during the intervention phase students’ disruptive behavior decreased and active responding increased. Furthermore, during the intervention phase each student made progress on words read correctly per minute. This is represented by each student attaining the goal of typical reading growth, compared with only one student achieving this goal during the baseline phase. Limitations of the study and implications for practice and future research are discussed.

Students with emotional or behavioral disorders (EBD) exhibit both academic and behavioral deficits, and these “dual deficits” can make it difficult for practitioners to provide effective instruction (Kauffman, 2005). Students’ lack of motivation, classroom disruptions, and aggressive behavior often inhibit teachers’ attempts to provide instruction (Wehby, Symons, Canale, & Go, 1998). Not surprisingly, as students with EBD advance in school, they often do not progress academically at a rate equal to that of their peers, including those with learning disabilities (Anderson, Kutash, & Duchnowski, 2001). By the time students with EBD reach middle school, their teachers are often facing students with severe academic challenges as well as established negative behaviors. Students’ academic problems often require intensive remediation efforts to improve basic skills, particularly in reading (Lane, 2004). Greenbaum et al. (1996) found that the percentage of children with EBD who were reading below grade level increased from 54% to 85% across the study’s 7-year span. Nelson, Benner, Lane, and Smith (2004) reported that 83% of their study’s sample of children with EBD scored below the norm group on a standardized measure of reading skill.

Researchers have suggested a variety of means, including academic and behavioral interventions, to address the learning and behavior problems of students with EBD (Kern & Clemens, 2007; Lewis, Hudson, Richter, & Johnson, 2004). However, these interventions are often implemented in classrooms characterized by coercive interchanges between teachers and students, and these contexts might contribute to the maintenance of learning and behavior problems, not to mention variations in intervention implementation integrity and teacher attention. Researchers have suggested that aggressive behavior patterns increase the likelihood that children will develop negative relationships with their teachers (Ladd & Burgess, 1999), and recent research has indicated that problematic relationships between teachers and students with behavior problems in kindergarten are associated with academic and behavioral problems through eighth grade (Hamre & Pianta, 2001). These problematic relationships with teachers may contribute to students with EBD having low rates of positive teacher attention, such as academic talk, teacher praise, and opportunities to respond to academic requests (e.g., Van Acker, Grant, & Henry, 1996; Wehby, Symons, & Shores, 1995). In light of the less-than-desirable classroom environments available to students with EBD, peer-mediated interventions, such as peer tutoring, which is designed...
to increase response opportunities (Falk & Wehby, 2001), and self-management interventions, such as self-graphing (Bloom, Hursh, Wienke, & Wold, 1992), might be particularly efficacious. This is because students are taught to implement intervention components and monitor their own progress while teachers can provide individual attention to students who need extra support.

One peer-mediated intervention with a strong evidence base is Peer-Assisted Learning Strategies (PALS; D. Fuchs et al., 2001). Procedures for PALS include structured activities, frequent interaction and feedback between tutors and tutees, and role changes between tutor and tutee (D. Fuchs et al., 2001). Numerous studies have validated the effectiveness of PALS on the academic measurements of students with and without disabilities. Mathes, Torgesen, and Allor (2001) found that low-achieving students benefited significantly from the PALS intervention in the area of reading achievement ($ES = .59$) and average-achieving ($ES = .34$) and low-achieving ($ES = .41$) students generally scored higher than students in the contrast condition. Furthermore, low-achieving students who participated in the PALS intervention scored at or near the average-achieving and high-achieving students on phonological awareness at the end of the study.

In a study conducted by L. S. Fuchs, Fuchs, and Karns (2001), data indicated that kindergarten students’ growth in mathematics in 10 PALS classrooms exceeded the growth of students in 10 no-treatment classrooms ($ES = .24$). Stronger effects were noted for students of medium initial achievement ($ES = .53$) and low achievement ($ES = .46$), and for students with disabilities ($ES = .41$). Finally, Rohrbeck, Ginsburg-Block, Fantuzzo, and Miller (2003) conducted a meta-analysis on the effects of peer-assisted learning interventions, such as PALS, on learning outcomes of elementary school students. Results indicated an effect size of .59 across 81 group-design intervention studies. The authors noted that the greatest intervention effects occurred with learners with the greatest need (i.e., high poverty, urban, young children, low achievement).

A number of studies have also examined the effects of peer tutoring on students with EBD (e.g., DuPaul, Ervin, Hook, & McGoey, 1998; Falk & Wehby, 2001; Maheady, Harper, & Mallette, 2003). Of particular interest to the current investigation are studies that have examined the effects of peer tutoring on middle school students with EBD during academic instruction. Franca, Kerr, Reitz, and Lambert (1990) found that peer tutoring had a positive effect on students’ correct responses, attitudes toward mathematics, and peer interactions. Spencer, Scruggs, and Mastropieri (2003) found that students in the peer tutoring condition scored higher on social studies tests and demonstrated increased on-task behavior when compared with traditional instruction. Locke and Fuchs (1995) reported improved on-task behavior and increased positive peer-to-peer comments during the implementation of peer tutoring.

Although PALS is a peer-mediated intervention that provides increased response opportunities for students, self-evaluation is a form of self-management whereby the student compares her or his progress to established criteria (Mercer & Mercer, 2005). Self-graphing is a component of self-evaluation in which students graph their own data, which allows a visual representation of their performance over time (Gunter, Miller, Venn, Thomas, & House, 2002). L. S. Fuchs and Fuchs (1986), in a meta-analysis of formative evaluation procedures for students with disabilities, found that these procedures resulted in significant increases in academic achievement. Moreover, effect sizes were enhanced when data from formative evaluations were graphed, perhaps because graphing of performance data provided more frequent performance feedback to students. Bloom et al. (1992) found that behavioral interventions were more effective when paired with self-monitoring, and DiGangi, Maag, and Rutherford (1991) demonstrated that the benefits of self-monitoring were enhanced when students with disabilities also graphed the results of their progress. Finally, Gunter et al. (2002) discussed how teachers can use computer-based spreadsheets to allow students to graph their progress. Such self-graphing also had positive effects on students’ academic and social behavior.

