Helping Students With Emotional and Behavioral Disorders Solve Mathematics Word Problems

PETER ALTER

Saint Mary's College of California, Moraga, CA, USA

The author presents a strategy for helping students with emotional and behavioral disorders become more proficient at solving math word problems. Math word problems require students to go beyond simple computation in mathematics (e.g., adding, subtracting, multiplying, and dividing) and use higher level reasoning that includes recognizing relevant information, disregarding irrelevant information, and choosing the correct arithmetic operation. By teaching a multistep problem-solving strategy and reinforcing completion of each step through a token economy, 4 students improved in their math problem-solving ability and their overall on-task behavior during the work sessions. In addition, 1 student transitioned from a teacher-driven token economy to use the token economy as a self-monitoring system. The author describes this strategy and the accompanying token economy system and provides the results of the intervention for 4 students.

Keywords: emotional and behavioral disorders, math word problems, strategy

Students with emotional and behavioral disorders (EBD) fall below national levels in all areas of academics, but most notably in reading and mathematics (Anderson, Kutash, & Duchnowski, 2001; Nelson, Benner, Lane & Smith, 2004). The need to remediate the academic deficiencies of students with EBD by using evidence-based instructional strategies continues to be an imperative in the field (Lane, 2004; Scott, Nelson, & Liaupsin, 2001). One specific area to be targeted for remediation is in the area of math word problems. In Hodge, Riccomini, Buford, and Herbst’s (2006) comprehensive review of studies investigating instructional interventions in mathematics for students with EBD only one study addressed problem-solving skills for students with EBD. Mathematics problem solving is defined as the presentation of a novel problem that requires the student to determine an appropriate course of action for attaining a goal before implementing a strategy to address the problem (Bottge, 2001; Fuchs et al., 2003; Polya, 1971). Mathematics problem solving is most commonly presented in the form of math word problems (Bottge, 2001).

Moving beyond these skill deficits are the potential performance deficits that students may demonstrate. That is, students may be unwilling to solve math word problems. Discrete task completion reinforcement may explain why students willingly solve a number of problems that involve a single step (e.g., adding two numbers or subtracting two numbers) but avoid word problems that require multiple steps to complete a single problem. For example, when completing a word problem, the computation is only one part in the multiple steps required to arrive at the correct solution. When students receive reinforcement only for deriving the correct final response, numerous additional opportunities for reinforcement are missed. Such a lean reinforcement schedule is likely to reduce the probability of students engaging in future problem-solving behavior.

The author presents a strategy for helping students with emotional and behavioral disorders become more proficient at solving math word problems. Math word problems require students to go beyond simple computation in mathematics (e.g., adding, subtracting, multiplying, and dividing) and use higher level reasoning that includes recognizing relevant information, disregarding irrelevant information, and choosing the correct arithmetic operation. By teaching a multistep problem-solving strategy and reinforcing completion of each step through a token economy, 4 students improved in their math problem-solving ability and their overall on-task behavior during the work sessions. In addition, 1 student transitioned from a teacher-driven token economy to use the token economy as a self-monitoring system. The author describes this strategy and the accompanying token economy system and provides the results of the intervention for 4 students.

Keywords: emotional and behavioral disorders, math word problems, strategy

Students with emotional and behavioral disorders (EBD) fall below national levels in all areas of academics, but most notably in reading and mathematics (Anderson, Kutash, & Duchnowski, 2001; Nelson, Benner, Lane & Smith, 2004). The need to remediate the academic deficiencies of students with EBD by using evidence-based instructional strategies continues to be an imperative in the field (Lane, 2004; Scott, Nelson, & Liaupsin, 2001). One specific area to be targeted for remediation is in the area of math word problems. In Hodge, Riccomini, Buford, and Herbst’s (2006) comprehensive review of studies investigating instructional interventions in mathematics for students with EBD only one study addressed problem-solving skills for students with EBD. Mathematics problem solving is defined as the presentation of a novel problem that requires the student to determine an appropriate course of action for attaining a goal before implementing a strategy to address the problem (Bottge, 2001; Fuchs et al., 2003; Polya, 1971). Mathematics problem solving is most commonly presented in the form of math word problems (Bottge, 2001).

