To the Graduate Council:

I am submitting herewith a dissertation written by Danielle Nicole Pappas entitled "Interdependent Group Contingencies with Randomly Selected Components Applied to Class-wide Performance in the Accelerated Reader Program." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Education.

Christopher Skinner, Major Professor

We have read this dissertation
and recommend its acceptance:

Sherry Bain

Charles Hargis

Accepted for the Council:

Vice Chancellor and Dean of
Graduate Studies

Thesis
2006b
.P35

INTERDEPENDENT GROUP CONTINGENCIES WITH RANDOMLY SELECTED COMPONENTS APPLIED TO CLASS-WIDE PERFORMANCE IN THE ACCELERATED READER PROGRAM

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Danielle Nicole Pappas
August 2006

Copyright © 2006 by Danielle N. Pappas
All rights reserved.

DEDICATION

This dissertation is dedicated to my family, Dave, Devanee, and Drew Pappas, and my mother, Donna Bleiler. To my husband, Dave, whose confidence in me has inspired me to reach for higher goals than I ever thought possible and whose unending love has supported me throughout my entire graduate career. To Devanee for her willingness to help her mother in all activities including the dissertation research and practicum practice. To Drew, whose impending arrival helped drive this research project. To my mother, for always believing in me, supporting me, and instilling confidence in me to achieve my goals.

ACKNOWLEDGMENTS

I would like to thank all those who helped me complete my Doctorate of Philosophy degree in Education, with a concentration in School Psychology. I would like to thank Dr. Christopher Skinner for his guidance, direction, support, and assistance throughout my graduate career and especially this dissertation research project. I am grateful to each individual who served on my committee: Dr. Sherry Bain, Dr. Richard Saudargas, and Dr. Chuck Hargis for serving on my dissertation committee.

I would also like to thank the Knox County School System. Working with the teachers, administrators, and students at Christenberry Elementary School was a pleasure. A special thank you to the 4th grade teachers who allowed me to take over their classrooms without hesitation during this project.

Lastly, I would like to thank my family and friends, whose encouragement and support made this achievement possible.

ABSTRACT

The Accelerated Reader (A/R) program is a computer-based reading program employed by numerous school systems that provides comprehension quizzes over 10,000 books that have been rated for difficulty level. The quiz information is used to provide feedback to educators. The A/R program is designed to incorporate the elements of choice and reinforcement. Despite the research that supports the use of choice and reinforcement to increase academic behaviors, the A/R reading program has not been successful in increasing the sustained silent reading behaviors of all students enrolled in the schools using this program.

The use of group contingencies has been effective in increasing academic behaviors. The current study was designed to extend this research by investigating the effects of a group contingency on the reading behavior of 4th grade students.

Thirty-two students across three 4th grade intact classes participated. The students continued to read A/R books and take quizzes as was the current practice. A multiple baseline design was employed to implement an interdependent group contingency with randomized components in each classroom. The intervention involved providing the entire class with a group reward based on the class-wide performance on the A/R quizzes. At the end of each week, a randomly selected criterion was compared to the class performance on the number of A/R quizzes passed. A randomly selected reward was implemented if the class performance exceeded the criterion.

Visual analysis of time-series graphs provides some support for an increase in A/R quizzes taken and passed immediately after the intervention was applied. This

increase was not maintained. Only one class showed a clear increase in percent of comprehension questions answered correctly after the group contingency was applied. Across all three classes there were no clear changes in reading level of books selected after the intervention was applied. 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, percent correct, and book level). However, the middle and high performing students did not show any statistically significant change.

TABLE OF CONTENTS

Chapter		Page
1.	**INTRODUCTION**	**1**
	Choice	3
	Why Providing Choices Works	4
	Empowerment	4
	Interest	5
	Effort	5
	Reinforcement	6
	Group Contingencies	6
	Purpose	15
2.	**METHODOLOGY**	**17**
	Participants and Setting	17
	Materials	18
	Procedure	19
	Experimental Design, Dependent Variables, and Data-Analysis Procedures	21
	Treatment Integrity	25
3.	**RESULTS**	**26**
	Visual Analysis of Time-Series Graphs	26
	A/R Quizzes Passed	26
	A/R Quizzes Taken	26
	Book Level	29
	Percentage Correct	29
	Effect Size Analysis	32
	A/R Quizzes Passed	32
	A/R Quizzes Taken	33
	Book Level	33
	Percentage Correct	34
	Summary of Visual and Effect Size Analyses	35
	Statistical Analysis by Groups	36
	A/R Quizzes Passed	36
	A/R Quizzes Taken	39
	Book Level	44
	Percentage Correct	47
	Summary of Analysis by Groups	51
	Teacher and Student Acceptability	51
	Teacher Acceptability	52
	Student Acceptability	52
4.	**DISCUSSION**	**55**
	Summary	55

Limitations and Future Research..... 57
Summary..... 63

LIST OF REFERENCES..... 66

APPENDIX..... 74

VITA..... 79

LIST OF TABLES

Table Page

1. Teacher Acceptability Rating Scale..... 23
2. Student Acceptability Rating Scale..... 24
3. Quizzes Passed, Means, Standard Deviations, and Effect Sizes..... 32
4. Quizzes Taken, Means, Standard Deviations, and Effect Sizes..... 33
5. Average Book Level, Means, Standard Deviations, and Effect Sizes.....34
6. Average Percent Correct, Means, Standard Deviations, and Effect Sizes.....35
7. Tukey's Post Hoc Tests, Multiple Comparisons.....37
8. Main Effect and Interaction Effects..... 38
9. Average Quizzes Passed, Means, Standard Deviations, and Effect Sizes.....38
10. Tukey's Post Hoc Tests, Pairwise Comparisons.....40
11. Tukey's Post Hoc Tests, Multiple Comparisons.....41
12. Main Effect and Interaction Effects..... 42
13. Average Quizzes Taken, Means, Standard Deviations, and Effect Sizes..... 42
14. Tukey's Post Hoc Tests, Pairwise Comparisons.....44
15. Tukey's Post Hoc Tests, Multiple Comparisons.....45
16. Main Effect and Interaction Effects..... 46
17. Average Book Level, Means, Standard Deviations, and Effect Sizes..... 46
18. Tukey's Post Hoc Tests, Pairwise Comparisons.....46
19. Tukey's Post Hoc Tests, Multiple Comparisons.....48
20. Main Effect and Interaction Effects..... 49
21. Average Percent Correct, Means, Standard Deviations, and Effect Sizes.....50
22. Tukey's Post Hoc Tests, Pairwise Comparisons..... 50
23. Teacher Responses on the Acceptability Scale..... 53
24. Student Responses on the Acceptability Scale..... 54

LIST OF FIGURES

Figure Page

1. Average daily quizzes passed by class..... 27
2. Average daily quizzes taken by class..... 28
3. Weekly average book levels by class..... 30
4. Weekly average percent correct by class..... 31
5. Quizzes passed by low, middle, and high groups..... 40
6. Quizzes taken by low, middle, and high groups..... 43
7. Book level by low, middle, and high groups..... 47
8. Percent correct by low, middle, and high groups..... 50
9. Average daily quizzes passed by group and by class..... 75
10. Average daily quizzes taken by group and by class..... 76
11. Weekly average book levels by group and by class..... 77
12. Weekly average percent correct by group and by class..... 78

CHAPTER 1: INTRODUCTION

The United States of America is evolving from an industrial nation to an information-based economy (Stuckey, 1991). This evolution places emphasis on knowledge stores and information technology. Literature is the means to learning the required knowledge necessary for this new economy. Personal literacy becomes an intricate part of our status in society, job opportunities, and income (Wilkinson, 1999). Because of this change in our society, media, politicians, and the general public have called for an improvement in reading skills of young people (Green, Hodgens, & Luke, 1997).