The study’s purpose was to examine the effects of an intervention consisting of reciprocal peer tutoring and self-graphing of reading data on the disruptive behavior, active responding, and reading fluency of students with EBD. The study was designed to address several issues regarding the relationship between reading fluency and disruptive behavior. For example, EBD students with reading problems have been well-documented (Lane, 1999; Levy & Vaughn, 2002; Vaughn, Levy, Coleman, & Bos, 2002), and evidence-based practices in reading instruction, particularly for middle and high school students with EBD, are sorely needed. Further, older students with reading difficulties tend to have problems reading text fluently, and this skill is critical to comprehending text (Strong, Wehby, Falk, & Lane, 2004).

In addition, whereas Spencer et al. (2003) reported positive effects of peer tutoring on on-task behavior, the current study used active student responding as a dependent measure. Percentage of time on-task, though an important measure of students’ classroom engagement, is not a direct measure of students’ active engagement. Active student responding, on the other hand, is a measure of a student’s observable response to academic instruction and is a direct measure of a student’s responding during academic instruction (Heward, 1994). Finally, because disruptive classroom behavior is a common characteristic of students with EBD (e.g., Kauffman, 2005), the current study examined the effects of reciprocal peer tutoring on students’ disruptions during reading instruction. The hypotheses for this study were that students with EBD would have the following:

1. increased active responding,
2. decreased disruptive behaviors, and
3. increased reading fluency during reciprocal peer tutoring with self-graphing rather than during typical classroom instruction.
METHOD

Setting and Participants

The setting for this study was a self-contained classroom for students with EBD in a middle school in a medium-size southeastern U.S. city. There were 1,507 students in Grades 6 through 8, and approximately 50% were White, 41% Black, 5% Hispanic, and 4% of other racial/ethnic origin; approximately 25% of the students in the school received a free or reduced-priced lunch. A female teacher and a female teaching assistant taught students in the classroom. The teacher was a 38-year-old White woman with 16 years of teaching experience and a master’s degree in education. Her assistant was a 41-year-old White woman with 4 years’ experience working with students with EBD. The classroom, located in a trailer next to the school, consisted of four parallel rows of student desks, two teacher desks, one study carrel, two small-group instruction areas (round tables and chairs), and three computers. In addition to the 4 student participants in this study, during reading instruction there were typically 4 other students (all boys; 3 Black, 1 Hispanic) present in the classroom. The teacher used a point-and-level system to reward and consequate student behavior, and this system remained in place throughout the study.

Participants were selected by the classroom teacher after she indicated that all of the students in her class could benefit from the intervention. Consent forms were sent home to all students in the class, and five (of nine) were returned. One student, (Jacob), moved during the 9th week of data collection before he received the intervention and was excluded from the sample. The final sample consisted of four students (see Table 1 for characteristics of participating students), all of whom were identified as having emotional disturbance (ED) according to Virginia guidelines, which are consistent with the federal definition. Latasha also received special education services for speech and language impairments; her teacher noted that she rarely volunteered information during class discussions and typically would work quietly on tasks rather than interact with her peers, even during free-time activities. The teacher reported that Mel, who received Risperdal for anxiety and hallucinations, often needed redirection to complete tasks and follow directions during instructional times; he often talked to peers, made jokes, and was out of his seat. The teacher reported that Dee was off task frequently, didn’t complete assignments, verbally threatened other students, and would argue with adults (and peers) when redirected. Finally, James was described by his teacher as attention-seeking and often off task, needing frequent redirection to maintain attention to tasks. James was often out of his seat and completed few assignments.

Design and Procedures

For three reasons, a multiple-baseline-across-subjects design (Kazdin, 1982) was used to examine the effect of the intervention on the dependent variables. First, the nature of the intervention (implemented with dyads) allowed for a sequential introduction of the intervention across participants. This allowed for an analysis of change with each participant. Second, it was hypothesized that the effects of the intervention would be difficult to reverse. Third, after exposure to the intervention, it would have been unethical to withdraw the intervention components.

Student Pairs. To overcome possible difficulties experienced by EBD students in peer-mediated instructional interventions (Sutherland, Wehby, & Gunter, 2000), the teacher used several data sources to group her students into peer tutoring dyads. Because students with EBD are characterized by social skills deficits that impact their ability to participate in peer-mediated interventions (Sutherland, Wehby, & Gunter, 2000), data from the Interpersonal Competence Scale for Teachers (ICS-T; Cairns, Leung, Gest, & Cairns, 1995) were collected on participating students to assist the teacher in pairing students during the intervention phase of the study. Researchers suggested using social competence data, in addition to academic achievement data, to group students into peer-mediated interventions (Sutherland, Wehby, & Gunter, 2000), data from the Interpersonal Competence Scale for Teachers (ICS-T; Cairns, Leung, Gest, & Cairns, 1995) were collected on participating students to assist the teacher in pairing students during the intervention phase of the study. Researchers suggested using social competence data, in addition to academic achievement data, to group students into peer-mediated interventions (Slavin, 1994; Sutherland, Wehby, & Gunter, 2000). The teacher used ICS-T data, her knowledge of her students and