Moving beyond these skill deficits are the potential performance deficits that students may demonstrate. That is, students may be unwilling to solve math word problems. Discrete task completion reinforcement may explain why students willingly solve a number of problems that involve a single step (e.g., adding two numbers or subtracting two numbers) but avoid word problems that require multiple steps to complete a single problem. For example, when completing a word problem, the computation is only one part in the multiple steps required to arrive at the correct solution. When students receive reinforcement only for deriving the correct final response, numerous additional opportunities for reinforcement are missed. Such a lean reinforcement schedule is likely to reduce the probability of students engaging in future problem-solving behavior.

To begin to address these issues, the strategy described in this article paired problem-solving steps with a token economy system and reinforced completion of each step to improve on-task behavior and problem-solving ability.
Table 1. Problem-Solving Steps

1) READ THE PROBLEM ALOUD
2) PARAPHRASE
   a. Give important information
   b. Repeat question aloud
   c. What is asked? What am I looking for?
3) VISUALIZE
   a. Draw a diagram
4) STATE THE PROBLEM
   a. I have . . . I want to find . . .
5) HYPOTHESIZE
   a. If I . . . then . . .
   b. How many steps?
6) ESTIMATE
   a. Round the numbers
7) CALCULATE
   a. Label
   b. Circle
8) SELF-CHECK
   a. Check every step
   b. Check calculation
   c. Does the answer make sense?

economy in order to provide extrinsic reinforcement for the completion of each step in the problem-solving strategy. To date, less focus has been given to the use of token economies for reinforcing academic behaviors as it is typically couched as a behavior management system. The guiding questions for this study were as follows: (a) Does this intervention improve on-task behavior during the mathematics instructional period? and (b) Does this intervention improve students’ daily percentage of math word problems completed correctly?

Method

Participants and setting

Participants included three boys in the same classroom of an alternative public school in a Midwestern city. Consent forms were sent home for all 7 students in the class, and 5 were returned. The fourth student was removed from school because of a family emergency, and the fifth student was moved to a secured 24-hr facility before receiving the intervention and was excluded from the sample.

Table 2. Participant Demographics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Ethnicity</th>
<th>Age (years)</th>
<th>Grade</th>
<th>Full-scale IQ</th>
<th>Exceptionalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Steve</td>
<td>Caucasian</td>
<td>11</td>
<td>5</td>
<td>68</td>
<td>EBD, MMD, ADHD, ODD</td>
</tr>
<tr>
<td>Eric</td>
<td>African American</td>
<td>11</td>
<td>5</td>
<td>81</td>
<td>EBD, PTSD, ODD</td>
</tr>
<tr>
<td>Travis</td>
<td>African American</td>
<td>11</td>
<td>5</td>
<td>88</td>
<td>EBD, ADHD, PTSD</td>
</tr>
<tr>
<td>Nathaniel</td>
<td>African American</td>
<td>10</td>
<td>4</td>
<td>68</td>
<td>EBD, MMD, ADHD</td>
</tr>
</tbody>
</table>

Note. EBD = emotional behavioral disorder, MMD = mild mental disability, ADHD = attention deficit hyperactivity disorder, ODD = oppositional defiant disorder; PTSD = posttraumatic stress disorder.

Table 2 presents demographic information for the three participants.

The exceptionalities listed in the final column of Table 2 were those identified on each students’ individualized education plan. An emotional and behavioral disorder was identified as the primary disability for all participants. The classroom teacher reported that all 4 participants had extensive histories of frequent, high-magnitude classroom disruptions, severe aggression, chronic noncompliance, and poor academic performance. She identified their time off-task as one of the most problematic areas that she dealt with on a day-to-day basis. The alternative school is described on their website as “a special school for students with severe and profound emotional and behavioral disabilities.” All procedures described in this study took place in these students’ classroom.

