

2010

Supplementing Accelerated Reading with Class-wide Interdependent Group-Oriented Contingencies

Danielle N. Pappas

Christopher H. Skinner

Amy L. Skinner

SUPPLEMENTING ACCELERATED READING WITH CLASSWIDE INTERDEPENDENT GROUP-ORIENTED CONTINGENCIES

DANIELLE N. PAPPAS

The University of Tennessee and Linn Benton Lincoln Education Service District

CHRISTOPHER H. SKINNER AND AMY L. SKINNER

The University of Tennessee

An across-groups (classrooms), multiple-baseline design was used to investigate the effects of an interdependent group-oriented contingency on the Accelerated Reader (AR) performance of fourth-grade students. A total of 32 students in three classes participated. Before the study began, an independent group-oriented reward program was being applied (i.e., a student received access to a tangible reward after passing each AR comprehension test). This program was supplemented with an interdependent group-oriented contingency, and results showed that the number of quizzes passed per week increased immediately after the intervention was applied; however, this increase was not maintained. When students were divided into ability groups based on their average baseline performance, the lowest performing students exhibited a statistically significant increase in quiz performance (i.e., quizzes taken, quizzes passed, and book level), but no significant changes were found in the average and high performing groups. © 2010 Wiley Periodicals, Inc.

Proficient reading is the result of a hierarchical process of skill development, and these skills include orthographic processing, phonemic awareness, phonics, fluency, and vocabulary (Adams, 1990; Denton & West, 2002; Johnston, Anderson, & Holligan, 1996; National Reading Panel, 2000; Pugh et al., 2001; Rego, 2006; Shanahan, 2005). The primary goal associated with developing these skills is comprehension. Although various procedures have been developed to enhance prereading and comprehension skills (e.g., text–structure awareness and the ability to think aloud, predict, and create graphic representations of text; see Duke & Pearson, 2002), little skill development will occur unless students choose to read (Skinner, Pappas, & Davis, 2005). Two strategies that have been shown to be effective for increasing the probability of students choosing to engage in academic behaviors are to provide choices (e.g., Dunlap et al., 1994) and to strengthen reinforcement for desired academic behaviors (e.g., Popkin & Skinner, 2003).

Allowing Students to Choose Work

Allowing students with disabilities to choose assignments has been shown to reduce problem behaviors and increase the probability of students engaging in academic behaviors in school settings (Dunlap et al., 1994; Dyer, Dunlap, & Winterling, 1990; Kern, Bambara, & Fogt, 2002). Several hypotheses may explain these findings. Providing choices in academic environments may instill a sense of empowerment and control, two factors that underlie intrinsic motivation, which may increase the probability of students choosing to engage in academic as opposed to other tasks (Spaulding, 1992). Allowing students to choose reading material may enhance their engagement because students often choose material that they find most interesting. Martin-Palmer, Codling, and Gambrell (1994) reported that 75% of the participants in their study chose books according to personal interests. When asked to talk about the most interesting books they had read, the participants repeatedly discussed the books they had chosen themselves.

The principle of least effort may also explain increases in engagement resulting from assignment choice (Friman & Poling, 1995). Researchers have shown that, when given the choice of

two academic behaviors, students will choose the academic behaviors that require the least effort (Billington, Skinner, & Cruchon, 2004; Billington, Skinner, Hutchins, & Malone, 2004; Martin, Skinner, & Neddenriep, 2001). Furthermore, as the difference in relative effort increases, so does the probability of students choosing to engage in the less-effortful academic activity (Billington & Ditommaso, 2003; Billington & Skinner, 2002; Cates & Skinner, 2000). When providing students with a choice of reading material, they may be more likely to choose material that requires less effort to read (Skinner et al., 2005). This research on effort and choice suggests that providing choices may cause students to read material that is not challenging (e.g., below their current reading level and/or containing little or no unfamiliar information, concepts, or vocabulary), which may hinder reading skills development.

The principle of least effort also may explain why some students with poorly developed reading skills fall farther behind. Students with poor reading skills often read slower and require more effort to read than do those with stronger reading skills (Stanovich, 1986). Consequently, when a group of students are given identical reading assignments, poor readers may be less likely to choose to read the material because the effort required to complete the task is so large (Skinner, 1998). As reading skill development is dependent on students choosing to read, these students may fall even farther behind their peers.

Reinforcement Strength and Choice

One strategy to increase the probability of students choosing to engage in academic behaviors that require more effort than competing behaviors (e.g., staring out the window) is to strengthen the reinforcement for those behaviors. Specifically, educators can increase the probability of students choosing to engage in high-effort assignments by enhancing the rate, quality, and immediacy of reinforcement (Billington et al., 2004; Cates & Skinner, 2000; Mace, Neef, Shade, & Mauro, 1996; Meadows & Skinner, 2005; Neef, Mace, & Shade, 1993; Neef, Shade, & Miller, 1994). These findings suggest that an interaction between reading skill development (e.g., reading fluency) and reinforcement (intrinsic and extrinsic) can have a strong influence on student reading behavior (Skinner, 1998; Skinner, Williams, & Neddenriep, 2004).

Consider two students who both read assignment material that is humorous. Also, their reading comprehension is assessed with a subsequent quiz. One student with weak skills reads the material slowly (takes 25 minutes), and the other student with stronger skills reads the material rapidly (5 minutes). Both are reinforced with (a) five chuckles as they read, (b) one belly laugh that occurs after they read the last line of the material, and (c) a score of 85% on a quiz they take the following day. An examination of relative reinforcement rates reveals that the rapid reader has one chuckle (lower-quality intrinsic reward) per minute of reading, whereas the slower reader has one chuckle per 5 minutes. Also consider the delay or time to reinforcement with respect to the belly laugh (high-quality intrinsic reward). After beginning to read, the rapid reader accesses this high-quality reinforcer in 5 minutes; the slower reader does not gain access to the high-quality reinforcer until 25 minutes has elapsed. Consequently, even if the quality and time to the external reinforcement is held constant (85% quiz grade the following day), the schedule of intrinsic reinforcement for the two students (e.g., rate and immediacy of reinforcement) and the additional effort required for the slow reader to read and comprehend make it less likely that the slow reader will choose to read (Skinner, 1998). Instead, the poor reader is more likely to choose to engage in competing behaviors (e.g., watch a situational comedy on television) that result in higher rates and more immediate reinforcement than the rapid reader. To address this problem, educators may have to further strengthen reinforcement so that the poor readers choose to read (Skinner et al., 2005; Skinner, Skinner, & Armstrong, 2000).