<table>
<thead>
<tr>
<th>Student/ethnicity</th>
<th>Age (yrs-mos)</th>
<th>Special education eligibility</th>
<th>Full-scale IQ</th>
<th>Instructional reading level</th>
<th>Medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latasha/African American</td>
<td>12-7</td>
<td>ED/SLI</td>
<td>81</td>
<td>Grade 7</td>
<td></td>
</tr>
<tr>
<td>Mel/African American</td>
<td>12-9</td>
<td>ED</td>
<td>101</td>
<td>Grade 5</td>
<td>Risperdal</td>
</tr>
<tr>
<td>Dee/African American</td>
<td>13-6</td>
<td>ED</td>
<td>82</td>
<td>Grade 6</td>
<td></td>
</tr>
<tr>
<td>James/Caucasian</td>
<td>11-4</td>
<td>ED</td>
<td>106</td>
<td>Grade 6</td>
<td></td>
</tr>
</tbody>
</table>

Note. ED = emotional disturbance; SLI = speech and language impairment.
the classroom social network, and reading assessment data (see section on Curriculum-Based Measurement Data) to create dyads that had a greater likelihood of working together during peer tutoring. The use of multiple sources of data to make grouping decisions created dyads that had a greater likelihood of working well together, as researchers have indicated that randomly assigning students with EBD to peer-mediated interventions is not effective (Pomplun, 1997). The ICS-T is an 18-item questionnaire consisting of 7-point Likert scales that the teacher completes for each participant prior to data collection. The ICS-T yields composite scores on popularity, Olympian (good at sports), affiliative, academic, aggressive, and internalizing components. Reliability coefficients are typical of similar self-ratings and other ratings on the factors assessed (i.e., .50-.70). The ICS-T has convergent validity with direct observation, student record (i.e., grades, discipline reports), and peer-nomination measures (Cairns & Cairns, 1994; Cairns et al., 1995), and it has predictive validity over an 8-year period for adult adjustment, early school dropout, and teenage parenthood (Cairns & Cairns, 1994). Table 2 summarizes the ICS-T factor scores for each student participant. These data, along with the teacher’s knowledge of the classroom social structure and the instructional reading level of each student, were used to pair students.

Baseline. During the baseline phase, no changes in the students’ academic instruction were made. Typical instruction involved students using spelling books or language arts worksheets. The teacher and teacher assistant would assist individual students with their work. Direct observation data were collected on disruptive behavior and active responding of target students, and weekly curriculum-based measurements (CBM) were collected on target students.

Intervention. The first author trained the teacher on the implementation of the intervention and provided ongoing technical support during the intervention. PALS was the model for the peer-tutoring intervention, and laminated copies of instructions for each of the three activities (i.e., partner reading, paragraph shrinking, prediction relay) to be used were provided to the teacher. The teacher gave the instructions to each peer-tutoring dyad at the beginning of each session. Students were paired in accordance with the approach previously discussed.

Latasha and Mel formed the first dyad. They began the intervention on the 19th day of data collection while the other students in the class received their typical instruction. The intervention was implemented with the first dyad after some stability in disruptive behavior was noted, particularly for Mel. Latasha’s disruptive data remained stable throughout the baseline phase. Similarly, on the 31st day of data collection, Dee and a peer began the peer-tutoring intervention after Dee had 4 days of stable disruptive data. Finally, on the 41st day of data collection, James and a peer, in addition to the remaining students in the classroom, began the peer-tutoring intervention. Although a downward trend was noted in James’ disruptive behavior, the teacher was anxious to implement the intervention classwide. On the day prior to the implementation of the peer-tutoring intervention, the teacher would meet with the respective dyad and go over the procedures for the intervention. During the course of the study, she also provided feedback to the students on the implementation of the peer-tutoring intervention.

During the peer-tutoring sessions, two students would sit at a table across from each other. One student (higher performing) would begin reading from children’s literature (e.g., James and the Giant Peach, Charlie and the Chocolate Factory) that had been selected for each group by the teacher. Like PALS, the sessions consisted of three activities. First, during partner reading, the higher-performing student in the dyad would read aloud for 5 min; after 5 min, the lower-performing student would reread the same text passage. During partner reading, the tutor would follow the error-correction procedures outlined by Fuchs et al. (2001).

Following partner reading, the dyads would begin the paragraph shrinking activity. First, the higher-performing student in the dyad would read orally, stopping at the end of each paragraph to identify the main idea. Tutors used questions from the

| Table 2 |
| Interpersonal Competence Scale–Teacher Scores for Study Participants |

<table>
<thead>
<tr>
<th>Student</th>
<th>Popularity</th>
<th>Olympian</th>
<th>Affiliative</th>
<th>Academic competence</th>
<th>Aggressive</th>
<th>Internalizing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latasha</td>
<td>4.0</td>
<td>4.7</td>
<td>0.5</td>
<td>4.0</td>
<td>1.3</td>
<td>6.0</td>
</tr>
<tr>
<td>Mel</td>
<td>4.0</td>
<td>4.7</td>
<td>4.0</td>
<td>4.0</td>
<td>3.7</td>
<td>3.7</td>
</tr>
<tr>
<td>Dee</td>
<td>4.3</td>
<td>5.3</td>
<td>4.5</td>
<td>4.5</td>
<td>4.3</td>
<td>3.3</td>
</tr>
<tr>
<td>James</td>
<td>5.3</td>
<td>4.7</td>
<td>6.5</td>
<td>5.0</td>
<td>4.7</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Note. Interpersonal Competence Scale for Teachers (Cairns et al., 1995); 1 = never, 7 = always.
laminated card, including “who or what the paragraph is mainly about” and “the most important thing about the who or what” to guide the tutee’s identification of main ideas. The reader then attempted to summarize the main idea in 10 words or fewer. If the main idea was not correctly identified, error correction procedures were followed according to D. Fuchs et al. (2001). After 5 min of this activity, the students switched roles.