Materials

I created all of the word problem worksheets for the pretest/posttest and the daily work. The word problem content was pretest the fourth- and fifth-grade curriculum and was similar to the word problems that had been presented throughout the year. The four key arithmetic strands identified in the curriculum and addressed in the word problems were as follows: (a) two-digit-by-one-digit multiplication of whole numbers, (b) division as the inverse of multiplication, (c) algebraic thinking/counting patterns, and (d) identification of fractions. The pretest–posttest contained 10 problems addressing all four strands and the daily worksheets included between 20 and 25 word problems. All worksheets and the pretest–posttest addressed the four strands in roughly equal proportions. A large number of problems were included in the daily worksheets to provide enough opportunities to keep all students engaged during the entire 15-min work session thus allowing for an accurate assessment of on-task behavior. Because all 4 participants were permitted to use calculators during daily math and as an accommodation in statewide standardized testing, all students used calculators for the pretest–posttest and for the daily math worksheets.

I used a variety of materials in this study. During the intervention condition, student tokens consisted of holes punched in an index card with a standard hole puncher. Each student received a new index card at the beginning
of every session; the index cards had “Follow Directions” and “Try Your Best,” written on them to serve as visual cues. A laminated sheet with the eight problem-solving steps printed in 22-point font was given to each student during the intervention condition. Backup reinforcers for the token economy included classroom items that students typically used during free time such as the following: computer games, a football, a magnetic dartboard, and preferred snacks that were typically dispensed by the teacher as a reward during a normal school day. Last, two stop-watches were used by the data collectors.

Procedures
Two upper-level undergraduate students in a university teacher preparation program assisted in implementing this intervention. The first research assistant served in as the teacher. In a 2-hr training before the first baseline session, the first author used role-playing and modeling to prepare her to assume her various responsibilities in the study. The second research assistant served as the primary data collector. The first author trained her for 1 hr on the definition of student’s on-task behavior and how to use momentary time sampling to evaluate students’ on-task behavior. She then trained with me to 80% interobserver agreement on on- and off-task behavior during the administration of two pretests; they also trained to 95% on interobserver agreement in scoring the permanent product worksheets using the first pretest.

Baseline. All 4 participants completed two pretests before the first baseline session. The second pretest matched the posttest. During the baseline condition, all 4 students were given the daily math word problem worksheets consisting of 20–25 math word problems and a calculator. For the first four sessions, the work period lasted 10 min, but for the rest of the sessions, in the baseline and intervention conditions, they lasted 15 min to more closely approximate typical independent work time in the classroom. All sessions for baseline and intervention took place four times per week for 30 min during baseline and 60 min once participants began the intervention phase.

The research assistant acting as the teacher circulated around the students’ desks and helped students read words aloud but did not provide any other assistance in solving the math word problems. No verbal praise or prompts were given. The second research assistant collected direct observation data on students’ on-task behavior. The second research assistant also collected permanent product data in the form of the competed math worksheets. After the 15-min interval, she calculated the number of problems attempted and number of problems correct. Students were not informed about the number of problems that were completed correctly.

Intervention. When a student began the intervention condition, he was given the worksheet, calculators like the ones used in the baseline condition, a laminated sheet with the problem-solving steps printed on them, and an index card that served as the student’s token point card. The decision to implement the intervention was based on the criteria outlined in Kennedy (2005, p. 152): “Individual baselines are established (for each student’s behavior), consistent patterns are observed and then the independent variable is systematically introduced to one baseline at a time.” Steve was the first student to receive the intervention beginning with the ninth session. The intervention was implemented for Eric after seeing an initial increase in Steve’s performance and after stable data patterns were observed for Eric during the eighth and ninth baseline sessions for on-task behavior and problems solved correctly. However, on the 16th session, because of the high variability in Eric’s rates of on-task behavior and percentage of problems completed correctly the intervention was adjusted for Eric. This adjustment is described in the self-management section.