A report from the National Reading Panel (2000) suggests that independent silent reading can help students develop comprehension skills necessary for success. Educators and parents can allot time for silent reading, yet skills are unlikely to improve unless students choose to read (Skinner et al., 2005). Computerized reading programs, such as Accelerated Reader (AR; Reading Renaissance I, 1998) and Electronic Bookshelf (Carter, 1996) may enhance the probability of students choosing to read by allowing students to choose from many options. The AR program is designed to encourage more reading practice by providing comprehension quizzes for more than 50,000 books and data that can be used for monitoring and feedback. After reading a book, students can take a multiple-choice quiz that assesses their comprehension. The computer calculates a quiz score, and points can be awarded based on the student's performance. A passing score is 60% accuracy on the comprehension quizzes (Reading Renaissance I, 1998).

The AR program is designed to incorporate the elements of choice and reinforcement for reading activities. Students are allowed to choose material to read. When an AR quiz is passed, the student is awarded points that can be exchanged for rewards. This reward system is an independent group-oriented contingency because all students have access to the same rewards based on their own performance meeting the same criteria (Litow & Pumroy, 1975; Sharp & Skinner, 2004). Although independent group-oriented rewards are considered fair, they may not be effective with students who require much effort to perform the behavior (Skinner, Skinner, & Burton, 2009). Supplementing independent group-oriented contingencies with interdependent group-oriented contingencies may increase the probability of these students choosing to read (Popkin & Skinner, 2003). With interdependent group-oriented contingencies, the entire group receives access to a reward when they meet a group-oriented goal (e.g., number of AR quizzes passed by the class). Researchers have used interdependent group-oriented contingencies to enhance performance in vocabulary (Lew, Mesch, Johnson, & Johnson, 1986), spelling (McLaughlin, Herb, & Davis, 1980; Saigh, 1987; Shapiro & Goldberg, 1986), math and language arts (Wilson & Williams, 1973; Winnett, Battersby, & Edwards, 1975), and reading (Sharp & Skinner, 2004; Stewart & McLaughlin, 1986).

There are several reasons why interdependent group-oriented reward procedures may lend themselves to application with the AR program. With the AR program, students can choose material that is appropriate for their reading skills. For example, in a fourth-grade class, strong readers may choose materials from the fifth-grade reading level, and weak readers may choose materials from a second-grade level. When independent group-oriented rewards are applied, stronger readers may complain that it is not fair that weaker readers can earn points and rewards by reading easier and typically briefer books and answering easier questions (Skinner et al., 2009). Additionally, stronger students may ridicule weak readers (Skinner, Cashwell, & Dunn, 1996). When interdependent group-oriented rewards are applied, however, all students may benefit when their classmates succeed. Consequently, stronger readers are likely to prefer that weaker readers have easier and briefer books and may encourage their less proficient peers (Cashwell, Skinner, Dunn, & Lewis, 1998; Pigott & Heggie, 1985). Also, when a class is engaged in sustained silent reading, students who choose not to read are unlikely to disrupt classmates who choose to read because each student is more likely to be rewarded when classmates perform well (Skinner et al., 2009).

It is often difficult for educators to deliver activity rewards (e.g., listening to music during independent seat-work) to some students, but not others. Therefore, when independent group rewards are used, educators often use tangible rewards. With interdependent group rewards, however, all or none gain access to rewards, making it easier for school personnel to administer and deliver activity rewards, which are often free and powerful (Skinner, Skinner, Skinner, & Cashwell, 1999). Finally,

when an entire group learns that they met a criterion, they are likely to engage in social behaviors (e.g., a group cheer) that serve as additional reinforcement for the behavior (Skinner et al., 2009).

Supplementing AR with Interdependent Group Contingencies

Sharp and Skinner (2004) used an A-B phase design to evaluate the effects of two interdependent group-oriented contingencies on AR reading quizzes passed in an intact second-grade class of 13 Black students. The entire class was given an ice-cream party if each member of the class passed an AR quiz within 6 weeks. The other contingency was weekly and included a randomly selected criterion. On Friday, the teacher would randomly select a criterion slip of paper (13 in the pool, with criteria 1–13 written on the slip) that represented the number of quizzes the class had to pass to earn the reward, extra recess that Friday. The criteria were randomly selected so that students would do their best, as opposed to giving up when they judged that they could no longer meet a known criterion or stop when they already exceeded the known criterion (Popkin & Skinner, 2003). For the first week, students read in pairs; after that, they were allowed to continue paired reading or they could choose to read independently. Prior to the program (A phase) the class passed an average of 0.7 quizzes per week. After the two interdependent group-oriented reward programs were applied, they passed an average of 7.5 quizzes per week.

Purpose

The design of the AR program allows students to choose their reading and is easily supplemented with independent group-oriented rewards delivered based on quizzes passed; however, it may not be effective with students with poorly developed reading skills because the effort required to read may not be worth the reinforcement they receive for reading. Therefore, students with weaker reading skills may need more powerful reinforcement to cause them to choose to read AR books. Although Sharp and Skinner (2004) found evidence that interdependent group-oriented reward programs may increase the number of AR quizzes passed, they used an A-B design that did not control for threats to internal validity (e.g., history, testing, maturation).

The primary purpose of the current study was to extend this research by using a multiple-baseline-across-classrooms design to evaluate the effects of an interdependent group-oriented reward program on AR quizzes passed. To explore effectiveness across groups, we used analysis of variance (ANOVA) to analyze quizzes-passed data in poor, average, and strong readers across baseline and intervention phase means. We conducted similar analyses on our other secondary data, the number of quizzes taken and reading level of material that students chose to read. The data on quizzes taken were collected to ensure that students did not start taking quizzes one after another, without reading the material, in hopes of increasing quizzes passed. The data on book level was collected to ensure that students did not choose to read easier material in an attempt to increase the probability of passing more quizzes to earn rewards.

METHOD

Participants and Setting

Participants were drawn from a pool of 51 students enrolled in an inner-city public school in the southeastern United States that served children in kindergarten through fifth grade. The students enrolled in this school were 74% white, 21% Black, and 4% other minorities. More than 90% of the students qualified for free or reduced lunches. All three teachers were female. Teachers A, B, and C had 1, 6, and 20 years of teaching experience, respectively. Although all students in the three participating classrooms were included in the intervention, data were collected and analyzed only for those students whose parents consented. Thus, the study included 32 (17 boys and 15 girls; 24 White

and 8 Black) fourth-grade students, 8–11 years old. Researchers were not given permission to obtain information regarding disabilities or other school records. The AR program data indicated, however, that the participants were reading at between the first- and sixth-grade levels.