The purpose of reading is to comprehend the text. Durkin (1993) characterizes comprehension as the "essence of reading." Comprehension is crucial to academic success and life long learning. There are many techniques and skills necessary to increase reading comprehension. Phonemic awareness (e.g., segmenting and blending), fluent reading (rapid and accurate decoding), prediction, the ability to think-aloud, text structure awareness, the ability to create graphic representations of text, and summarization skills may enhance comprehension (Duke & Pearson, 2002).

While various procedures have been developed to enhance these skills, none will be effective unless students choose to read. Researchers have found that two strategies are effective for increasing the probability of students choosing to engage in academic behaviors (Dunlap, et al., 1994; Popkin & Skinner, 2003). First, instead of assigning one academic task, researchers have found that students are more likely to choose to engage in academic tasks when they are given a choice of tasks (Dunlap, et al., 1994). Second,

strengthening reinforcement for engaging in academic tasks can enhance the probability of students choosing to do that task (Popkin & Skinner, 2003).

A popular trend in independent silent reading curriculum is computerized reading programs such as Accelerated Reader and Electronic Bookshelf. Although, we investigated a procedure designed to enhance student reading via the Accelerated Reader (A/R) program, the Electronic Bookshelf functions basically the same way (Carter, 1996). A/R is based on the premise that the primary obstacle to increasing reading skills is the lack of reading practice (Reading Renaissance I, 1998).

The Report of the National Reading Panel (2000) suggests that independent silent reading can help students develop comprehension skills necessary for success. While educators or parents can allot time for silent reading, skills are unlikely to improve unless students choose to read (Skinner, Pappas, & Davis, 2005). The A/R program is designed to encourage more reading practice through providing a list of reading level books (i.e., to make book selection easy), comprehension quizzes, and teacher reports for continual monitoring and feedback. Students choose from over 50,000 published books. Once a student completes the book, he/she takes a multiple-choice comprehension quiz that assesses his or her reading comprehension. The computer then calculates a quiz score from the items answered correctly and awards points based on the difficulty level of the text. A passing score is 60% accuracy on the comprehension quizzes (Reading Renaissance I, 1998).

The A/R program is designed to incorporate the elements of choice and reinforcement for reading activities. A/R is designed to give students the opportunities to choose their own silent reading books from the school library, classroom library, or home

resources. When an A/R quiz is passed, the student is awarded points. The point system can be used to determine eligibility for rewards. This is an independent group contingency because all students have access to the rewards based on the same criterion.

Choice

Researchers have shown that allowing students with emotional-behavioral disorders (EBD) to choose assignments can reduce problem behaviors and increase desired academic behaviors in school settings (Dunlap et al., 1994; Dyer, Dunlap, & Winterling, 1990; Kern, Bambara, Fogt, 2002). Dunlap et al. (1994) used withdrawal designs (i.e., A-B-A-B phases) to determine if providing choices can influence the behavior of elementary students diagnosed with EBD. Two 11-year-old boys participated in the first experiment during their general education English and Spelling class. During the baseline phases (A phases), academic tasks were assigned by the teacher. During the intervention phases (B phases), students were provided a menu with several assignment options and allowed to choose their assignments. Results showed both an increase in task engagement and a decrease in disruptive behavior during the choice phases of the study.

In the second study, during the intervention phases (B phases), Dunlap et al. (1994) allowed a 5-year-old boy to choose the book the teacher would read to him during one-on-one reading time. In the A phase, the teacher chose the book. Results showed an increase in task engagement and a decrease in disruptive behavior during the choice phases of the study.

Dyer, Dunlap, and Winterling (1990) investigated choice making as an intervention to reduce problem behaviors in students with severe handicaps. Each of the three participants had severe cognitive impairments and exhibited serious behavioral

problems such as aggression, self-injury, and tantrums. The participants ranged in age from 5 to 11 years old. Investigators used an ABAB design. During the A phases, both the tasks and reinforcers were selected by the teacher. During the B phases, students were allowed to choose tasks and reinforcers that were consistent with the ongoing curriculum of each individual participant. The tasks were prevocational or preacademic activities that the students had demonstrated the ability to complete accurately. Although no differences in response accuracy occurred across experimental phases, the results showed lower rates of problem behaviors (e.g., aggression) during choice phases.

Kern, Bambara, and Fogt (2002) modified class-wide science curriculum to investigate the effects of choices on academic task engagement and destructive behavior of six junior high school students diagnosed with EBD. Researchers employed an alternating treatment design (i.e., ABAB). During the A phases, all tasks were assigned by the teacher. During the B phases, students were provided two choice opportunities during the science class period. At least one of those choice opportunities included an activity that was a high interest activity as determined by the teacher. Results showed a significant increase in academic engagement and a decrease in destructive behaviors during the choice phases.

Why Providing Choices Works

Empowerment

Providing choices in academic environments instills a sense of empowerment and control in the students. Empowerment and control are two personal characteristics that underlie intrinsic motivation and should increase academic engagement (Spaulding, 1992). It makes intuitive sense to provide opportunities for choice making during some

reading activities. Allowing students to choose books that interest them should increase the student's engagement in sustained silent reading activities (Martin-Palmer, Codling, & Gambrell, 1994).

Interest

Providing choices in academic environments strengthens the reinforcing element of satisfying personal interests (Spaulding, 1992). According to cognitive psychology, self-generated goals such as a desire to learn more about a certain topic or incongruous information that produces uncertainty is a representation of a problem (Gagne, Yekovich, & Yekovich, 1993). To resolve the problem, individuals are motivated to learn or find a solution. With respect to reading, the cognitive dissonance associated with a self-generated desire to learn about a topic, may lead a student to read a particular text. Martin-Palmer, Codling, & 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.

Effort

When providing choices in academic environments, educators should be aware of another principle that mediates students' choice behavior. When the reinforcement is equivalent across behaviors (e.g., quality, rate, and immediacy), students are more likely to choose the behavior that requires the least amount of effort (Billington & Ditommaso, 2003; Billington & Skinner, 2002; Cates & Skinner, 2000). This principle means that when providing choices to students in an academic environment, the student will choose the assignment that requires the least effort. Thus, assignments that are more challenging

and may provide better skill development may not be selected (Billington, Skinner, Hutchins, & Malone, 2004). Providing unequal assignment choices (i.e., low-effort assignments compared to high-effort assignments), may reduce the learning rates if the high-effort assignments results in greater skill development (Skinner, Pappas, & Davis, 2005). This principle creates a challenge to the A/R program. The A/R program is designed to allow students to choose their own books. However, according to the principle of least effort, students are more likely to choose to read texts that require less effort.

Reinforcement

Student academic behavior is directly related to the reinforcement for that behavior (Skinner, Williams, Neddenriep, 2004). Educators can increase the probability that a student will engage in assigned tasks, even low-interest and/or high-effort academic behaviors by increasing the rate of reinforcement, increasing the quality of reinforcement, making the reinforcement more immediate, or reducing the reinforcement for competing behaviors (Neef, Mace, & Shade, 1993; Neef, Shade, & Miller, 1994; Skinner, Williams, & Neddenriep, 2004). Group contingency programs are one way an educator can enhance reinforcement for A/R reading comprehension.

Group Contingencies

Researchers have identified three types of group contingencies; independent, dependent, and interdependent (Litow & Pumroy, 1975). Numerous researchers have shown that these contingencies can be used to alter students' behavior. Researchers also have identified strengths and weaknesses associated with using these contingencies to influence the behavior of students.

With independent group contingencies, each student in a group (e.g., a class) receives access to the same consequence(s) based upon his/her own behavior (independent) meeting the same criterion (Litow & Pumroy, 1975; Skinner, Williams, & Neddenriep, 2004). Independent group contingencies are often used in academic settings. One example is the grading system in effect in most schools. A student earning 90% or higher on an assignment or in a subject area receives an A and a student earning 80% to 89% on an assignment receives a B. This is considered an independent group contingency because the consequences (grades) and criterion (percent correct on assignment) are the same for each member of the group (class). It is considered independent because each student receives access to the consequence base on his/her own performance (Litow & Pumroy, 1975).