Because consistent increases were observed for both dependent variables for Steve and because of initial increases in Eric’s data following intervention, the intervention phase for Travis began after the 11th baseline session following a stable data pattern on three baseline sessions. Last, the intervention was implemented for Nathaniel on the 16th session. Although there was some instability in the data for Travis and there was an upward trend in on-task behavior for Nathaniel, extenuating family circumstances that included a deterioration of his mother’s health and the possibility of moving to another county for this participant made it necessary to implement the intervention at that point in time. Nathaniel’s mother passed away during this final week of intervention, explaining why he only received two sessions of the intervention condition. He returned to see his classmates at the same time they were taking the posttest and asked whether he could take it also. Thus, his intervention results should be interpreted with caution.

During the intervention condition, the first research assistant followed the strategy acquisition training model outlined by Montague and Bos (1986). In the first intervention session, the research assistant reviewed the strategy and modeled it by solving the first problem on the worksheet. Next, the student was given an opportunity to solve another problem using the step-by-step strategy while receiving corrective feedback. In subsequent interventions sessions, the research assistant reviewed strategy steps, monitored verbal rehearsal and provided corrective feedback on randomly selected problems.

In addition, during this condition the first research assistant monitored the completion of each step for all the students. After completion of each step (e.g., Read aloud, Paraphrase, Draw a picture …) the students were given reinforcement through the token economy system. On the index card that represented the student’s ‘point card,’ the first research assistant dispensed ‘points’ by punching holes...
in the card with a hole puncher for the completion of each of the steps of the problem-solving heuristic (i.e., one hole punch per completion of one step). Verbal praise such as, “Great job figuring out what the problem is asking,” was paired with the dispensing of points. For example, to solve the following problem, “A store sold 3 boxes of pencils for 6 dollars. If there are 44 pencils in a box, how many pencils were sold in all?” students would apply each of the eight problem-solving steps. First, students would read the problem aloud, then they would identify important information (e.g., “I want to know how many pencils were bought in all. I know how many boxes were sold and I know how many pencils are in each box”). Students would then visualize how to solve the problem and often draw a diagram to aid in comprehension (e.g., “I can draw 3 boxes and label each box with the number 44 to represent 44 pencils in each box”). Students would then summarize the problem elements (e.g., “I have the total number of boxes which is 3, and I have the total number of pencils in each box, which is 44. Now I want to find out how many pencils were sold in all”). Then students would hypothesize how to solve the problem (“If I multiply the number of boxes—3 by the number of pencils in each box—44 that will tell me how many pencils were sold in all”). Students would then estimate at what would be a reasonable answer by rounding the numbers in the problem (e.g., “Because 44 is close to 50 and I know 50 × 3 = 150, an answer close to that number would make sense.”). Students would then calculate their answer and check it by one more time.

When the 15-min work session was completed these points were exchanged for snacks or minutes of engaging in a preferred activity (the backup reinforcers). The first research assistant also played with students during their earned free time in the intervention condition. For the students who were still in the baseline condition, after they completed their math worksheet, they returned to their daily seatwork in another part of the classroom. The 4 participants then completed a posttest 51 days after taking the same test as a pretest. When the pretest and posttest were given, no components of the intervention condition were used. This included not allowing students to have the laminated problem-solving steps sheets on their desks. Students were given as much time as they needed and the first research assistant, acting as the teacher, only gave assistance in reading unknown words.

**Self-management.** A phase change for the intervention was implemented in the 15th session for Eric. Because the first step in the problem-solving heuristic was “Read the problem aloud,” Eric expressed difficulty concentrating because of the other students reading their problems aloud. This difficulty was reflected in his highly variable data (for problems solved correctly and on-task behavior) in the first six sessions of the intervention condition. Eric was moved to another part of the classroom, approximately 20 feet away from the other students, and allowed to self-manage his token economy. He was given verbal instructions by the first research assistant to punch a point for every step in the problem-solving strategy that he completed and reminded that his total number of points should not exceed the maximum number of points any other student received. Upon completion of the 15-min work session, Eric was able to spend his points on backup reinforcers with his peers.