Finally, the students participated in the prediction relay activity. First, the higher-performing reader made a prediction about what would happen in the next half page of text. Next the reader read the half page aloud while the tutor identified and corrected reading errors. The reader then confirmed or disconfirmed his or her prediction. Next, the reader summarized the main idea of the text read. The tutor would then judge whether the prediction was correct. Finally, the procedures were repeated with the roles reversed—the tutor became the reader on the next half page of text. The entire peer tutoring session lasted approximately 20 min daily.

**Self-Graphing.** During the intervention phase, teachers taught students to self-graph their words read correctly per minute and errors per minute from their weekly CBM assessments. Following procedures outlined in Gunter et al. (2002) the research assistant (RA) created a folder for each target student on a computer in the classroom. An Excel spreadsheet for each student included baseline CBM data entered and graphed by the RA. The RA taught the students to enter CBM data on the spreadsheet and graph their effort on the embedded graph. Students graphed their data after each weekly CBM assessment.

**Treatment Fidelity.** Treatment fidelity data for the implementation of each component of the intervention were collected daily. Data included partner reading, paragraph shrinking, prediction relay, and self-graphing of CBM data. The RA recorded on a checklist the occurrence or nonoccurrence of each component for each dyad during the intervention phase.

**Follow-Up.** Direct observation data were collected on three separate occasions (Days 63, 68, and 80) to measure both student behavior and the teachers’ continuation of the intervention; CBM data were collected on one occasion (Week 15). Our data collection ended during Week 11 (Day 48) so the follow-up data collection gave an indication of the teacher’s continued implementation, as well as maintenance of any treatment effects.

**Measures**

Before data collection began, the teacher identified when she taught reading to her students. Data were collected daily during this time period, from 10:00 until 10:45 A.M. Curriculum-based measurements were conducted on Thursday of each week for project participants; if a participant was absent, the CBM was conducted on Friday.

Observer training included the primary observer, an undergraduate psychology student, and the secondary data collector, a doctoral student in education. The observers memorized the behavioral codes of disruptive behavior and active responding, which included both examples and nonexamples (see Note). Next, the two observers practiced data collection in a separate classroom using the coding procedures until 90% agreement was achieved for each direct observation variable. During this training, the observers would meet with the first author following practice observations to discuss agreements and disagreements; examples and nonexamples of disruptive behavior and active responding were added to the coding manual that the observers continued to memorize. The first author also trained both observers to conduct CBM procedures during separate training sessions.

Prior to data collection, the primary observer spent time in the classroom to familiarize the students with her presence. The primary observer used a paper-and-pencil data sheet to record a frequency count of disruptive behavior and a duration measure of active responding for each participant during the first 20 min of each lesson. During interobserver agreement checks, the primary observer and the second observer sat at the rear of the class and did not interact with the teacher, teacher assistant, students, or each other during the observation sessions; interobserver agreement checks were also conducted for CBM data collection.

To acquire a sample of participants’ behavior across the daily 20-min observations, students’ disruptive behavior and active responding were recorded during 1-min intervals. The order of participant observation was randomly assigned for each observation session at the beginning of the study. During each 1-min interval the observer (a) recorded the occurrence of disruptive behavior and (b) documented the duration of active responding of the target student using a stopwatch (with 5 s to record active responding at the end of each 1-min interval). Thus, at the end of each 1-min interval, the observer would have noted the frequency of disruptive behavior as well as the number of seconds the target student was actively responding. During the next 1-min interval, the next target student was observed, and so forth. At the end of each 20-min observation session, each student participant was observed for a total of 5 min.

**Disruptive Behavior.** Disruptive behavior was recorded using a frequency count when the target student performed a behavior that interrupted, or had the potential to interrupt, the instruction in the classroom during each 1-min interval. Examples of disruptive behaviors included students (a) calling out when the teacher’s expectation was for them to quietly raise a hand, (b) getting out of their seats without permission, (c) throwing paper, and (d) tapping their pencils on the desk. Nonexamples included students raising their hands, asking their partner to try again, and staring at an object.

**Active Responding.** Active student responding was defined as observable behaviors made by students following specific instructional stimuli. Reading words orally, answering questions orally, and writing words are all examples of active
student responding. Nonexamples include looking at the teacher, looking at a book, and drawing or doodling. During each 1-min interval, a stopwatch started when the target student engaged in active responding and stopped when active responding ended. This continued until the beep on the headphone indicated the end of the 1-min interval, at which time the total amount of time (in seconds) spent responding was noted on the data collection sheet. At the end of each session, the total number of seconds of active responding recorded for each participant was divided by 300 (5 min) to determine the percentage of time spent actively responding.

Curriculum-Based Measurement. Prior to the weekly administration of CBM, an RA assessed the reading level of each participant using procedures outlined by Hosp and Hosp (2003). The RA asked the teacher to estimate the reading ability of each participant by grade level, and then three different grade-level equivalent passages of at least 200 words from the Qualitative Reading Inventory (Leslie & Caldwell, 2000) were identified for each student. The student read each passage aloud for 1 min, and the RA recorded words read correctly and incorrectly. If the median score of these three passages was between 70 and 100 words correct per minute, with six or fewer errors, the student’s instructional level was identified (Fuchs & Deno, 1982). If the student had fewer than 70 words read per minute, or more than six errors, three more passages from the next lower grade level were identified and administered. This process continued until each student’s instructional level was identified.