Independent group contingencies appear to be fair to all students because the consequence and criterion are held constant. However, this constancy rules out the ability to alter the target behaviors (e.g., assignments), criteria, and consequences based on individual student's idiosyncratic levels of skill development or preferences (Skinner, Williams, & Neddenriep, 2004).

When access to consequences is public, independent group contingencies provide students with an indication of how their classmates performed. Classmates know the students who met a criterion and received access to a consequence and those who did not. Additionally, because the criterion is the same for everyone, students often know the specific criterion that their peers met or failed to meet. This can have negative social side effects. For example, students who receive access to rewards may be identified by their classmates as smart or hard working. However, when students observe other classmates

not receiving access to rewards, they may consider them dumb or lazy (Cashwell, Skinner, Dunn, & Lewis, 1998). This public feedback with respect to classmates' performance can be detrimental to the social fabric of the classroom.

With dependent group contingencies all students in a group receive access to the same consequence based on the performance of a designated individual or group of individuals meeting a criterion. The group's access to rewards is dependent on the behavior of an individual or smaller group of individuals (Smith & Misra, 1994). An example of this contingency would be when the entire class earns a reward based on the spelling performance of Johnny.

With dependent group contingencies, peer pressure can support appropriate behaviors and censure inappropriate behaviors. Additionally, these contingencies do not require students to compete, can increase group cohesion, and reduce disruptive behavior (Gresham & Gresham, 1982). However, with dependent group contingencies, there is tremendous pressure on the selected individuals (Romeo, 1998), it appears to be unfair to the majority of students (Skinner, Williams, & Neddenriep, 2004), and there is an increased probability of peers threatening students whose performance is targeted. When the target student(s) meets the criterion, classmates may praise the student. However, when they fail to meet the criterion, peers may punish or ostracize the target student(s) (Cashwell, et al., 1998).

With interdependent group contingencies the entire group (e.g., class) receives access to the same consequence contingent upon the group's behavior meeting a group-oriented criterion (Litow & Pumroy, 1975; Slavin, 1987). One example of this type of contingency is when the entire class earns a reward if the entire class's spelling quiz

average was above 80%. Each student's performance contributes to the group's average. This establishes interdependency where each students access to consequences is influenced by his/her own performance and the performance of his/her classmates.

With interdependent group contingencies, the class is not divided into those who receive access to consequences and those who do not. Working towards a common goal may increase cooperative and prosocial behaviors among classmates (Pigott & Heggie, 1985). For example, a student who finishes work early may help a peer with his/her work in order to increase the probability of the entire group earning access to a reward. Another advantage of interdependent group contingencies is that it is easy for school personnel to administer and manage the delivery of rewards (Cashwell et al., 1998). When an entire group learns that they met a criterion, they are likely to engage in social behaviors that serve as additional reinforcement for the behavior (e.g., a group cheer). However, when a class learns that they did not earn a reward, the entire group could respond negatively. Another related limitation is that when students who perform well do not earn the reward, they may become upset and even aggress against the students they perceive to be responsible for the loss of reward (Cashwell et al., 1998).

Group contingencies have been used to reduce socially inappropriate behaviors and increase academic behaviors. Researchers reviewing the literature typically look at each type of contingency and try to determine the effectiveness of the type of contingency employed. However, Pigott and Heggie (1985) reviewed the group contingency literature by dividing the research into target behavior categories of social behaviors or academic behaviors. In this review, the use of group contingencies to reduce socially inappropriate behavior was inconsistently effective. Also, the reduction of

socially inappropriate behavior was accompanied by negative side effects such as increased negative social interactions between peers (i.e., threats). In the academic behavior category, the group contingency was consistently superior to an individual contingency in increasing academic behavior. The academic interventions were not accompanied by the negative side effects found with social behaviors. Pigott and Heggie (1985) suggested that the success of the studies using group contingencies for academic behaviors was due to an increase in cooperative behavior, an increase in peer social reinforcement, and a reduction in the negative effects any one student can have on the overall level of reinforcement.

There is a multitude of research articles using group contingencies to decrease socially inappropriate behavior. Studies have shown that group contingencies are effective at modifying the behavior of a hyperactive child (Patterson, Jones, Whittier, & Wright, 1965), decreasing the rate of disruptive behavior (Crouch, Gresham, & Wright, 1985; Gresham & Gresham, 1982; Hall et al., 1971, Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Madsen, Becker, Thomas, Koser, & Plage, 1968; Theodore, Bray, Kehle, & Jenson, 2001), decreasing uncontrolled verbalizations (Davies & Witte, 2000), and decreasing noise levels in a lunchroom (Davey, Alexander, Edmonson, Stenhoff, & West, 2001).

There have been fewer studies investigating the effects of group contingencies on academic behaviors. Researchers have shown an improvement in academic behavior including 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, 1997), and reading (Sharp & Skinner, 2004; Stewart & McLaughlin, 1986).

Lew, et al. (1986) investigated the effects of collaborative skills and academic group contingencies on vocabulary achievement and mainstreaming. In this empirical study, the authors investigated four group conditions including (a) an opportunity to interact with classmates, (b) use of a positive goal interdependence, (c) use of a positive goal interdependence with a collaborative-skills group contingency, and (d) use of a positive goal interdependence with both collaborative skills and academic group contingencies. There were 83 participants from the 8th grade and four 8th and 11th grade students identified as social isolates. The results showed the most positive relationships, frequent collaborative skills, and the highest vocabulary achievement in the positive goal interdependence with both collaborative skills and academic group contingency condition.

Shapiro and Goldberg (1986) compared independent, interdependent, and dependent group contingencies for spelling performance. Using an alternating treatment design, 53 participants from the 6th grade were exposed to the three different group contingencies targeting spelling. Results initially showed a greater increase in spelling scores for the independent group contingency. However, that result did not continue and the rapid decline for the independent group contingency category resulted in no significant differences across any of the group contingency conditions. All group conditions showed a significant improvement over baseline with the greatest improvement among the lowest achieving students.

Saigh (1987) investigated an academic achievement game utilizing an interdependent group contingency to improve the spelling performance of students in an English as a Second Language (ESL) class. There were 32 participants in the 4th grade. An ABAB design was employed. During the B phase of the design, the class was divided into two teams that competed for the rewards. Results showed a significant improvement in spelling scores during the group contingency phases.

McLaughlin, Herb, and Davis (1980) investigated the effects of individual and group contingencies on spelling performance in a special education classroom. A multi-element design with alternating experimental conditions was employed. There were 10 participants that ranged in age from 10 to 12 years and ranged in standardized spelling performance from kindergarten level to grade 3.2. Results showed an increase in spelling accuracy in the group contingency phase above both baseline and individual contingency conditions.

Winett, Battersby, and Edwards (1975) looked at the effects of an architectural change in the classroom, individualized instruction, and academic group contingencies on social behavior and academic performance in math and language. Although behavioral data was only collected on 10 pre-selected students, academic performance data was collected on all 27 participants in the 6th grade. An alternating treatment design was employed. The results showed that the highest academic work completion and improved social behavior was achieved in the individualized instruction with the group contingency condition.

Wilson and Williams (1973) applied group contingencies in the 1st grade and analyzed its effects on academic and social behavior. There were 110 participants in the

4th grade that were divided into groups of 9-12 students each. An ABAB design was employed. The dependent variables included writing assignment accuracy and inappropriate behavior such as time off-task, motor activity, nonverbal noise-making, inappropriate verbalizations, and aggressive responses. Results showed a significant increase in work completion and appropriate behavior during the group contingency condition.

Stewart and McLaughlin (1986) investigated the effects of group and individual contingencies on reading performance. There were nine participants from a junior high school. All participants were Native American from the Peigan Indian tribe and ranged in age from 11 to 15. Using an ABAB design, the researchers investigated seven dependent variables. The results showed that both the individual and the group contingency phases produced (a) a higher number of words read orally correct, (b) an increase in the number of students participating in oral reading, (c) an increase in the number of pages read silently, and (d) comprehension quiz scores of at least 85%. However, there was no significant difference found between the individual and group contingency condition.