**Dependent measures**

Baseline and intervention assessment data were collected for three dependent measures for this study: (a) percentage of problems completed correctly on daily worksheets, (b) percentage of problems completed correctly on a pretest–posttest measure, and (c) percentage of time on task. The word problems were scored as correct if the appropriate number (the answer) was written below the problem. The percentage correct was calculated by dividing the number of the number of problems completed correctly by the number of problems completed. On-task behavior was defined as each student having his eyes oriented toward his paper and working on the assigned math word problems. Looking at the teacher during communication about the math problems (e.g., when the teacher was reading difficult words aloud to the student) was also defined as on-task behavior. Percentage of time on task was measured using momentary time sampling in 30 second intervals for each participant across the 15-min work session. Percentage of time on task was calculated by dividing the number of intervals the student was rated as on-task by the total number of intervals.

**Interobserver agreement**

During 35% of the daily sessions for all conditions, interobserver agreement was assessed for on-task behavior. The first author assessed on-task behavior simultaneously with the second research assistant, who acted as the primary data collector. An interobserver agreement estimate was calculated using the interval-by-interval method—dividing the number of intervals agreed by the number of intervals agreed and disagreed and multiplying by 100 (Cooper, Heron, & Heward, 2007). Interobserver agreement ranged from 83% to 97%, with a mean of 92%. In the baseline phase, the average interobserver agreement for Steve was 95%, for Eric it was 90%, for Travis it was 93% and for Nathaniel it was 94%. In the intervention phase the average interobserver agreement for Steve was 96%, for Eric it was 91%, for Travis it was 91% and for Nathaniel it was 94%. For Eric, in the self-management phase, the average interobserver agreement was 98%. Interscorer agreement on the permanent product word problems was also measured for 17% (4/24) of the worksheets including both pretests, the posttest and daily worksheets. Interscorer agreement on the total number of problems scored was 99.2% (135/136).
Treatment fidelity

Treatment fidelity was monitored directly by the first author during 42% (5/12) of the intervention sessions. The following elements of the intervention were monitored. The research assistant (a) either taught or reviewed the problem-solving strategy, (b) modeled the strategy for participants during their first intervention session, (c) provided points to all participants for completion of each step, (d) paired points with verbal praise, and (e) provided preferred backup reinforcers immediately upon the conclusion of the work session. If any element of the intervention was not implemented, I provided immediate feedback. In addition, Eric’s treatment fidelity during the self-monitoring phase was monitored for 66% (2/3) of the sessions, in which I tracked the number of problems completed and the number of steps completed while Eric was working to insure that he gave himself a point for each problem completed. These five (or six in the self-monitoring phase) elements were met 100% of the time for all intervention sessions.

Results

In this section, the results of the percentage of problems completed correctly on the daily worksheets and the scores from the Pretest to the Posttest are presented for each participant. Then the rates of on-task behavior during the daily 15-min work sessions are presented for each participant.