Once the appropriate reading level was identified for each student, CBMs were collected weekly on each participant using procedures described by Shinn (1989) in Hosp and Hosp (2003). Passages taken from Spargo (1989) were identified for each participant one grade level above that identified by the reading assessments. Student participants read a novel passage aloud for 1 min each Thursday, and the RA recorded words read correctly and incorrectly. If the participant was not available to read the passage on Thursday, the RA administered the reading probe on Friday of the same week.

Interobserver Agreement

During 19% of the daily observation sessions (across phases), interobserver agreement (IOA) was assessed for the occurrence or nonoccurrence of disruptive behavior and percentage of active responding. On 39% of the weekly CBM assessments (across phases) IOA was assessed for words read correctly and errors per minute. Interobserver agreement was measured when the two observers collected data simultaneously, and an IOA estimate was calculated. The IOA estimate for disruptive behavior was calculated by dividing the number of agreements by agreements plus disagreements and multiplying by 100% (for each of the IOA sessions). The IOA estimate for active responding was calculated by dividing the smaller percentage of time spent actively responding by the larger percentage and multiplying by 100%. The IOA estimate for words read correctly and errors per minute was calculated by dividing the smaller number of words read correctly, and errors, by the larger number of words read correctly, and errors, respectively, and multiplying by 100%. Mean agreement estimates of the dependent variables were 87.6% for disruptive behavior (range = 75.2%–100% per session), 86.1% for active responding (range = 69%–100% per session), 99.4% for words read correctly per minute (range = 93.5%–100% per session), and 94.7% for errors made per minute (range = 80%–100% per session).

Student Satisfaction Surveys. Because students with EBD are often not motivated to participate in academic instruction, we were interested in their satisfaction with the different components of the intervention. The first author developed a 6-item student satisfaction survey to assess the students’ perception of the intervention. At the end of the intervention phase, the RA asked each student the following:

1. What did you like about partner reading?
2. What didn’t you like about partner reading?
3. What did you like about graphing your reading data?
4. What didn’t you like about graphing your reading data?

Students were also asked how often they would like to continue partner reading (i.e., every day, 2–3 times per week, once a week, or never). Finally, each student was asked to rate on a scale of 1 to 10 (1 = not at all, 10 = very much) their likes and dislikes of the regular reading instruction and partner reading.

Treatment Acceptability Rating Form–Revised. We were also interested in the teacher’s rating of the acceptability of the intervention because this would affect her use of the intervention in the future. The teacher filled out a modified version of the Treatment Acceptability Rating Form–Revised (TARF-R; Reimers & Wacker, 1988) at the end of the intervention phase. The purpose of the TARF-R was to determine how acceptable and feasible the teacher judged the intervention to be within the ongoing classroom curricula and schedule. The TARF-R consisted of 20 questions in three categories:

1. teacher willingness to implement the intervention
2. teacher expected effectiveness of the intervention
3. teacher disadvantages of implementing the intervention

The questions were listed in random order so that the three categories could not be ascertained.

RESULTS

Treatment Fidelity

Treatment fidelity data were collected for each dyad to assess the implementation of the peer tutoring intervention (N = 86). Data indicated that partner reading was implemented on 62 of
86 opportunities (72.1%). Reasons for no implementation included one of the students in a dyad being absent or suspended, student refusal to participate (one occasion), and teacher decision not to implement. When partner reading was implemented, paragraph shrinking was implemented 87.1% of the time, and the prediction relay was implemented 88.7% of the time. Treatment fidelity data during follow-up indicated that the peer tutoring intervention was not implemented during any of the three observations. Students self-graphed their CBM data during the intervention phase 100% of the time.

Disruptive Behavior

Table 3 summarizes the means and standard deviations for disruptive behavior by phase. Figure 1 depicts the mean rate of disruptive behavior per session across students. Latasha’s data remained stable throughout all phases. Mel’s data indicated a great deal of overlap, and despite a slight mean increase in disruptive behavior, his data stabilized by the end of the intervention phase. Dee and James had the highest rate of disruptive behavior during baseline, and although there was much data overlap during the intervention phase, decreasing trends in the data were noted. By the end of the intervention phase, both students’ disruptive behavior data had stabilized.

Active Responding

Table 3 summarizes the means and standard deviations for percentage of time actively responding by phase. Figure 2 depicts the mean percentage of time spent actively responding per session across students. Increasing trends in data were noted for all students during the baseline phase, and this increasing trend continued during the intervention phase. Although phase means and percentage of time means indicate increased active responding during the intervention phase, much data overlap was also noted across phases.

Curriculum-Based Measures

**Words Read Correctly Per Minute.** Figure 3 depicts the words read correctly per minute across students. Each student in each phase of the study had a goal of words read correctly. This goal was based on earlier research by Deno, Fuchs, Marston, and Shin (2001), which indicated that students with disabilities who received effective reading instruction increased their words read correctly per minute at a rate of 1.39 per week. Thus, for each student, the first data point in the baseline phase served as the beginning of the growth line, and the number of subsequent weeks in the respective phase was multiplied by 1.39 to establish the goal at the end of the phase; this procedure was repeated to establish the goal for the intervention phase, with the last data point in the baseline phase serving as the beginning of the growth line.

Table 4 provides the means, standard deviations, slopes, goals, and final CBM data for words read correctly per minute for each student during each phase of the study. Data indicate that during the baseline phase, one student, Dee, exceeded his identified goal on the final CBM. During the intervention phase,
all four students exceeded their identified goal on the final CBM, and for each student the final CBM during the intervention phase exceeded the final CBM during the baseline phase. Follow-up data indicated that at Week 15 all students read fewer words correctly per minute than they did during the final CBM at the end of the intervention, indicating that the effects did not maintain. Slope data indicated that two students, Latasha and Mel, had negative linear trends in growth, and one student, Dee, had no discernible growth during the baseline phase. Data during the intervention phase indicated that these three students had positive linear trends in words read correctly per minute, while James had a negative trend during intervention. Only two data points, both above his targeted goal, were available to compute James’ slope.