Materials

Students were asked to continue participating in the AR reading program already in progress at the school. There was a minimum of 100 books per reading level in the school library. Each classroom also contained small libraries of AR books. When students completed reading a book, they could elect to take a 10-question, multiple-choice comprehension quiz. The quiz scores were calculated by the AR program and maintained in the program's database. Students were not permitted to retake a quiz for a book they had already passed (60% accuracy). After a student scored 60% or higher on a quiz, that student was allowed to choose a reward. The reward pool included school supplies (e.g., pencils, pens, erasers), edibles (e.g., candy), and small inexpensive toys (e.g., balloons, tops).

The AR program generated weekly printouts for each student indicating the number of quizzes passed that week, the primary dependent variable. We also examined the data on quizzes taken and book reading level average. The students' names were eliminated and replaced with numbers to ensure confidentiality. For each classroom, two plastic containers were used to hold slips of colored paper with criteria and rewards written on them. A third container was used to store students' reward suggestions.

Procedures

Participants were introduced to the intervention on a Monday. Information was provided by the teacher and the primary experimenter. Students were informed that their current AR program would remain in place and that they would continue to earn a reward when they passed an AR quiz. Also, they were told that the entire class would have an opportunity earn additional rewards, contingent on the number of AR quizzes passed by the group. Specifically, students were told that on Friday the number of quizzes passed for the entire classroom would be checked. This number then would be compared to a number drawn from the criteria container. If the class met or exceeded the randomly selected criterion, a reward would be randomly selected and be delivered to the entire class within a week. Students were not informed of the required criterion or the reward as the drawing did not occur until Friday, but the students were informed of the range of the criterion and a menu of rewards available in the drawing.

To ensure that students received access to the group reward, the criterion for the first week of the intervention was rigged to ensure that each met the criteria (Sharp & Skinner, 2004). For example, if the class passed 14 quizzes that week, all criterion slips of paper were 14 or lower, so that the class was guaranteed to win. After the first week, the criterion was set at 20% above and 20% below the intervention phase average. For example, if during the first 3 weeks of intervention, the average number of quizzes passed was 20, the range of possible criteria would have been 16–24. The class was informed on Monday of the criterion range for that week.

The rewards were chosen by the teacher. Teacher C already had a list of rewards that she considered to be high quality. Teachers A and B agreed that those rewards were appropriate for this intervention. The rewards included (a) an ice cream party, (b) a popcorn party, (c) lunch in the classroom, (d) music during seatwork, (e) board game day, (f) pajama day, (g) treat day (i.e., candy bars or cookies), (h) computer time, (i) free pencils, (j) extra free time, and (k) arts and crafts day. The students were allowed to make a reward suggestion at any time. Students could write a suggestion on a blank slip of paper and deposit it in the suggestion box. The primary researcher would discuss the suggestion with the teacher and determine whether the new reward would be included. If the

suggestion was accepted, it was announced to the class at the beginning of the week and added to the pool. During the study, only one reward suggestion was received; this particular reward was not included in the intervention because it was too costly. Therefore, the list of rewards remained the same throughout the intervention phase.

Each teacher provided approximately 30 minutes per day for independent silent reading. In addition, students were encouraged to read their AR books when independent seatwork was completed. Upon completion of an AR book, teachers often reminded students to take the computer quiz as soon as possible on the classroom or library computers.

Each Friday morning, student assessment data from the AR program were printed. The students were told the number of quizzes they had passed that week. Then either the teacher or a student was asked to draw one slip of paper out of the criteria container and read the number aloud. If the class passed more AR quizzes than the randomly selected criterion, a slip of paper from the reward container was randomly selected and read aloud. Then the teacher and primary researchers decided on a day to deliver the reward. Rewards were delivered sometime the following week. If the class did not exceed the criterion number, a reward was not drawn; however, the students were reminded that a new drawing would occur again the next Friday and were encouraged to continue reading AR material and passing quizzes.

Experimental Design

A multiple baseline design across classrooms was used to compare students' performance from baseline to intervention. Collaboratively, the participating teachers decided the order of implementation in the three classrooms with Class A, B, and C applying the intervention following 4, 5, and 6 weeks of baseline, respectively. Because low attendance could impact the number of AR quizzes passed, we planned to eliminate any data when attendance fell below 75% for the week. This situation never occurred as attendance ranged from 75% to 92% over the study.

Dependent Variables and Data-Analysis Procedures

The primary dependent variable was the number of AR quizzes passed. The number of AR quizzes taken and the reading level of books also were analyzed. Although the typical school week included 5 days, during baseline, week 2 was a 4-day week due to a national holiday, and, during the intervention phase, weeks 5, 7, 8, 9, 10, and 12 were 4-day weeks due to holidays, in-service, and school closings. To compare data, for each week the average number of quizzes taken and passed per school day was analyzed. Visual analysis of time series graphs was used to interpret these data.

Additionally, to make comparisons across students, baseline phase and intervention phase mean performance was calculated for each student across each measure. For each measure, participants were divided into low-performing, middle-performing, and high-performing groups using baseline means and a .5 standard deviation (*SD*; i.e., low-performing average baseline was $< 0.5 SD$ from the mean; high-performing baseline was $> 0.5 SD$, and middle-performing baseline was $\pm 0.5 SD$). For each measure, a three (groups) \times two (baseline and intervention mean) ANOVA was used to test for significant differences. Tukey's post hoc test was used to specify significant differences. Bonferroni's correction was used to maintain an experiment-wise error rate and alpha level of 0.05.

Treatment Acceptability and Integrity

Data on student and teacher acceptability of the interdependent group contingency were collected on the last day of the intervention, before the weekly drawing. Teachers completed an acceptability form with 16 questions and a 6-point Likert scale for responses (see Appendix A). Students completed an acceptability form with 10 questions and a 4-point Likert scale for responses

(see Appendix B). For both Likert scales, the minimum value represented “strongly disagree” and the maximum value represented “strongly agree.” The students were encouraged to answer the items honestly because the forms were anonymous and there was no right or wrong answer. The primary researcher was available to answer questions while the students and teachers completed the forms.