Percentage of problems completed correctly

Figure 1 presents the percentage of problems completed correctly on the daily worksheets across students. For the daily worksheets, the percentage of problems completed correctly increased for all 4 participants from baseline to intervention, and for Eric they increased from baseline to self-monitoring. Steve’s baseline percentage of problems completed correctly (M = 13%, range = 5–25%; SD = 6.7) steadily improved throughout intervention (M = 91%, range = 75–100%; SD = 10.3). Visual analysis indicated a clear differentiation of levels from baseline to intervention with excellent stability across both conditions. The total number of problems completed decreased from baseline (M = 19.1) to intervention (M = 5.7). Steve’s percentage of problems correct had no overlapping data points and demonstrated a rapid immediacy of effect from baseline to intervention. For Eric, his average percentage of problems completed correctly improved from baseline (M = 13%, range = 0–33%; SD = 10.4) to intervention (M = 42%, range = 0–75%; SD = 33.0) to self-monitoring (M = 64%, range = 54–81%; SD = 15.6). The total number of problems completed decreased from baseline (M = 19.5) to intervention (M = 11.8) to the self-monitoring phase (M = 9.7). The poor stability of his performance in the initial intervention phase was precipitated by the behavior issues noted in the self-monitoring section. Eric’s percentage of problems completed correctly demonstrated high variability, no clear differentiation of levels and multiple overlapping data points from baseline to the initial intervention. However, when comparing baseline to the self-monitoring phase, there are no overlapping data points, good stability, and low variability. For Travis, his average percentage of problems solved correctly improved from baseline (M = 15%, range = 4–25%; SD = 6.9) to intervention (M = 44%, range = 0–75%; SD = 31.9). The total number of problems completed decreased from baseline (M = 22.4) to intervention (M = 7.6). There was high variability and poor stability in Travis’ intervention phase, but only one overlapping data point when comparing baseline data to intervention data and medium overall level change from baseline to intervention. Further, a trend line indicates an accelerating trend in the intervention phase after accounting for data gathered in Session 15, where behavioral issues caused him to not solve a single problem correctly. Nathaniel’s baseline percentage of problems completed correctly (M = 19%, range = 9–42%; SD = 10.0) improved to (M = 67%, range = 50–83%; SD = 14.1). The total number of problems completed decreased from baseline (M = 15.6) to intervention (M = 5.0). Nathaniel’s data indicated no overlapping data points from baseline to intervention and indicated a positive trend in the intervention phase. The stability of performance in baseline data and the clear differentiation between baseline and intervention are suggestive of the effect of the intervention; however, there are only two data points in the intervention phase. In regard to the results of the pretest–posttest measure for problems solved correctly: Steve improved from 10% (1/10) to 60% (6/10), Eric improved from 20% (2/10) to 30% (3/10), Travis’ score decreased from 30% (3/10) to 10% (1/10) and Nathaniel improved from 10% (1/10) to 60% (6/10).

On-task behavior

Figure 2 presents the percentage of intervals participants were on-task during the daily 15-min work sessions. The average percentage of time on-task stayed the same or improved from baseline to intervention; however, reactivity appeared to have affected early baseline data points for on-task behavior. Although some measures were undertaken to control for reactivity (e.g., researchers spent time with the students in their classroom on at least two separate occasions before beginning the study), all 4 participants had initially high levels of on-task behaviors that eventually began to decrease. The classroom teacher noted after the second baseline session that the students were “putting on a show for you all.” Habituation appeared to occur for different students for a different number of sessions and comparisons of the last two sessions in baseline versus intervention are illustrative of the effect of the intervention.
Fig. 1. Percentage of problems solved correctly across participants.
Fig. 2. Percentage of on-task behavior across participants.
For Steve, his baseline percentage of on-task behavior 70% (50–100%; SD = 20.3) improved in the intervention condition to 93% (76–100%; SD = 11.5). There was fair stability of performance and the final data point indicated an accelerating trend bringing him up to 97% on-task behavior for the entire 15-min interval. Steve’s percentage of intervals on-task had a decreasing trend in the baseline condition, had a rapid immediacy of effect after the intervention and had low variability in the intervention phase. A comparison of the final two data points in baseline were 63% and 60% and in intervention were 83% and 97% suggesting considerable difference from baseline to intervention.

Eric’s baseline rate of on-task behavior 67% (43–90%; SD = 15.9) stayed the same during intervention 67% (6–100%; SD = 40.2) but improved in the self-monitoring phase 99% (97–100%; SD = 1.7). After a rapid immediacy of effect from baseline to intervention, Eric’s intervention data had considerable variability and a number of overlapping data points. Because the participant’s data fluctuated in successive sessions, it is difficult to discern a trend. It is similar to his percentage of problems completed correctly but the self-management phase had stable performance and improvement in level from the baseline to the self-management phase. A comparison of the final two data points in baseline (53% and 43%) suggests considerable difference when compared to the final two points in the self-monitoring phase (97% and 100%).