Popkin and Skinner (2003) used an interdependent group contingency program with five students with serious emotional disturbance. The students ranged in age from 11 to 14 years old. The group contingency program was applied first to spelling assignments, second to mathematic assignments, and lastly to English assignments so that by the end all three types of assignments were included in the program. The program included randomly selected target behaviors, criteria, and rewards. The results showed increases in academic performance for the target assignments.

Sharp and Skinner (2004) investigated the effect of an interdependent group contingency coupled with class-wide paired reading on group reading performance. There were 13 African American participants in a 2nd grade class. An AB design was employed. A reversal phase was planned but rejected by the teacher. The B phase involved a paired reading program where the students read at least one A/R book with a partner. Once the pair of students passed an A/R quiz, they had the option of continuing to read together or separately. The B phase also employed the interdependent group contingency. The entire class earned access to a reward based on the number of A/R quizzes passed by the entire class at the end of the week. The criterion and the reward were randomly selected. The results showed a significant increase in the number of A/R quizzes passed per week for the entire class during the B phase.

One of the challenges for group contingencies applied to academic behaviors is setting appropriate criteria. If the criterion is set too low, the students may only put forth the minimum effort to meet the criteria, but not give their best effort. If the criterion is set too high, the students may get frustrated attempting to reach the goal without receiving rewards (Skinner, Skinner, & Armstrong, 2000). One solution to this dilemma is randomizing the criterion (Skinner, Cashwell, & Dunn, 1996). By randomly selecting the group-oriented criterion, the students do not know in advance the minimal effort required to reach the criterion. This component is designed to maximize student effort toward an unknown criterion.

Four studies have employed the randomized criterion component with group contingencies. Sharp and Skinner (2004), previously reviewed, found that randomizing the criterion resulted in an increase in the class-wide reading quizzes passed per week.

Popkin and Skinner (2003) used randomly selected criteria to increase class-wide spelling, mathematics, and English achievement. Two studies adapted the Good Behavior Game to show that randomizing the criteria and/or target behaviors decreased inappropriate social behaviors (Kelshaw-Levering, et al., 2000; Theodore, et al., 2001).

Another challenging component of group contingencies is the selection of rewards. Some rewards are more reinforcing to some students than to others. If the reward is not reinforcing or is adverse to one student, that particular student may intentionally attempt to prevent the class from earning the reward (Kelshaw-Levering, et al., 2000). Using an interdependent group contingency reduces the negative effects one student may have on the class's ability to meet the criterion (Skinner, Williams, & Neddenriep, 2004).

Randomizing the rewards reduces the possible negative side effect of reward selection. Because the students are not aware of the reward prior to the randomized selection, those students who find one or more of the rewards adverse, will not be able to intentionally sabotage the classes ability to earn that reward. By randomly drawing the reward after the group has earned one, it minimizes the potential of a reward to negatively influence the target behavior of particular students (Popkin & Skinner, 2003).

Purpose

The purpose of the current study was to extend the research on interdependent group contingencies with randomized components and increase sustained silent reading behaviors for 4th grade students as measured by A/R quizzes. The students in this study were exposed to an interdependent group contingency for the number of A/R reading quizzes passed per week. The criterion for earning a reward would be randomly selected.

The reward would be randomly selected. The random selection of the criterion and the reward took place on the last day of the week. An analysis would be conducted to determine if there's a significant increase in number of A/R quizzes taken, number of A/R quizzes passed, percent of comprehension questions correct, and reading book level per week for the entire class.

CHAPTER 2: METHODOLOGY

Participants and Setting

Participants were 51 students enrolled in an inner-city school in the southeastern United States. It was a public school and served children in kindergarten through 5th grade. The students enrolled in this school were 74% Caucasian, 21% African American, and 4% other minorities. The school provided free or reduced lunches to 90% of the enrolled students. Participants ranged in age from 8 to 11, and were in the 4th grade.

Approval was obtained from the school district's research coordinator, the school principal, and the University Institutional Review Board prior to soliciting informed consent for participation. Following the approval process, 4th grade classroom teachers at the school were informed of the research project and asked to participate. Three teachers were interested and volunteered their classrooms. The teacher participants ranged in years of teacher experience from 1 to 20 years. All students enrolled in the three participating classrooms were given parental permission letters to return to their teachers with parental signatures. All students in the three participating classrooms were included in the intervention. The parental consent was for inclusion of the student's data in the research project. Nineteen parents did not consent to participation or did not return consent forms. The data for those students were excluded in evaluation of the intervention. This resulted in 32 students participating throughout the entire study. At the introduction of the intervention, a written assent was obtained from all students in each classroom.

Materials

Students were asked to continue participating in the A/R reading program already in progress at the school. Each student chose from over 50,000 published books listed in the A/R system to read for leisure activity. There is a minimum of 100 books per reading level in the school library. Students were allowed to check out a library book, return it, and check out another book. The classrooms also contained small libraries of A/R books. When the student completed reading a book, he/she took a multiple-choice comprehension quiz of 10 questions. The quiz scores were calculated by the computer and maintained in a database. Students were not permitted to retake a quiz for a book they had already passed (60% accuracy). The books in the A/R system were rated by level of difficulty. The teachers printed a weekly report that showed the A/R activity for her class that week. A 4th grade reading book is a short-chapter book, less than 200 pages long, in a 5-inch by 7-inch format with double spacing between lines and large print. The student participants in this study read between 1st grade and 6th grade reading levels.

The intervention materials included containers for holding slips of paper. Three disposable plastic storage containers were purchased for each classroom. One container held the slips of paper with the criterion for earning the reward, one container held the slips of paper with the rewards on them, and one container was for student reward suggestions. Colored cardstock was used for the criteria, reward, and suggestion slips.

Before drawing the criterion for earning rewards, the classroom's A/R quizzes passed for the week were printed using a computer. Any school computer connected to a printer worked. Most reports were printed in the school library. The computer report listed all classroom students, the number of quizzes taken that week, the number of

quizzes passed that week, book level average, average percent correct, points earned in the A/R system, diagnostic codes for at-risk students, and the percent of reading that was fiction. Of these items, the data used in the current study included the total quizzes passed, total quizzes taken, book level, and percent correct. The student's names were blackened and replaced with random numbers to ensure confidentiality. The weekly totals were adjusted by hand to exclude the nonparticipating students.

Procedure

Participants were introduced to the intervention on Monday. They were told that on Friday, the number of quizzes passed for the entire classroom would be checked. This number would then be compared to a number drawn from the criteria container. If the class exceeded the number on the criterion slip of paper, a reward would be drawn. The reward would be delivered within a week. Students were not informed of the required criterion or the reward since the drawing did not occur until Friday. However, the students were informed of the range of the criterion and a list of rewards available in the drawing. The students were asked to continue reading in the A/R reading program.

The criterion for the first week of the intervention was rigged to ensure the class earned the reward. 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 average. For example, if the intervention quizzes passed average was 20 for 3 weeks of intervention, the criterion included would have been 16 through 24. The class was informed on Monday what the range of criterion would be for that week.

The rewards were chosen by teacher input. 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 an (a) ice cream party, (b) 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. A student would write his/her suggestion on a blank slip of paper kept in the suggestion box, and leave it in the 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 into the intervention, it was announced to the class at the beginning of the week. Only one reward suggestion was received throughout the intervention phase. 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 reading activities when the students were encouraged to read A/R books. In addition, students were encouraged to read their A/R books when independent seatwork was completed. Upon completion of an A/R book, students were encouraged to take the computer quiz as soon as possible on the classroom or library computers.

On Friday of each week, a computer report from the A/R program was printed. The students were told the number of quizzes they had passed that week. Then a student was asked to draw one slip of paper out of the criteria container and read the number aloud. If the students exceeded the criterion number, another student was asked to draw a

slip of paper from the reward container and read it aloud. Then the teacher and principle investigator decided upon a day to deliver the reward. Rewards were delivered within the next week. If the students did not exceed the criterion number, a reward was not drawn, the students were encouraged to continue reading in the A/R program, and the students were reminded that a new drawing would occur again the next Friday.