Each teacher was presented with checklists and asked to monitor the primary experimenter’s implementation across 20% of the sessions where the group contingency was applied. The steps and procedures monitored to ensure consistency were (a) printing a weekly report, (b) excluding names and data of those students without consent, (c) setting up criteria and reward containers, (d) obtaining the class’s attention, (e) informing the class about added or removed rewards, (f) shaking the criteria container, (g) drawing a slip of paper from the criteria container, (h) recording the criterion drawn, (i) comparing the criterion to class performance for that week, (j) determining whether the reward was earned, (k) drawing the reward, (l) recording the reward earned, (m) determining reward implementation and recording it, (n) fulfilling the reward, and (o) reminding the teacher to continue procedures. Average integrity was 98% across all classroom administrations.

RESULTS

Figure 1 shows an increase in the number of quizzes passed for each class immediately after the intervention was implemented. Class A had a stable baseline, but during the intervention phase Class A showed high variability in the number of quizzes passed with an increasing trend followed by a decreasing trend. Figure 1 suggests that the intervention may have had an immediate effect on Class A that was not maintained. Class B had a decreasing baseline trend with low variability. Again, Figure 1 suggests that the intervention may have had an immediate effect on Class B that was not maintained. Class C showed an increasing baseline trend that hinders the conclusions regarding possible intervention effects. The intervention phase data were cyclical, with an immediate increase followed by a rapid decline and another increasing trend.

Groups × Condition Interactions

The effect of the group contingency on quizzes passed, quizzes taken, and book level across high-, middle-, and low-performing groups also was analyzed (see Figures 2, 3, and 4). Students were placed into groups based on their average baseline performance relative to their peers. For each dependent variable, students were placed in the low group if their average baseline performance was $.5 SD$ below the mean and in the high group if their average baseline performance was $.5 SD$ above the mean. The average baseline performance of the middle group fell within $\pm .5 SD$ of the mean. Because there were three ANOVAs run, Bonferroni’s correction was applied. Consequently, to maintain an experiment-wise alpha level of 0.05, differences were considered significant when $p < .017$.

For all students, the baseline mean daily quizzes passed was 0.70 ($SD = 0.81$). Consequently, the low group included 12 students with average daily quizzes passed $\leq .31$. The high group included 6 students with average daily quizzes ≥ 1.14 , and the 14 students whose average daily quizzes passed fell between 0.32 and 1.13 were in the middle group. Figure 2 depicts the mean baseline and intervention quizzes passed for each group. ANOVA revealed a significant main effect for groups, $F(2, 29) = 5.48, p = .01$. Post hoc comparison using Tukey’s test showed that the high group passed significantly more quizzes than did the low group ($p = .01$), but not significantly more than the middle group ($p = .03$). Also, there was no significant difference between the low and middle groups on daily quizzes passed ($p = .80$). ANOVA revealed that the main effect for condition was not significant, $F(1, 29) = 5.38, p < .03$; however, this effect must be interpreted in light of the significant interaction found, $F(1, 29) = 77.85, p < .001$. Figure 2 shows an increase in quizzes

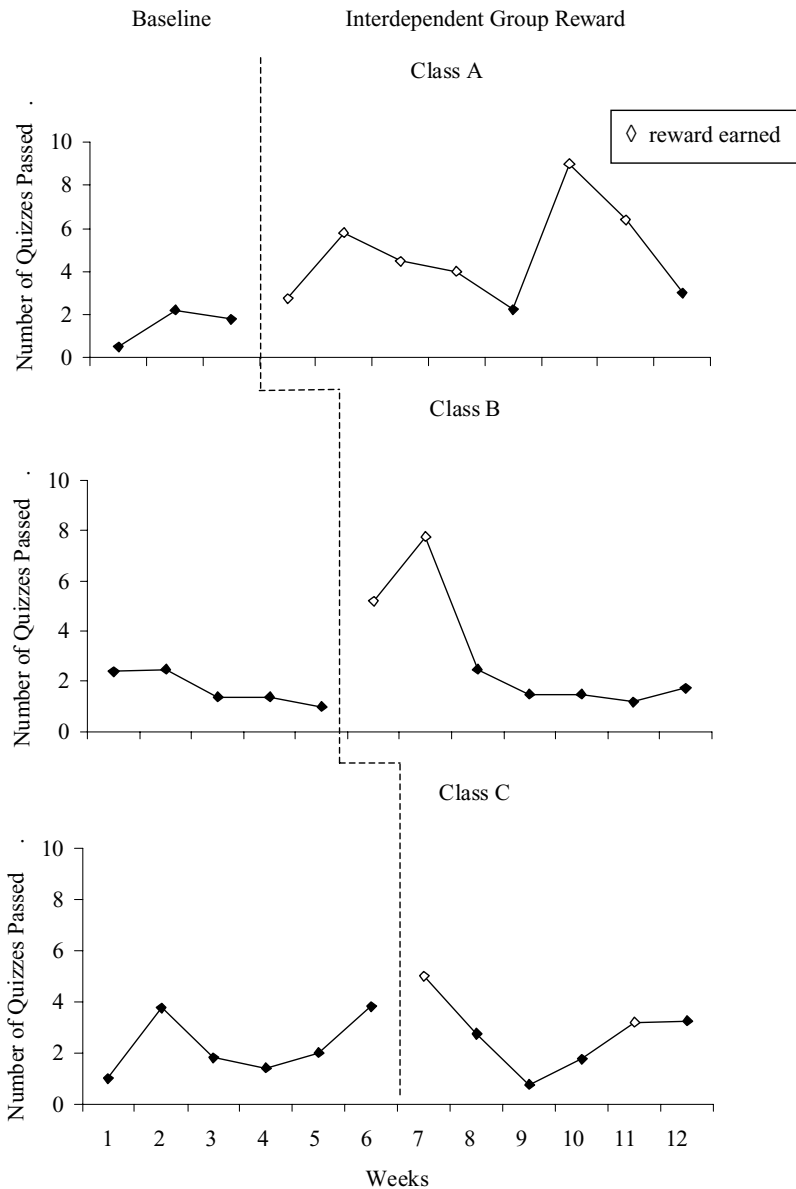


FIGURE 1. Average number of daily quizzes passed each week.

passed for both the low (from .002 to 1.46) and middle groups (from 0.67 to 1.27) but a decrease in quizzes passed for the high group (from 2.12 to 1.69). Pairwise comparisons revealed a significant increase in daily quizzes passed for the low group ($p < .001$) but not for the middle group ($p = .06$). Although the high group had a slight decrease in quizzes passed, this difference was not significant ($p = .38$).