Errors Per Minute. Figure 4 depicts the errors per minute across students. Visual inspection of the data indicates that three of the four students had decreasing trends in errors per minute from baseline to intervention. Latasha’s mean rate of errors per minute was stable from baseline ($M = 1.5, SD = 0.6$) to inter-
vention \( (M = 1.5, SD = 1.0) \). Mel’s mean rate of errors per minute decreased from baseline \( (M = 4.2, SD = 1.3) \) to intervention \( (M = 1.2, SD = 0.8) \), as did Dee’s \( (M = 5.7, SD = 2.4 \) to \( M = 3.3, SD = 1.5) \) and James’s \( (M = 2.9, SD = 1.4 \) to \( M = 0.5, SD = 0.7) \), respectively.

**Student Satisfaction**

In response to the question “What did you like about partner reading?” student responses indicated that they liked “stories,” “correcting errors and learning new words,” “getting better at reading,” and “reading.” Responses to the question “What did you like about self-graphing?” included “working on PC,” “liked it,” “the touch mouse,” and “it was fun.” In response to the question “What didn’t you like about partner reading?” students responded “5 minutes was too short,” “having to read what the other person read,” “doing it every day,” and “everything.” No students indicated that they disliked the self-graphing procedure. Two students indicated that they would like to continue partner reading once a week, and two students indicated that they would like to continue partner reading 2 to 3 times a week. Finally, on a scale of 1 to 10 (1 = like not at all, 10 = like very much), the average student satisfaction was 8.5.

---

**FIGURE 2.** Percentage of time spent actively responding per session.
much), students indicated a preference for partner reading ($M = 9.8, SD = 0.5$) over typical reading instruction ($M = 8.5, SD = 1.0$).

**TARF-R**

The teacher completed the TARF-R at the end of the intervention phase. High mean scores (range 1–7) on the willingness to implement ($M = 6.8, SD = 0.5$) and expected effectiveness factors ($M = 5.0, SD = 1.7$) suggest that the teacher would be willing to implement the intervention and to expect that it would be effective for students. A lower mean score (range 1–7) on the perceived disadvantages factor ($M = 1.7, SD = 0.8$) suggests that the teacher perceived few adverse consequences as a result of the intervention.

**DISCUSSION**

This study examined the effects of an intervention on reciprocal peer tutoring and self-graphing of reading data on the dis-
ruptive behavior, active responding, and reading fluency of middle school students with EBD. Results indicated that during the intervention phase students’ disruptive behavior decreased and their active responding increased, in comparison to data from the baseline phase. During the intervention phase, each student made progress on words read correctly per minute, represented by their attainment of a goal representing typical reading growth. This compared with only one student achieving this goal during the baseline phase. Decreases in errors per minute were noted for three of the four students during the intervention phase. Students indicated that they liked both the peer tutoring and self-graphing components of the intervention, and the teacher indicated that she was willing to implement the intervention and expected it to be effective with her students. However, three follow-up direct observations indicated that the intervention was not continued in the classroom, and final CBM data indicated that students were reading fewer words per minute at Week 15 than they were at Week 11, when the intervention ended.

The intervention appeared to have some positive effects on students’ disruptive behavior. Two students demonstrated little change in their rate of disruptions (one student had few disruptions during baseline), and two students demonstrated marked decreases in disruptions from baseline to intervention. Moreover, on days during the intervention phase when the intervention was not implemented, three of the four students indicated increases in disruptive behavior. Interestingly, the rate of disruptive behavior during baseline corresponded closely to the internalizing, aggressive, popularity, and affiliative factors of the ICS-T. Latasha was rated highest on the internalizing factor, and she had the lowest rate of disruptive behavior, with only one occurrence during the study. Meanwhile, Dee and James had the highest ratings on the aggressive, popularity, and affiliative factors of the ICS-T, and each of these students had high rates of disruptive behavior during the baseline phase. In turn, the intervention appeared to have the greatest effect on the disruptive behavior of Dee and James. The peer tutoring intervention might have had effects, such as increases in positive social interactions, on the more withdrawn student participants (Latasha and Mel). However, that was not measured. At a minimum, it does not appear that the intervention had a detrimental effect on Latasha’s disruptive behavior. Although Mel’s disruptive behavior increased slightly (from 0.7 to 0.8 per session) during the intervention phase, on days that peer tutoring was not implemented, it increased to 1.2 per session.

For the two students who did show improvements, several factors may explain the positive effect of the intervention on their disruptive behavior. First, increasing trends in active responding, with much variability, were noted for each of the four students. Active student responding is a measure of students’ observable responses and is typically incompatible with inappropriate classroom behavior. Thus, if students were actively responding more often, they might have been less likely to disrupt class. This assertion is supported by active responding data from the days during the intervention phase when peer tutoring was not implemented. Each student’s percentage of time spent actively responding was lower on these days than it was on days when peer tutoring was implemented. This corresponded with increases in disruptive behavior on the same days.