Experimental Design, Dependent Variables, and Data-Analysis Procedures

A multiple baseline design across classrooms was used to compare students' performance from baseline to intervention. The participating teachers negotiated the order of implementation in the three classrooms. The intervention was implemented in Class A for two weeks. While Class A continued the intervention, Class B began the initial implementation of the intervention, Class A and Class B continued for one week, and then Class C began the initial implementation of the intervention. The intervention continued in all classrooms for an additional 6 weeks for a total of 9 weeks of intervention from the initial implementation in Class A.

Data were collected and analyzed each week for each classroom including 4 weeks of baseline prior to Class A implementation of the intervention. The dependent variables were the total number of A/R quizzes passed per week, total number of A/R taken per week, average percent correct on A/R quizzes, and reading book level. Quizzes not completed by the time of the criterion drawing on Friday were not included in the data. Students whose parents did not consent to participation were excluded from the data.

There are limitations with this study related to school attendance. The group contingency intervention depends on regular attendance of the students in order to take

the computer generated quizzes. If school-wide attendance falls below 75% for the week, the data for that week will not be included in this study. Data showed the lowest weekly attendance was 87% while the average during intervention was 92%. Therefore, all intervention weeks were included in the data collection.

A related limitation is the number of school days in a week. The typical school week is 5 days. However, there were several 4-day weeks during the intervention phase. During baseline, week 2 was a 4-day week due to a national holiday. During the intervention phase, weeks 5, 7, 8, 9, 10, and 12 were 4-day weeks due to holidays, in-service, and school closings. In order to compare 5-day weeks with 4-day weeks, the data was prorated by adding 20% to the 4-day weeks. All weekly data is reported with prorated figures. The quizzes passed and quizzes taken data is reported in daily averages instead of weekly to increase the comparability between 4-day and 5-day weeks.

Student performance data were analyzed for effect size by dividing the difference between the intervention and baseline means by the standard deviation during baseline (Busk & Serlin, 1992). ANOVA was used to analyze data across ability groups and intervention phases.

Data on student and teacher acceptability of the interdependent group contingency were collected on the last day of the intervention. Teacher's completed an acceptability form with 16 questions and a 6-point Likert scale for responses (Table 1). Student's completed an acceptability form with 10 questions and a 4-point Likert scale for responses (Table 2). Both Likert scale responses ranged from 1 for strongly disagree to 4 or 6 for strongly agree.

Table 1
Teacher Acceptability Rating Scale

Academic Intervention Rating Scale The procedures used for the A/R 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.	1	2	3	4	5	6
2. Most teachers would find this intervention appropriate for academic problems in addition to the A/R reading program.	1	2	3	4	5	6
3. The intervention should prove effective in changing the class's academic behavior.	1	2	3	4	5	6
4. I would suggest the use of this intervention to other teachers.	1	2	3	4	5	6
5. The class's A/R reading is severe enough to warrant use of this intervention.	1	2	3	4	5	6
6. I would be willing to use this intervention in the classroom setting.	1	2	3	4	5	6
7. The intervention would *not* result in negative side-effects for the class.	1	2	3	4	5	6
8. The intervention would be appropriate for a variety of children.	1	2	3	4	5	6
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 A/r 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

Table 2
Student Acceptability Rating Scale

A/R Reading Reward System Rating Scale					
		Strongly Disagree	Disagree	Agree	Strongly Agree
1.	The reward system is good for the A/R reading program.	1	2	3	4
2.	I like the A/R reward system.	1	2	3	4
3.	I would read more for the A/R reward system.	1	2	3	4
4.	Most kids would read more for the A/R 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 A/r reading reward system.	1	2	3	4
7.	The A/R 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

The teachers and students completed treatment acceptability scales after data collection was complete. Scales were completed the last day of data collection just prior to the weekly drawing for their group contingent reward. The acceptability scales were completed prior to the reward drawing to reduce the effect that the drawing results (earned a reward or did not earn a reward) would influence the acceptability rating. The students and teachers were instructed to read the document and respond to the items. 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 answered any questions while the students and teachers completed the forms.

Treatment Integrity

The three classroom teachers at 20% of intervention drawings were presented checklists and asked to monitor the primary experimenter's implementation. The steps and procedures monitored to ensure consistency were (a) printing of the weekly report, (b) exclusion of student names and student data without permission, (c) set up of criteria and reward containers, (d) obtaining classroom attention, (e) telling the class about any added or removed reward items, (f) shaking the criteria container to mix up the numbers, (g) drawing of one slip of paper from the criteria container, (h) recording of the criterion drawn, (i) comparison of criterion to class performance for that week, (j) determination of earning reward, (k) drawing of reward, (l) recording of reward earned, (m) determination of reward implementation and recording it, (n) reward fulfillment, and (o) reminder of continued intervention for next week. Experimental integrity was 98% for all classroom administrations.

CHAPTER 3: RESULTS

Time series graphs containing the data for each dependent variable were visually analyzed. Effect size data were calculated for each class and each dependent variable by dividing phase mean difference by the standard deviation from the baseline phase (Busk & Serlin, 1992). Finally, the entire group was divided into three sub-groups (e.g., high, middle, and low performing groups) and ANOVA was used to evaluate the effects of the intervention on the three different groups across each dependent variable.

Visual Analysis of Time-Series Graphs

A/R Quizzes Passed

Figure 1 shows an immediate increase in each class's quizzes passed after the intervention was implemented. Class A had a stable baseline, but during the intervention phase Class A showed high variability in 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 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 are cyclical, with an immediate increase followed by a rapid decline and another increasing trend.

A/R Quizzes Taken

Figure 2 shows an immediate increase in each class's rate of quizzes taken after the intervention was implemented. Class A had a stable baseline, but during the intervention phase Class A showed high variability including an increasing trend