For all students, the baseline mean daily quizzes taken was 1.2 ($SD = 0.89$). Consequently, the

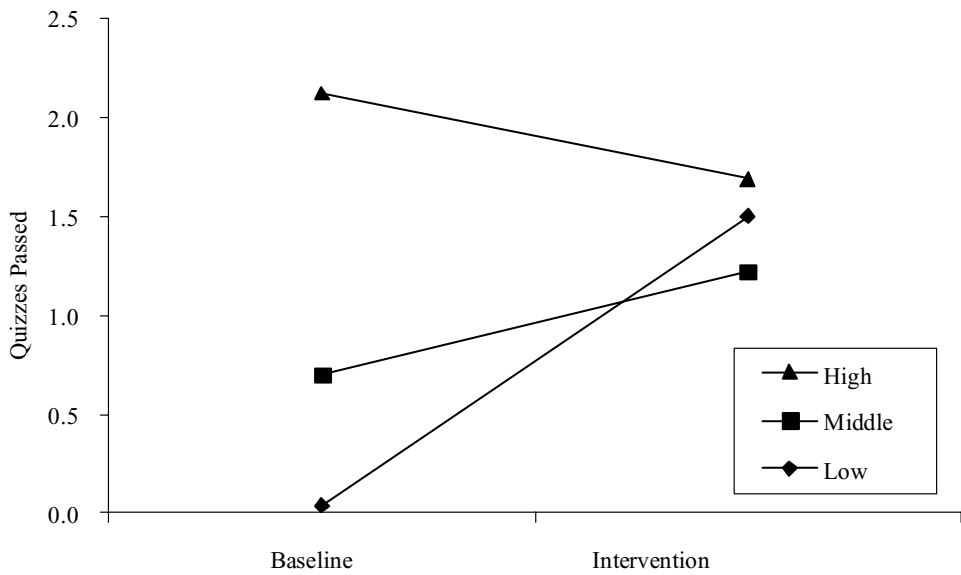


FIGURE 2. Average number of daily quizzes passed during baseline and intervention phase for the low, middle, and high groups.

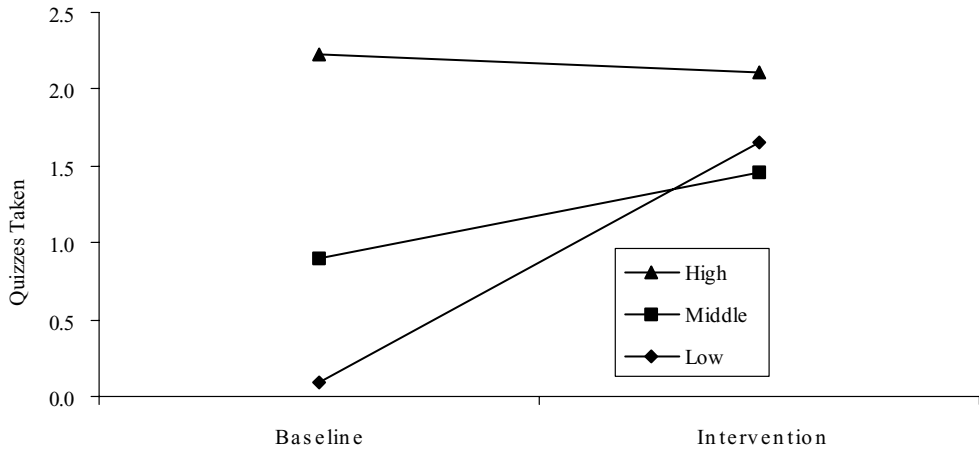


FIGURE 3. Average number of daily quizzes taken during baseline and intervention phase for the low, middle, and high groups.

low group included 13 students with average daily quizzes taken $< .46$. The high group included 7 students with average daily quizzes taken > 1.34 , and the 12 students whose average daily quizzes taken fell between $.47$ and 1.33 were in the middle group. Figure 3 depicts the mean baseline and intervention quizzes taken for each group. Although ANOVA revealed an insignificant main effect for group, $F(2, 29) = 4.10$ ($p = .03$), post-hoc comparisons using Tukey's test showed that the high group attempted significantly more quizzes than the low group ($p = .005$), but no significant differences in quizzes passed were found between the high and middle groups ($p = .04$)

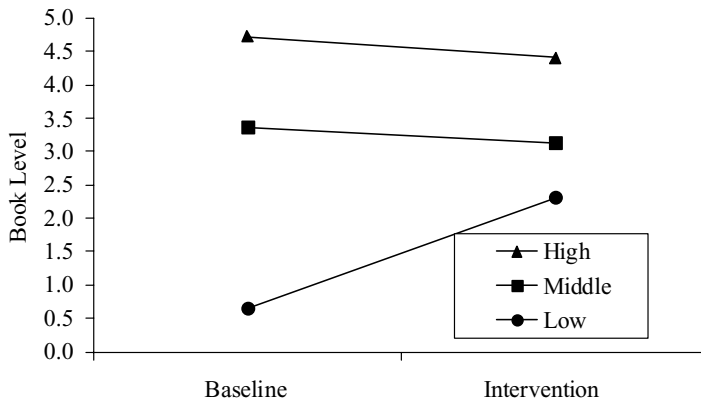


FIGURE 4. Average book level of quizzes taken during baseline and intervention phases for the low, middle, and high groups.

or between the low and middle groups ($p = .62$). ANOVA indicated that the main effect for condition was significant, $F(1, 29) = 7.69$ ($p < .001$), which must be interpreted in light of the significant interactions found, $F(1, 29) = 93.06$ ($p < .001$).

The quizzes taken data are similar to the data for quizzes passed and show an increase in quizzes taken for both the low (from 0.01 to 1.66) and middle groups (0.90 to 1.45), with the low group showing the largest increases, but the high group showing a slight decrease in quizzes passed (from 2.22 to 2.11). Pairwise comparisons revealed a significant increase in daily quizzes taken for the low group ($p < .001$) but not for the middle group ($p = .157$). Although the high group had a slight decrease in quizzes taken, this difference was not significant ($p = .827$).

Although students were encouraged to check out books only from their reading level, they were allowed to select books from higher or lower levels if those books interested them. For all students, the baseline mean book level, based on the quizzes they took, was 3.29 ($SD = 1.02$). The 15 students whose baseline book level was 2.27 or below were placed in the low group, and the 4 students whose baseline book level was 4.33 were placed in the high group, leaving 13 students whose average baseline book level fell between 2.28 and 4.32 in the middle group.