Researchers (e.g., Gunter & Denny, 1998; Lewis et al., 2004; Steinberg & Knitzer, 1992; Wehby et al., 1998) have documented and lamented the lack of effective instructional practices in classrooms for students with EBD. Teachers often focus on classroom management to the detriment of effective academic instruction, resulting in academic instruction similar to that present during the baseline phase of this study. In light of instructional procedures—namely independent practice—used during the baseline phase of this study, it is not surprising that active responding increased during the intervention phase. Active student responding is typically topographically incompat-

<table>
<thead>
<tr>
<th>Student</th>
<th>M (SD)</th>
<th>Slope</th>
<th>Goal</th>
<th>Final CBM</th>
<th>M (SD)</th>
<th>Slope</th>
<th>Goal</th>
<th>Final CBM</th>
<th>CBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latasha</td>
<td>84.5 (22.2)</td>
<td>−12.3</td>
<td>121.2</td>
<td>67</td>
<td>83.3 (8.9)</td>
<td>2.2</td>
<td>75.4</td>
<td>88</td>
<td>82</td>
</tr>
<tr>
<td>Mel</td>
<td>68.6 (12.4)</td>
<td>−3.2</td>
<td>77.6</td>
<td>56</td>
<td>84.8 (15.2)</td>
<td>1.3</td>
<td>64.4</td>
<td>95</td>
<td>75</td>
</tr>
<tr>
<td>Dee</td>
<td>67.5 (16.9)</td>
<td>0.2</td>
<td>62.0</td>
<td>67</td>
<td>68.8 (4.4)</td>
<td>3.3</td>
<td>72.6</td>
<td>73</td>
<td>68</td>
</tr>
<tr>
<td>James</td>
<td>92.6 (16.2)</td>
<td>2.7</td>
<td>111.8</td>
<td>105</td>
<td>118.5 (3.5)</td>
<td>−5.0</td>
<td>107.8</td>
<td>116</td>
<td>106</td>
</tr>
</tbody>
</table>

Note. CBM = curriculum-based measurement.
ible with disruptive behavior. By focusing on instructional procedures that result in increased active responding, teachers might spend less time responding to disruptive behavior.

Another explanation for the effect of the intervention on disruptive behavior might be partially explained by the CBM data. While Greenwood, Horton, and Utley (2002) indicated that independent practice, similar to academic interactions characterized by both the baseline and follow-up phases of this study, was one of the best instructional practices for promoting academic engagement, it may not have as powerful an effect on academic achievement as active responding (Heward, 1994) or, as this study suggests, disruptive behavior. Curriculum-based measurement data in this study indicated that only one student, Dee, reached his goal of words read correctly per minute by the end of the baseline phase, whereas all four students reached their goals by the end of the intervention phase. In addition, three of the four students had decreasing trends and decreased means of errors per minute from baseline to intervention. Therefore, not only were students actively responding more during the intervention phase but they were also reading more words per minute.

![Figure 4: Errors per minute.](image-url)
with fewer errors. From a negative reinforcement perspective (Gunter & Coutinho, 1997), academic interactions during the intervention phase may have been less aversive to students as they were receiving increased opportunities to respond, with increased words read correctly and fewer errors. From a transactional perspective (Sutherland & Morgan, 2003), academic success during the intervention phase might have resulted in increased active responding and fewer disruptions, leading to even greater academic success. Greenwood (1996) demonstrated that student engagement mediates the relationship between academic instruction and school outcomes. In the current study, the intervention, representing academic instruction, might have led to increased engagement (represented by fewer disruptions and increased active responding), resulting in increased academic achievement (represented by CBM data).

A final explanation for the decrease in disruptive behavior might be the information from the student satisfaction surveys. Students clearly preferred the intervention components to their typical instruction, and they might have been more motivated to actively participate in the academic process as a result of this preference. While Latasha responded to the question “What didn’t you like about the partner reading?” with “Everything,” she also noted that she “liked reading” and rated both her normal reading instruction and the partner reading as a “10” on the student satisfaction survey. This disconnect might be explained by her discomfort at being forced by the intervention to interact with a peer, which coincides with her teacher’s rating of her shy and withdrawn behavior. At the same time, the results from the TARF-R indicated that the teacher was willing to implement the intervention and thought it would be effective, yet follow-up data indicated that at least on the three occasions that data were collected, the intervention was no longer being implemented. Findings from this study support research suggesting that behavior change for teachers of students with EBD is difficult to maintain. Gunter, Denny, Jack, Shores, and Nelson (1993) found that a teacher’s rate of praise decreased to near baseline levels when the use of prompts was withdrawn. Sutherland, Wehby, and Copeland (2000) demonstrated that when feedback on rates of behavior-specific praise was withdrawn, a teacher’s rate of behavior-specific praise decreased. Sutherland, Alder, and Gunter (2003) found that when feedback on rates of opportunities to respond was withdrawn, a teacher’s rate of providing opportunities decreased. Heward (1994) pointed out the importance of investigating both what must be done to facilitate teachers’ adoption and continued implementation of effective methods. An ongoing prompt or evaluation system might be needed to maintain a minimum rate of effective teaching practices.

A particularly disturbing aspect of the follow-up data was the number of words read correctly by the students at Week 15, 4 weeks after the intervention stopped. All students read fewer words correctly at this assessment than they did at the final assessment at Week 11, and two of the four students had more errors per minute. Though only one point in time, this indicates a lack of growth, which is particularly troubling given the positive growth indicated by each student during the intervention phase.