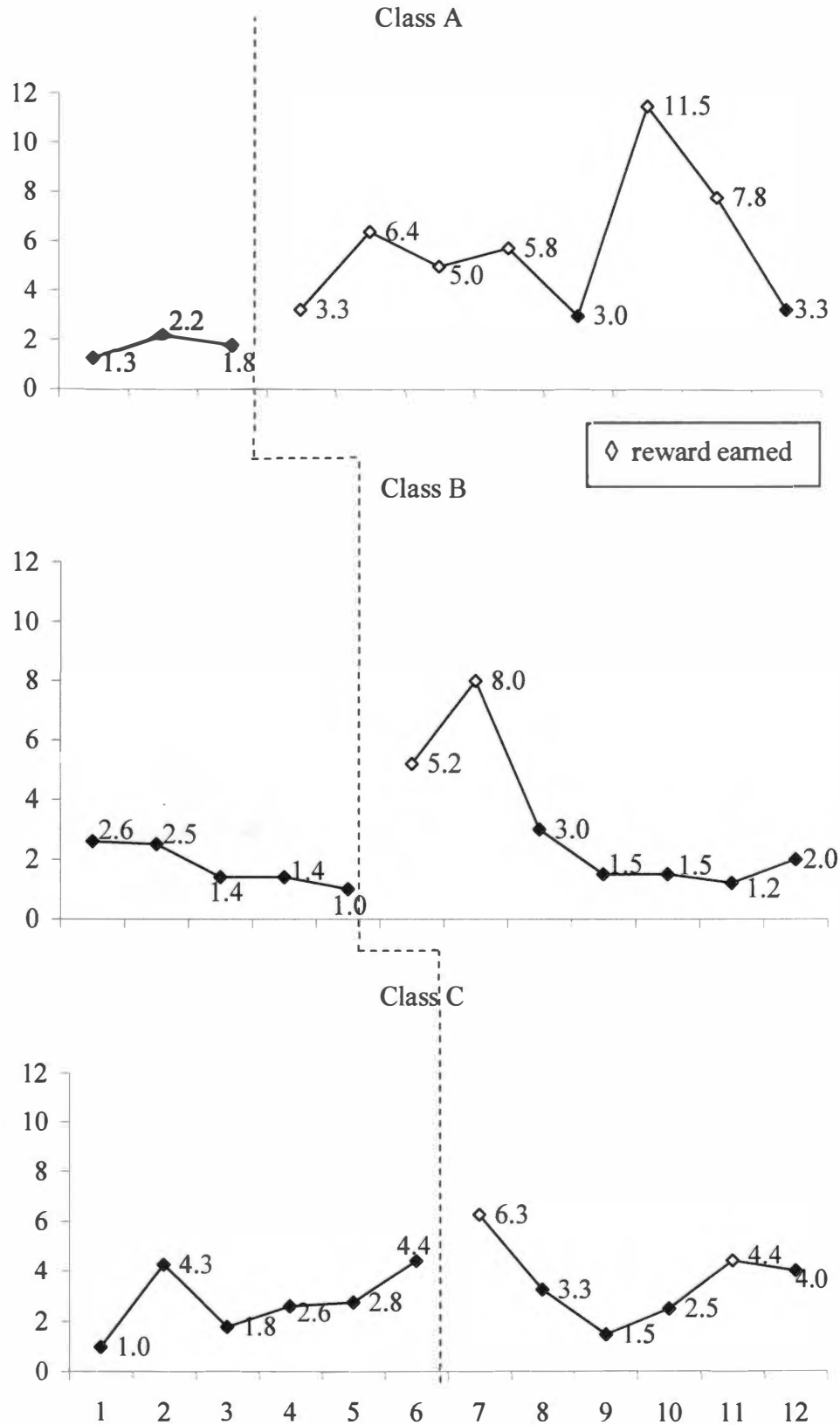


Figure 1. Average daily quizzes passed by class.

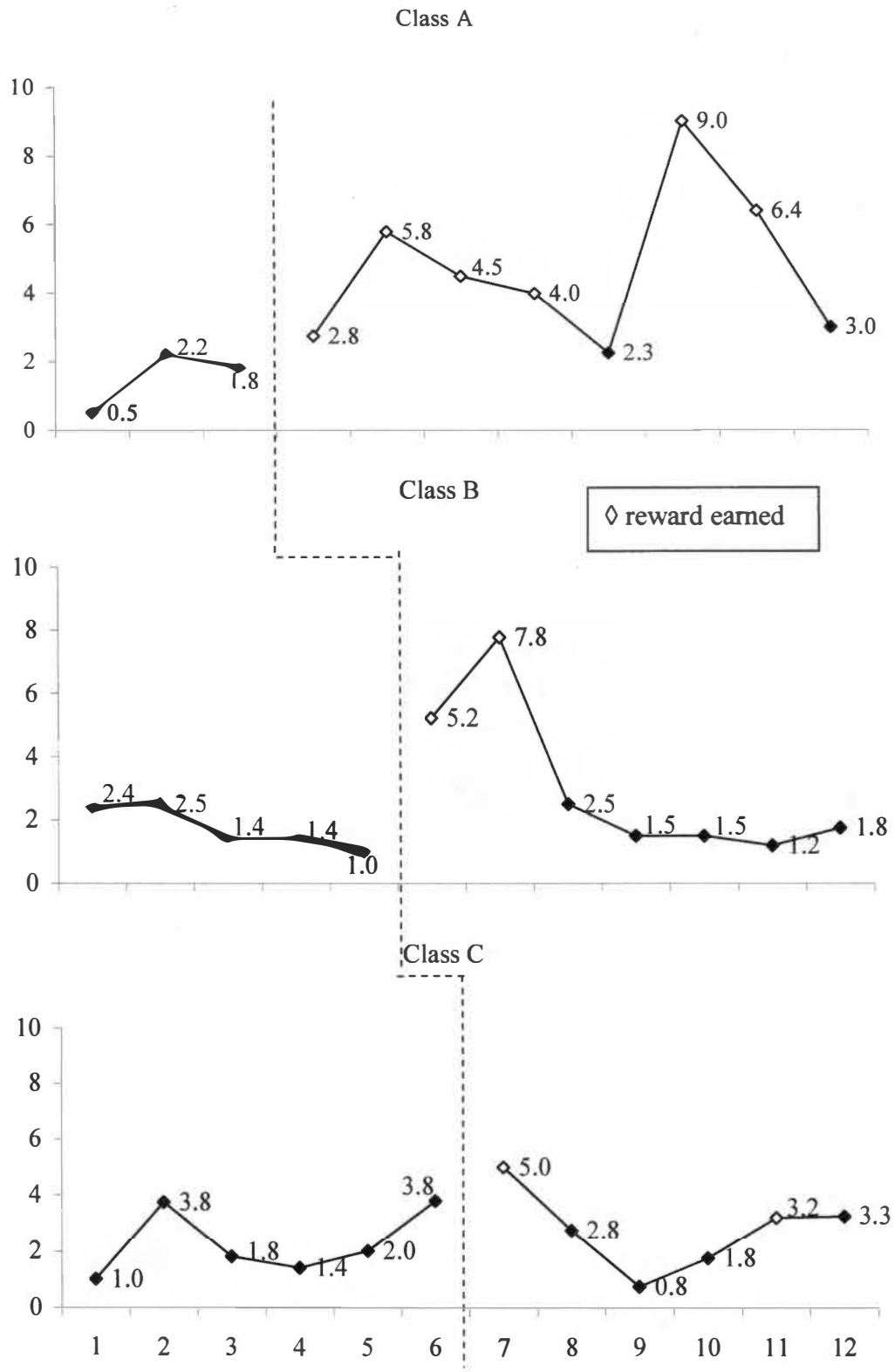


Figure 2. Average daily quizzes taken by class.

followed by a decreasing trend. The intervention may have had an immediate positive effect in Class A that was not maintained. Class B had a decreasing baseline with low variability. The intervention appeared to have an immediate positive effect on Class B that was not maintained. Class C showed an increasing baseline with high variability that hinders the ability to draw cause-and-effect conclusions. Again, the intervention phase data is cyclical with an immediate increase followed by a rapid decline and another increasing trend.

Book Level

Figure 3 shows no consistent, clear, immediate changes in each class's book level after the intervention was implemented. The data for Class C are more stable than for Class A and B. Across all three classes, Figure 3 suggests no change in book levels after the intervention was implemented.

Percentage Correct

Figure 4 shows no clear or consistent changes in each class's percent correct after the intervention was implemented. Figure 4 shows that Class A had a stable baseline with a declining slope. After the implementation of the intervention, Class A showed an immediate increase, but it was not maintained and performance was variable during the intervention phase. Class B had a stable baseline with low variability. The intervention had no immediate effect on Class B with continued low variability. Class C showed a decreasing baseline with variability. The intervention had little to no effect on Class C with an increase in variability during the intervention phase.