For Travis, his average baseline percentage of on-task behavior 75% (43–90%; SD = 18.7) improved to an intervention average of 87% (76–97%; SD = 9.6). Travis’s high percentage of on-task behavior did not begin to decrease in baseline until the sixth session. However, there was a decreasing trend in percentage of intervals on-task during baseline and an increasing trend during intervention. Travis’ data indicates considerable reactivity throughout the first five sessions of the baseline condition, there is then a decelerating trend until the intervention is implemented at which point the level has an initial increase and then an accelerating trend in the next three data points. In the final two data points of the baseline phase rates of on-task behavior were 67% and 67% and in the intervention phase they were 80% and 93%.

Nathaniel’s baseline rate of on-task behavior 76% (3–97%; SD = 25.7) improved in the intervention condition to 82% (67–97%; SD = 21.2). Nathaniel’s baseline data had high variability, and there are a large number of overlapping data points from baseline to intervention. This makes it difficult to discern a clear trend in the data. Furthermore, there is a decreasing trend during the intervention phase with the final data point being lower than the majority of the baseline data points. However, this is only based on two data points in the intervention phase and was during a period of extreme emotional duress for Nathaniel. Therefore, all results for Nathaniel should be interpreted with caution.

Discussion

This study examined the effects of an intervention using a behavior chain of math problem-solving steps and a token economy reinforcing the completion of each step to affect math word problem-solving skills and on-task behavior for elementary school students with EBD. Results indicated that during the intervention condition students’ problem-solving accuracy and on-task behavior increased, as compared to data from the baseline condition. Differentiation between baseline and intervention is apparent via visual inspection for all 4 participants. In regard to the daily percentage of problems completed correctly, it is salient to note that the mean number of problems attempted also decreased dramatically. This is indicative of students using the problem-solving strategy to some degree, rather than writing down any number or “I don’t care” or simply adding all of the numbers listed in the problem. One student, Eric, did not demonstrate consistent improvement in regard to on-task behavior or math word problems solved correctly in the initial intervention phase. However when a self-management system was put into place for Eric, following the initial intervention phase, it yielded improvements in both dependent variables.

In terms of the pretest–posttest designed to measure problem-solving ability, three students (Steve, Eric, and Nathaniel) demonstrated improvement. Travis got fewer correct on the posttest. The results of the pretest–posttest measure are surprising for two reasons. First, although Travis had the highest full-scale IQ of the 4 participants, he was the only participant to do worse on the posttest. Second, Nathaniel’s posttest score improvement from 10% (1/10) to 60% (6/10) was most surprising given that his mother had passed away earlier in the week. It should be noted that Nathaniel was actually discouraged from taking the posttest by the author, but repeatedly asked whether he could take it.

Results from the study expand on Alter, Wyrick, Brown, and Lingo (2008) by demonstrating the effects of this intervention on 4 more participants with challenging behaviors. Furthermore, this study contributes to the small but growing research base of viable interventions for remediating the academic difficulties of students with EBD. The use of a problem-solving heuristic combined with a token economy adds to the growing list of intervention strategies that can be used to increase on-task behavior and academic performance including the following: response cards (Lambert, Cartledge, Heward, & Lo, 2006), dependent group contingencies (Heering & Wilder, 2006), increased opportunities to respond (Sutherland, Wehby, & Yoder, 2002), interspersed easier problems (Skinner, Hurst, Teeple, & Meadows, 2002), and strategic self-monitoring (Rock, 2005).

While the intervention appeared to affect problem-solving skills and on-task behavior, there were a number of overlapping data points especially in terms of on-task.
behavior. This may be attributed to a high level of reactivity during the baseline phase as this study had three new people coming into the classroom (the research assistant acting as the teacher, the research assistant acting as the primary data collector, and the author measuring treatment fidelity and serving as the reliability data collector). In the initial sessions of the baseline phase, all students were content to write down any number in order to answer each problem, as evinced by the high average number of problems completed and the low number correct for all 4 participants in the baseline phase. This is consistent with Skinner’s (2002, p. 648) “discrete task completion” hypothesis that completed tasks (e.g., word problems) serve as conditioned reinforcers. It is interesting that it does not appear that the task had to be completed correctly to be a conditioned reinforcer, but rather simply having an answer for each problem was reinforcing. This may explain the high levels of initial on-task behavior, which the classroom teacher noted as atypical.