Figure 4 depicts the mean baseline and intervention book level for each group. ANOVA revealed a significant main effect for group, $F(2, 29) = 6.82$ ($p < .004$). Post-hoc comparison using Tukey's test showed that the low group read significantly lower level books than both the middle ($p < .001$) and high group ($p < .001$). Also, the middle group read significantly lower level books than the high group ($p < .006$). ANOVA revealed an insignificant main effect for condition, $F(1, 29) = 1.38$ ($p < .25$), but a significant interaction was found, $F(1, 29) = 468.77$ ($p < .001$). These data show an increase in book level for the low group (from 0.65 to 2.31) and small decreases in book level for the middle (from 3.37 to 3.13) and high groups (from 4.73 to 4.40). Tukey's post hoc, pairwise comparisons revealed a significant increase in average book level for the low group ($p < .001$) but no significant differences for the middle ($p = .577$) and high ($p = .636$) groups.

Acceptability

The teachers' average acceptability score across all items and teachers was 5.09. All items received a positive response (slightly agree to strongly agree) with the exception of the statement, "Most teachers would find this intervention appropriate for academic problems in addition to the AR reading program." One teacher underlined the words, "most teachers," and marked slightly disagree.

The students' average acceptability score across all items and students was 3.43. Across all classes, the statement receiving the highest acceptability rating was, "I'd like to continue the AR reading reward system" ($x = 3.73$). The statement receiving the lowest acceptability rating across all classes was, "The reward system would only have good results" ($x = 3.01$), which may have been influenced by the weeks when no reward was earned. Regardless, these responses indicate strong student and teacher acceptability for the group reward procedure.

DISCUSSION

Sharp and Skinner (2004) used an A-B phase design to evaluate the effects of interdependent group-oriented rewards with randomly selected criteria on AR quizzes passed and found a large increase after the intervention was applied. The current study was designed to extend this line of research by investigating the effects of an interdependent group-oriented reward system with randomly selected rewards and criteria on the AR reading behavior of fourth-grade students across three intact classes using a multiple baseline design to control for threats to internal validity. Across all three classrooms, visual analysis of time-series graphs provide some support for an increase in AR quizzes passed immediately after the intervention was applied. This increase, however, was not maintained throughout the intervention phase. These results were similar to those of Shapiro and Goldberg (1986), who found an immediate increase in sixth-grade students' spelling scores that were not maintained throughout the intervention phase.

When students were divided into three groups (high, middle, and low performers), analyses showed significant increases in reading performance for the low group across all three dependent variables but no significant changes for the middle and high groups. Again, these results are similar to those of Shapiro and Goldberg (1986), who found that the group contingency enhanced the spelling performance of the lowest performing sixth-grade spelling students but not that of the high performers. Together, both the visual and statistical analyses suggest that, although initial treatment effects were not maintained, these treatment effects appeared to be concentrated with the lowest performing students.

If poor readers are to benefit from sustained silent reading time, they must choose to read. Additionally, their reading skills are more likely to improve when they choose to read material that is more challenging. The current study suggests that, when first implemented, the interdependent group-oriented reward may have caused the low-performing students to increase their reading performance (i.e., took and passed more quizzes). Additionally, for the low performers these increases in comprehension occurred as they increased the grade level of the material they were reading. Although these statistically significant findings suggest that the intervention may have caused desirable outcomes with the low-performing students, these findings must be interpreted with caution because some threats to internal validity may have confounded these results.

One concern with selecting groups based on extreme scores is regression to the mean. With respect to quizzes passed, the initial *repeated* measures during baseline show relatively stable responding prior to application of the intervention and *immediate* increases after the intervention was applied, which suggests that regression to the mean does not account for the current findings. Regardless, future researchers should address the possible presence of regression to the mean. A multiple-baseline design where phase changes are made based on the performance of only the low-performing students may control this threat to internal validity.

Another concern was that, after the group contingency was applied, students would begin taking quizzes one after another, without actually reading the material. As quizzes taken and quizzes passed data followed almost identical patterns (compare Figures 2 and 3), the current study suggests that this possible negative side effect did not occur. Also, we were concerned that students would begin to choose easier reading material because it would require less time and effort, thereby increasing

the probability of the students gaining access to rewards. This concern was not supported by our analysis of the middle and high performers which showed no significant differences in reading level. Additionally, with the lower performers we found the opposite effect: Rather than reading lower level material, they read significantly higher level material during the intervention phase. Although this may have been caused by low performers' improved reading skills and/or their desire to read more challenging material, researchers should determine if other variables may have played a role. Specifically, these fourth-grade students may have begun to exhaust lower-level reading material. Although the AR program and the school library had many reading material options for lower grades, because these low-performing fourth-grade students and many other students in the school (e.g., first-, second-, and third-grade students) were reading at these levels, there may have been fewer available texts at the lower levels. Thus, a more restricted range of options may have caused these students to choose more difficult material.

Visual analysis of the time series data shows an increase in quizzes passed followed by a decrease. Future researchers should investigate several variables that may account for this pattern, as identifying the cause of this pattern may allow researchers to modify the interdependent group-oriented reward system so that initial increases are maintained. The initial increase may have been caused by novelty effects. If novelty effects caused the increase followed by the decrease, then the intervention could be modified to make it more effective. In the current study, efforts were made to keep the contingency novel by allowing students to suggest rewards. Unfortunately, only one student suggested a reward that was not included in the pool because it was too expensive. One way to encourage students to suggest rewards may be to have class discussions about possible rewards and/or the teachers could surreptitiously drop acceptable rewards in the suggestion box (Skinner et al., 2009). A related concern is reward quality. Specifically, students may have been highly motivated to read based on the chance to earn a specific, high-quality reward. After that reward was randomly selected, they may have been less motivated to read. If this variable reduced maintenance, then researchers could address it by replacing rewards after they are drawn (Skinner et al., 2004).

The decrease in quizzes passed during the intervention phase also may have been influenced by the length and difficulty of the material that the students chose to read. As the low-performing readers increased the reading level of their material, they may have taken longer to read the material (the material tended to become longer as grade level increased and, because the material was more difficult, students' reading may have been slower), so that the number of quizzes they could take per week decreased. Additionally, they may have been less likely to pass the quizzes because the material was more difficult to read and comprehension questions may have been more difficult. To control for this problem, future researchers should ensure that there are enough books at each grade level. Additionally, researchers could weigh points based on the reading level of the material (e.g., a low performer gets credit for two quizzes passed when he or she passes a quiz on material written above his or her grade level).