**Limitations**

Several limitations must be kept in mind as one attempts to interpret the findings of this study. First, this study was conducted in a single classroom for students with EBD. Thus, teacher factors, individual student factors, and classroom contextual factors, to name a few, must be considered before generalizing these findings to other classrooms of students with EBD. Second, because one of the foci of this study was active responding during reading, the definition of active responding used in this study does not allow an accurate analysis of the effects of the intervention on active responding represented by overt reading. For example, there was a great deal of data-point overlap across phases for active responding. However, during the baseline and follow-up phases, active responding was typically represented by students writing on worksheets, while during the intervention phase active responding was typically represented by students reading words aloud. Increases in words read correctly and decreases in errors per minute during the intervention provide some evidence of the effect of the intervention on reading fluency; at the same time, an operational definition of active reading responding would have provided greater evidence of the effect of the intervention on this dependent variable. Third, the self-graphing component of the intervention was implemented by the research assistant. Self-graphing is simple and cost-effective to implement, yet any results attributed to it must be tempered by the fact that it was not implemented by the teacher. Also, because the intervention was delivered as a package, it is not possible to determine what effects, if any, self-graphing had on the dependent variables beyond the effects of peer tutoring. Fourth, IOA checks for the direct observation variables were conducted on 19% of the sessions. Although these checks occurred across phases and agreement estimates across variables were acceptable, agreement checks on at least 20% of total observation sessions are generally recommended (Kennedy, 2005). Finally, treatment fidelity data indicated that the intervention was not implemented daily throughout the intervention phase. As a result, the effects of the intervention may have been diluted, providing a less than accurate picture of its effect on the dependent variables. Data collected during the days when the intervention was not implemented provide some evidence of its effectiveness, given the decreases in active responding and increases in disruptions on these days.

**Summary and Directions for Future Research**

Although the combined intervention of reciprocal peer tutoring and self-graphing had some positive effects on students’ reading fluency and classroom behavior, the delivery of the intervention does not allow for an evaluation of the effect of the individual intervention components on the dependent variables. It is particularly important for future studies to investigate the
Effect that self-graphing of academic data by students with EBD has on various academic and behavioral measures to determine the strength of this intervention component.

Implementation of the intervention, or lack thereof, raises important issues for both researchers and practitioners. Baker, Gersten, Dimino, and Griffiths (2004) investigated factors that influenced the sustainability of elementary teachers’ use of PALS in mathematics. These researchers found that among other factors, high-quality professional development and ongoing support during the research study influenced teachers’ use of PALS. The current study relied on the teacher to implement the intervention after the training was conducted, with little ongoing support. Future research should provide ongoing support to teachers to encourage the use of treatment packages, which provide a more efficacious evaluation of treatment effects. Singh and Oswald (2004) provided a methodology for practitioners and clinicians to use to develop evidence-based treatment options in educational and treatment settings. Preservice and in-service training in this model might help teachers become more effective consumers of research, facilitating the sustained use of evidence-based practices.

Some of the challenges the teacher faced in implementing peer tutoring in this setting were unavoidable. For example, in special education classrooms there are limited options available for pairing students due to smaller class sizes. In addition, the social skills deficits of students with EBD can further limit the available options for student pairs, as some students may not interact well with others. This can threaten the efficacy of the intervention (Sutherland, Wehby, & Gunter, 2000). In the current study, the teacher’s use of ICS-T ratings, CBM assessment data, and personal knowledge of the students appeared to result in dyads that used the peer-mediated intervention effectively. Other circumstances in the classroom often made implementation impossible. For example, once the students had been working with a partner for a period of time, they were very reluctant to exchange partners. This created several situations where the intervention could not be implemented, such as when a partner was absent, suspended, or refused to participate. Coleman and Vaughn (2000) noted the emotional variability cited by teachers of students with EBD, particularly as it relates to providing effective reading instruction. This emotional variability impacts a teacher’s ability to effectively implement any reading instruction, and particularly an intervention like peer tutoring where social interaction is required.

Finally, results from the student satisfaction survey provide several recommendations for both research and practice. First, the students indicated that they would prefer peer tutoring that was implemented 1 to 3 times weekly, rather than every day. Because one of the primary research questions was to examine the effect of peer tutoring on students’ active responding and disruptive behavior using daily direct observational data, more frequent implementation of the intervention was desired. Researchers have suggested implementation of peer tutoring 3 times a week (D. Fuchs et al., 2001), and the student interviews from this study support that notion. More difficult to overcome, however, was the indication that the students did not like the repeated readings component of the intervention. Although repeated readings are an important component of fluency training (Mastropieri, Leinart, & Scruggs, 1999), characteristics of students with EBD, such as a lack of motivation, indicate an area of future research to identify the most effective manner in which to include repeated readings in reading instruction.

In summary, reciprocal peer tutoring and self-graphing appear to have promise for improving both the classroom behavior and reading fluency of students with EBD. At the same time, logistical challenges inherent in implementing this intervention package with students with EBD in special classrooms provide hurdles for both practitioners and researchers. Results from the current study provide some rationale for further work in this area in both practice and research, as it appears the benefits for our students, with sustainable implementation, could be significant.

About the Authors
KEVIN S. SUTHERLAND, PhD, is an associate professor of special education and disability policy at Virginia Commonwealth University. His current interests include emotional or behavioral disorders, teacher–student interactions, and intervention research. ANGELA SNYDER, PhD, is research faculty in the Department of Special Education and Disability Policy at Virginia Commonwealth University. Her interests include educational research methodology and educational statistics, as well as emotional and behavioral disorders in children and youth. Address: Kevin S. Sutherland, Virginia Commonwealth University, 1015 W. Main St., PO Box 842020, Richmond, VA 23284; e-mail: kssuther@vcu.edu

Note
Contact first author for detailed definitions.

Authors’ Notes
1. Support for this research was provided in part by grants from the Great Cities Universities (R215K020369) and the Metropolitan Educational Research Consortium at Virginia Commonwealth University.
2. The authors would like to thank Tom Gumpel, Phil Gunter, Kathleen Lane, and Paul Morgan for their comments on an earlier draft of this manuscript, as well as the teacher and students who participated in this study.

References


Mathes, P. G., Torgesen, J. K., & Allor, J. H. (2001). The effects of...