Two participants demonstrated some instability for both dependent measures across the baseline and initial intervention phase. For Eric, the implementation of a self-management phase was necessary as a result of the first step to “Read the Problem Aloud,” in the problem-solving heuristic. As the other students read their problems aloud, Eric became more and more distracted explaining the dramatic decreases for both dependent variables in the 12th and 15th sessions of the intervention phases. During these sessions, he emitted numerous verbal outbursts asking peers to “Please read quieter” and “Please shut up.” The self-management phase, which allowed Eric to work in another part of the classroom, yielded consistently higher levels of on-task behavior and number of problems solved correctly and indicates a way that the token economy could be faded to use more self-management and less teacher-directed external reinforcement. Further, because Eric was allowed to punch his own points in his token economy, this adaptation may present a viable way for this method to be implemented by teachers with higher student to teacher ratios. In regard to Travis’ data for math problem-solving skills and on-task behavior, he demonstrated some variability throughout the intervention phase and a surprisingly lower score on the posttest. In the two sessions (Sessions 13 and 15) in which his levels decreased below his previous performance, there was same indication that he was angry over events that had happened earlier in the day. This variability in the data could be related to or indicative of the levels of emotional distress and behavioral difficulties that are experienced by students with EBD and ultimately affect their educational performance.

Several limitations must be considered when interpreting the findings from this study. First, this study was conducted in a single classroom with students in a highly specific environment (an alternative school for students with EBD). Related to this is the limited sample size of 4 students, which affect the generality of these findings and replication studies are necessary to improve the external validity of these findings. It is unfortunate that this limitation is prevalent among studies in the literature that involves interventions for the academic difficulties of students with EBD. Second, the issue of external validity must also be considered because the intervention was neither faded, nor was it turned over to the classroom teacher. Thus, there are no measures of treatment acceptability or durability of the effect of the intervention after it has been faded. Furthermore, the level of intensity and amount of time required in the instruction and reinforcement elements of this study may not be palatable or even possible for many teachers. The research assistant acting as the teacher was constantly moving, providing instruction, reading words, punching cards and monitoring for completion of steps for all participants. It is clear that this is not feasible with a larger numbers of students in the classroom. Third, in consideration of the described intervention, the issue of time and the total number of problems that could be solved may make this intervention prohibitive for some teachers. For a 15-min work session in which students completed on average seven word problems they then had a 25-min time for backup reinforcers. However, when considering the issue of quality versus quantity dedicating more time to solve fewer problems but with greater accuracy may be preferable. Fourth, the issue of reactivity affected the data, and the overall variability in the data for some of the students indicates that this intervention may not be suitable for all students. Fifth, it was noted by the research assistant acting as the teacher, that students did not complete every step for every math problem completed. However, exactly what steps were completed by the different students was not recorded. Sixth, because the intervention was presented with an instructional component (i.e., the behavior chain of problem-solving steps) and a reinforcement component (i.e., the token economy) it is impossible to know what the separate contributions were for each of these components. This does point to a future research direction that would involve separately presenting the problem-solving strategy and the token economy and then combining them to determine the singular effect of each component and then evaluating their combined effects on a group of students.

The results of this study support the use of this intervention to improve math problem-solving skills of students with EBD; however, other future research is necessary. Future research directions may include using different, briefer problem-solving heuristics combined with token economies. This may also lead to using different problem-solving strategies for different arithmetic strands beyond the four types detailed in this study. Also, fading the token reinforcement from a continuous reinforcement schedule to a less dense schedule of reinforcement will provide important insight into how quickly this artificial reinforcement system can be replaced with more natural, adaptive reinforcers. An alternating treatment design that uses the problem-solving heuristic and the token economy in
separate phases and measures the same dependent variables would provide important information for future research and practice. Last, transitioning to have classroom teachers implement this intervention with students, and implementing the intervention with larger classrooms of students in group design research will give further information into the efficiency, effectiveness and overall acceptability of the intervention.

References


Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. Journal of Emotional and Behavioral Disorders, 10, 5–15.