Another problem may have been related to the criteria. Although each criterion was randomly selected, after the first intervention week the criteria range was set at 20% above and 20% below the intervention phase average. Therefore, the initial increase in quizzes passed caused a large increase in the criteria needed to earn rewards, which may have decreased student motivation (Skinner et al., 2009). Compounding this problem is that the initial increase in AR quizzes passed may have been the result of finishing books that were started during the baseline phase. Once those materials were completed during the first week (it may have taken a total of 3 weeks), students may not have been able to start a new books and finish them as efficiently. Future researchers could address this applied problem by adding more variability to the criteria. For example, Sharp and Skinner (2004) had criteria that include 1–13 quizzes passed for 13 students.

Class size also may have influenced our results. Shapiro and Goldberg (1990) found that

sample size influences the effectiveness of an interdependent group contingency. Specifically, an interdependent group contingency was more effective at increasing spelling scores in low-performing students when they participated in a smaller group (i.e., group size of 4 vs. a group size of 48). Class A, B, and C had 16, 17, and 18 students, respectively (i.e., total students in each classroom including those not providing permission to participate in the current study). These students may have improved their reading performance more if they were placed into smaller groups (e.g., of 4–5 students); consequently, each student’s performance would have a larger impact on the probability that their group earned a reward.

Summary: It’s a Start

Although the AR program is an efficient and sustainable procedure for monitoring students’ sustained silent reading, it is not likely to be effective in remedying reading skills deficits unless students choose to read AR material. Both the current data and that of Sharp and Skinner (2004) suggest that merely providing reading choices and allowing students to read at their own pace is often not enough, especially for students with weaker reading skills. The current results suggest that supplementing the AR program with interdependent group-oriented rewards may enhance the probability of weaker readers actually choosing to read. These results, however, were short lived. If future researchers can determine which variable(s) account for this immediate but unsustainable change, they may be able to identify procedures that educators can use to enhance the probability of students choosing to read AR books, and consequently enhancing their reading skills.

APPENDIX A
Teacher Acceptability Rating Scale

Academic Intervention Rating Scale							
The procedures used for the AR Reading program is the “intervention” referred to in the questions.							
	Strongly		Slightly		Slightly		Strongly
	Disagree			Agree			
1.	This would be an acceptable intervention for a class with an academic problem.						
2.	Most teachers would find this intervention appropriate for academic problems in addition to the AR reading program.						
3.	The intervention should prove effective in changing the class’s academic behavior.						
4.	I would suggest the use of this intervention to other teachers.						
5.	The class’s AR reading is severe enough to warrant use of this intervention.						
6.	I would be willing to use this intervention in the classroom setting.						
7.	The intervention would <i>not</i> result in negative side effects for the class.						
8.	The intervention would be appropriate for a variety of children.						

(Continued)

9.	The intervention is consistent with those I have used in the classroom setting before.	1	2	3	4	5	6
10.	The intervention is a fair way to handle the class's academic problems.	1	2	3	4	5	6
11.	The intervention is reasonable for the AR reading program.	1	2	3	4	5	6
12.	I like the procedures used in the intervention.	1	2	3	4	5	6
13.	The intervention is a good way to handle this class's academic problems.	1	2	3	4	5	6
14.	Overall, the intervention would be beneficial for the class.	1	2	3	4	5	6
15.	The intervention would produce a lasting improvement in the class's academic behavior.	1	2	3	4	5	6
16.	Soon after using the intervention, the teacher would notice a positive change in the academic problem.	1	2	3	4	5	6

APPENDIX B
Student Acceptability Rating Scale

AR Reading Reward System Rating Scale					
		Strongly Disagree	Disagree	Agree	Strongly Agree
1.	The reward system is good for the AR reading program.	1	2	3	4
2.	I like the AR reward system.	1	2	3	4
3.	I would read more for the AR reward system.	1	2	3	4
4.	Most kids would read more for the AR reward system.	1	2	3	4
5.	The reward system is good for all the kids.	1	2	3	4
6.	I'd like to continue the AR reading reward system.	1	2	3	4
7.	The AR reward program would be OK for other school work.	1	2	3	4
8.	Most kids would find the reward system OK for other school work.	1	2	3	4
9.	The reward system is fair for the whole class.	1	2	3	4
10.	The reward system would only have good results.	1	2	3	4

REFERENCES

- Adams, M. (1990). *Beginning to read: Thinking and learning about print: A summary*. Champaign-Urbana, IL: Center for the Study of Reading.
- Billington, E. J., & Ditommaso, N. M. (2003). Demonstrations and applications of the matching law in education. *Journal of Behavioral Education, 12*, 91–104.
- Billington, E. J., & Skinner, C. H. (2002). Getting students to choose to do more work: Evidence of the effectiveness of the interspersal procedure. *Journal of Behavioral Education, 11*, 105–116.
- Billington, E. J., Skinner, C. H., & Cruchon, N. M. (2004). Improving sixth-grade students' perceptions of high-effort assignments by assigning more work: Interaction of additive interspersal and assignment effort on assignment choice. *Journal of School Psychology, 42*, 477–490.
- Billington, E. J., Skinner, C. H., Hutchins, H., & Malone, J. C. (2004). Varying problem effort and choice: Using the interspersal technique to influence choice towards more effortful assignments. *Journal of Behavioral Education, 13*, 193–207.
- Carter, B. (1996). Hold the applause! Do Accelerated Reader and Electronic Bookshelf send the wrong message? *School Library Journal, 41*(10), 22–25.
- Cashwell, C. S., Skinner, C. H., Dunn, M., & Lewis, J. C. (1998). Group reward programs: A humanistic approach. *The Journal of Humanistic Education and Development, 37*, 47–53.