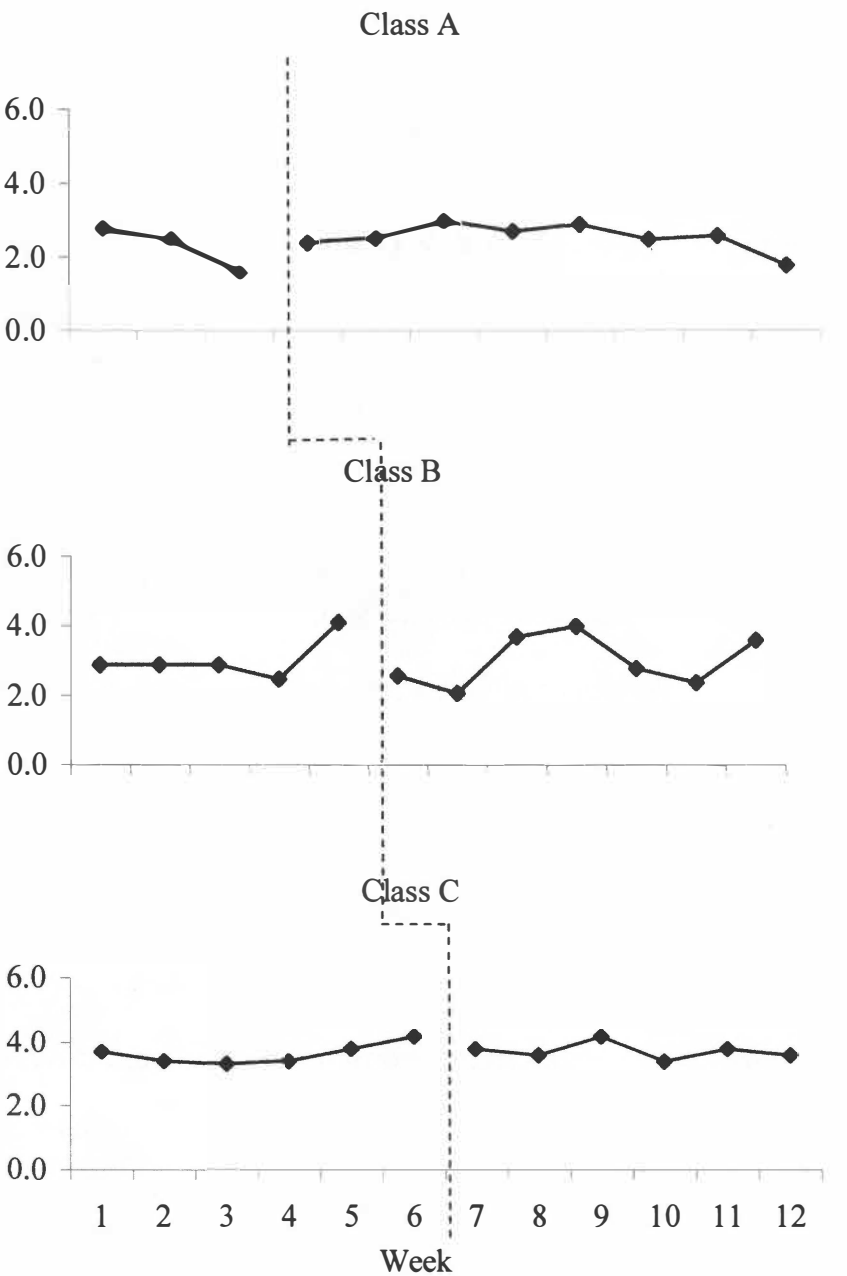


Figure 3. Weekly average book levels by class.

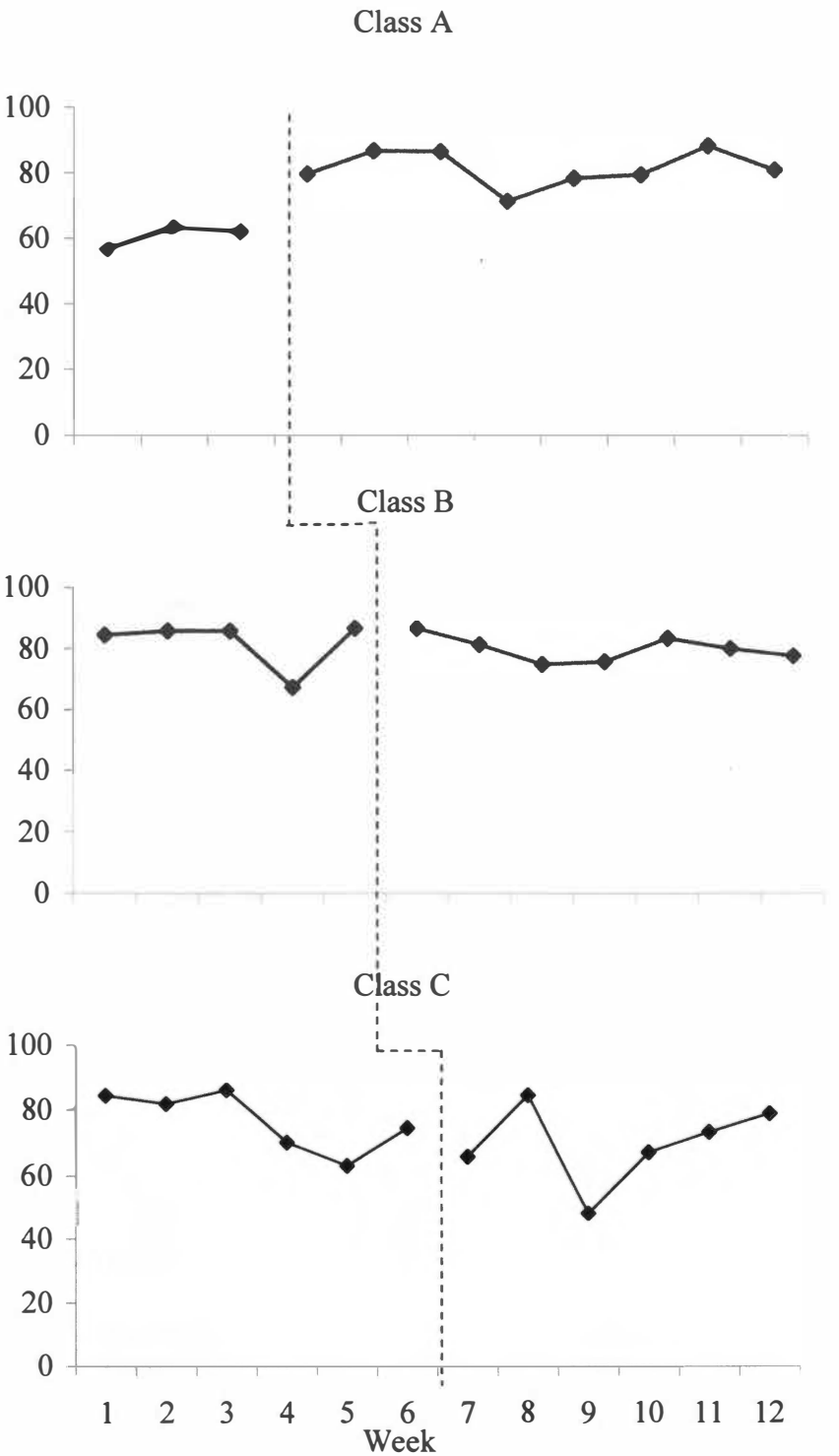


Figure 4. Weekly average percent correct by class.

Effect Size Analysis

A/R Quizzes Passed

Table 3 provides summary data for the daily average number of quizzes passed phase for Classes A, B, and C and the entire group. During baseline, the mean number of daily average quizzes passed per week for the total sample was 0.95. During the intervention phase, the mean number of daily average quizzes passed per week for the total sample was 1.87. Effect size calculated by dividing the difference between the intervention and baseline means by the standard deviation during baseline (Busk & Serlin, 1992) showed a large effect of 1.14 for quizzes passed (Cohen, 1962).

During baseline, the daily average number of quizzes passed for Classes A, B, and C respectively were 1.50, 1.74, and 2.29. During the intervention phase, the daily average number of quizzes passed for Classes A, B, and C increased to 4.71, 3.06, and 2.78 respectively. The effect size calculations revealed a large effect size (see Cohen, 1962) for Classes A and B (*ES* = 4.42, 1.63 respectively) and a small effect size (*ES* = 0.39) for Class C.

Table 3
Quizzes Passed, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
Class A	9	1.50	0.73	4.71	4.42
Class B	10	1.74	0.81	3.06	1.63
Class C	13	2.29	1.26	2.78	0.39
Total Sample	32	0.95	0.82	1.87	1.14

A/R Quizzes Taken

Table 4 provides the summary data for the daily average number of quizzes taken for Classes A, B, and C and the total group. For the total group, during baseline, the mean number of daily average quizzes taken per week was 1.07. During the intervention phase, the daily average number of quizzes taken increased to 2.17. The effect size was large (ES = 1.23).