- Cates, G. L., & Skinner, C. H. (2000). Getting remedial mathematics students to prefer homework with 40% more problems? An investigation of the strength of the interspersal procedure. *Psychology in the Schools*, 37, 339–347.
- Denton, K., & West, J. (2002). Children's reading and mathematics achievement in kindergarten and first grade (NCES-2002-125). Washington, DC: National Center for Education Statistics.
- Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (pp. 205–242). Newark, DE: International Reading Association.
- Dunlap, G., DePerczel, M., Clarke, S., Wilson, D., Wright, S., White, R., et al. (1994). Choice making to promote adaptive behavior for students with emotional and behavioral challenges. *Journal of Applied Behavior Analysis*, 27, 505–518.
- Dyer, K., Dunlap, G., & Winterling, V. (1990). Effects of choice making on the serious problem behaviors of students with severe handicaps. *Journal of Applied Behavior Analysis*, 23, 515–524.
- Friman, P., & Poling, A. (1995). Making life easier with effort: Basic findings and applied research on response effort. *Journal of Applied Behavior Analysis*, 28, 583–590.
- Johnston, R., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in pre-readers: The nature of the relationship. *Reading and Writing: An Interdisciplinary Journal*, 8, 217–234.
- Kern, L., Bambara, L., & Fogt, J. (2002). Class-wide curricular modification to improve the behavior of students with emotional or behavioral disorders. *Behavioral Disorders*, 27, 317–326.
- Lew, M., Mesch, D., Johnson, D. W., & Johnson, R. (1986). Components of cooperative learning: Effects of collaborative skills and academic group contingencies on achievement and mainstreaming. *Contemporary Educational Psychology*, 11, 229–239.
- Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis*, 8, 341–347.
- Mace, F. C., Neef, N. A., Shade, D., & Mauro, B. C. (1996). Effects of problem difficulty and reinforcer quality on time allocated to concurrent arithmetic problems. *Journal of Applied Behavior Analysis*, 29, 11–24.
- Martin, J. J., Skinner, C. H., & Neddrenrip, C. E. (2001). Extending research on the interspersal procedure to perceptions of continuous reading assignments: Applied and theoretical implications of a failure to replicate. *Psychology in the Schools*, 38, 391–400.
- Martin-Palmer, G., Codling, R. M., & Gambrell, L. B. (1994). In their own words: What elementary students have to say about motivation to read. *The Reading Teacher*, 48(2), 176–178.
- McLaughlin, T. F., Herb, C., & Davis, C. (1980). The effects of individual and group contingencies on spelling performance for a special education class. *B.C. Journal of Special Education*, 4(3), 263–270.
- Meadows, S. F., & Skinner, C. H. (2005). Causing students to choose more Language Arts work: Enhancing the educational and external validity of the additive interspersal procedure and discrete task completion hypothesis. *Journal of Behavioral Education*, 14, 227–247.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: National Institutes of Health.
- Neef, N. A., Mace, F. C., & Shade, D. (1993). Impulsivity in students with serious emotional disturbance: The interactive effects of reinforcer rate, delay, and quality. *Journal of Applied Behavior Analysis*, 26, 37–52.
- Neef, N. A., Shade, D., & Miller, M. S. (1994). Assessing influential dimensions of reinforcers on choice in students with serious emotional disturbance. *Journal of Applied Behavior Analysis*, 27, 575–583.
- Pigott, H. E., & Heggie, D. L. (1985). Interpreting the conflicting results of individual versus group contingencies in classrooms: The targeted behavior as a mediating variable. *Child and Family Behavioral Therapy*, 7, 1–15.
- Popkin, J., & Skinner, C. H. (2003). Enhancing academic performance in a classroom serving students with serious emotional disturbance: Interdependent group contingencies with randomly selected components. *School Psychology Review*, 32, 271–284.
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Lee, J. R., Katz, L., Frost, S. J., et al. (2001). Neuroimaging studies of reading development and reading disability. *Learning Disabilities Research & Practice*, 16, 240–249.
- Reading Renaissance I: Using learning information systems to create world-class readers* (1998). Madison, WI: The Institute for Academic Excellence.
- Rego, A. M. (2006). *The alphabetic principle, phonics, and spelling: Teaching students the code*. New York: Guilford Publications.
- Saigh, P. A. (1987). The effects of an academic achievement game on the spelling performance of limited English proficiency students. *B.C. Journal of Special Education*, 11, 73–80.
- Shanahan, T. (2005). *The national reading panel report: Practical advice for teachers*. Naperville, IL: Learning Point Associates.
- Shapiro, E. S., & Goldberg, R. (1986). A comparison of group contingencies for increasing spelling performance among sixth grade students. *School Psychology Review*, 15, 546–557.

- Shapiro, E. S., & Goldberg, R. (1990). In vivo rating of treatment acceptability by children: Group size effects in group contingencies to improve spelling performance. *Journal of School Psychology, 28*, 233–250.
- Sharp, S. R., & Skinner, C. H. (2004). Using interdependent group contingencies with randomly selected criteria and paired reading to enhance class-wide reading performance. *Journal of Applied School Psychology, 20*, 29–46.
- Skinner, C. H. (1998). Preventing academic skills deficits. In T. S. Watson & F. Gresham (Eds.), *Handbook of child behavior therapy: Ecological considerations in assessment, treatment, and evaluation* (pp. 61–83). New York: Plenum.
- Skinner, C. H., Cashwell, C. S., & Dunn, M. S. (1996). Independent and interdependent group contingencies: Smoothing the rough waters. *Special Services in the Schools, 12*, 61–78.
- Skinner, C. H., Pappas, D. N., & Davis, K. A. (2005). Enhancing academic engagement: Providing opportunities for responding and influencing students to choose to respond. *Psychology in the Schools, 42*, 389–404.
- Skinner, C. H., Skinner, A. L., & Armstrong, K. (2000). Shaping leisure reading persistence in a client with chronic schizophrenia. *Psychiatric Rehabilitation Journal, 24*, 52–57.
- Skinner, C. H., Skinner, A. L., & Burton, B. (2009). Applying group-oriented contingencies in classrooms. In K. A. Akin-Little, S. G. Little, M. Bray, & T. Kehle (Eds.), *Behavioral interventions in schools: Evidence-Based positive strategies* (pp. 157–170). Washington, DC: APA Press.
- Skinner, C. H., Skinner, C. F., Skinner, A. L., & Cashwell, T. C. (1999). Using interdependent contingencies with groups of students: Why the principal kissed a pig at assembly. *Educational Administration Quarterly, 35*, 806–820.
- Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review, 33*, 384–397.
- Spaulding, C. L. (1992). The motivation to read and write. In J. W. Irwin & M. A. Doyle (Eds.), *Reading/Writing connections: Learning from research* (pp. 177–201). Newark, DE: International Reading Association.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–406.
- Stewart, J. P., & McLaughlin, T. F. (1986). Effects of group and individual contingencies on reading performance with Native American junior high school students. *Techniques: A Journal for Remedial Education and Counseling, 2*, 133–144.
- Wilson, S. H., & Williams, R. L. (1973). The effects of group contingencies on first graders' academic and social behaviors. *Journal of School Psychology, 11*(2), 110–117.
- Winnett, R. A., Battersby, C. D., & Edwards, S.M. (1975). The effects of architectural change, individualized instruction, and group contingencies on the academic performance and social behavior of sixth graders. *Journal of School Psychology, 13*, 28–40.