During baseline, the daily average number of quizzes taken for Classes A, B, and C respectively were 1.75, 1.78, and 2.80. During the intervention phase, this increased to 5.74, 3.20, and 3.65 for Classes A, B, and C respectively. The effect size calculations revealed large effect sizes for Classes A and B (*ES* = 8.88 and 1.68 respectively) and a moderate effect size for Class C (*ES* = 0.66).

Book Level

Table 5 displays the weekly average reading book levels for Classes A, B, and C during baseline and intervention. During baseline, the average book level for the total sample was 3.29. During the intervention phase, the average book level for the total sample was 3.21. Effect size calculations showed no effect (ES = -.08; Cohen, 1962),

Table 4
Quizzes Taken, Means, Standard Deviations, and Effect Sizes (*N* = 32)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
Class A	9	1.75	0.45	5.74	8.88
Class B	10	1.78	0.84	3.20	1.68
Class C	13	2.80	1.29	3.65	0.66
Total Sample	32	1.07	0.89	2.17	1.23

Table 5
Average Book Level, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
Class A	9	2.30	1.49	2.55	0.17
Class B	10	3.06	0.61	3.03	-0.05
Class C	13	3.62	0.34	3.73	0.34
Total Sample	32	3.29	1.02	3.21	-0.08

which suggests that the intervention did not cause students to read more difficult or easier material.

Weekly average book levels 2.30, 3.06, and 3.62 during baseline phases and 2.55, 3.03, and 3.73 during the intervention phases for Classes A, B, and C respectively. Effect sizes showed little change across the Classes A, B, and C (*ES* = 0.17, *ES* = -0.05, and *ES* = .34 respectively). These analyses suggest that the intervention had little or no effect on book level for Classes A and B, but may have resulted in a small increase in book level for Class C (Cohen, 1962).

Percentage Correct

During baseline, the average percent correct for the total sample was 67.86 (Table 6). During the intervention phase, the average percent correct for the total sample was 78.19. Effect size showed no effect for percent correct (*ES* = .27; Cohen, 1962).

During baseline, the weekly average percent correct for Classes A, B, and C was 60.87, 81.96, and 76.70 respectively. During the intervention phase, the average percent correct for Classes A, B, and C was 81.61, 79.94, and 69.67 respectively. *T*-tests

Table 6
Average Percent Correct, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	N	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
Class A	9	60.87	3.68	81.61	5.64
Class B	10	81.96	8.24	79.94	-0.25
Class C	13	76.70	9.03	69.67	-0.78
Total Sample	32	67.86	38.27	78.19	0.27

revealed a statistically significant increase between the baseline and intervention phases for Class A ($p = .009$) but no significant change for Classes B ($p = .13$) and C ($p = .30$). The effect size for Class A revealed a large increase in percent correct ($ES = 5.64$), a moderate decrease in percent correct for Class C ($ES = -0.78$) and no effect on Class B ($ES = -0.25$).

Summary of Visual and Effect Size Analyses

Effect size analysis of class average data suggests that the intervention may have been effective in increasing the number of AR quizzes taken and passed, particularly for Classes A and B. However, visual analyses of time-series graphs show highly variable and possibly cyclical data that prevents drawing strong cause and effect conclusions. Both visual analysis of time series graphs and effect size analysis suggest that the intervention had little effect on the level book students read and the percent correct on their comprehension exams.

Statistical Analysis by Groups

The effect of the group contingency on different students was also analyzed. For each dependent variable the total group (all three classes) were divided into three groups (high, middle, and low) based on their average baseline performance relative to their peers. For each dependent variable students were placed in the low group if his/her average baseline performance was half a standard deviation below the mean. Each child was placed in the high group if his/her average baseline performance was half a standard deviation above the mean. The middle group fell within with +/- a half of standard deviation of the mean. ANOVA's and post-hoc analysis were then used to determine if the group contingency had different effects across the three groups.

A/R Quizzes Passed

The total sample was divided into three groups (low, middle, and high) according to baseline average daily quizzes passed. Students were placed in the low group if their average daily quizzes passed was a half of a standard deviation (SD=0.81) below the mean (x=0.70). All students averaging 0.31 quizzes passed daily or less during baseline were placed in the low group (n=12). All students averaging 1.14 quizzes passed daily or more, half of a standard deviation above the mean, were placed in the high group (n=6) and all those averaging between 0.32 and 1.13 quizzes passed daily were in the middle group (n=14).

A mixed model, between and within subject ANOVA was calculated to evaluate the differences between groups and between baseline and intervention means for each group. There was a main effect for the groups, $F(1, 5) = 5.48, p < .05$, which was expected as students were placed into groups based on relative performance. Post hoc

Table 7
Tukey's Post Hoc Tests, Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
high	low	1.1323*	.36323	.011	.2352	2.0294
	middle	.9483*	.35448	.032	.0728	1.8237
middle	high	-.9483*	.35448	.032	-1.8237	-.0728
	low	.1840	.28579	.797	-.5218	.8898
low	high	-1.1323*	.36323	.011	-2.0294	-.2352
	middle	-.1840	.28579	.797	-.8898	.5218

*the mean difference is significant at the .05 level.

comparison using Tukey's test showed that the high group was significantly different from both the middle and the low group ($p < .05$) while the low and middle groups were not significantly different from each other ($p > .05$; Table 7).

The ANOVA revealed a main effect for the intervention which showed that daily quizzes passed were significantly higher, $F(1, 5) = 5.38, p < .05$, with the intervention phase than on the baseline phase (Table 8). This main effect must be interpreted in light of the significant interactions found, $F(2, 29) = 5.48, p < .05$.

Table 9 displays the mean daily quizzes passed for the low, middle, and high groups during baseline and intervention phases. These data show an increase in quizzes passed for both the low and middle groups, but a decrease in quizzes passed for the high

Table 8
Main Effect and Interaction Effects

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
	Main Effect of Reinforcement				
Reinforcement	3.713	1	3.713	5.375	0.028
Error	20.035	29	0.691		
	Main Effect of Group				
Group	7.576	2	3.788	5.483	0.010
Error	20.035	29	0.691		
	Interaction Effects				
Intercept	82.171	1	82.171	77.850	0.000
Error	30.610	29	1.056		

Table 9
Average Quizzes Passed, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
High	6	2.12	0.72	1.69	-0.59
Middle	15	0.67	0.25	1.27	2.40
Low	11	0.002	0.005	1.46	291.60

group. A large effect size from baseline to intervention was calculated for the low and middle groups (low $ES = 291.6$, middle $ES = 2.40$). The effect size for the high group decreased moderately ($ES = -0.59$).

Tukey's post hoc, pairwise comparison revealed that the phase mean differences for the low performing group was statistically significant ($p < .05$, Table 10, Figure 5). The phase mean differences for the middle and high performing groups was not statistically significant. These analyses suggests that the reinforcement program may have increased the number of quizzes passed for the low performing group, while having no effect on the middle and high performing groups.

A/R Quizzes Taken

The total sample was divided into three groups (low, middle, and high) according to baseline average daily quizzes taken. Students were placed in the low group if their average daily quizzes taken was a half of a standard deviation (SD=0.89) below the mean (x=1.2). All students averaging 0.46 quizzes taken daily or less during baseline were placed in the low group (n=13). All students averaging 1.34 quizzes taken daily or more, half of a standard deviation above the mean, were placed in the high group (n=7) and all those averaging between 0.47 and 1.33 quizzes taken daily were in the middle group (n=12).

A mixed model, between and within subject ANOVA was calculated to evaluate the differences between groups and between baseline and intervention means for each group. There was a main effect for the groups, $F(1, 5) = 4.098$, $p < .05$, which was anticipated as students were placed into groups based on their relative performance. Post

Table 10
Tukey's Post Hoc Tests, Pairwise Comparisons

Group	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
High	.433	.487	.381	-.564	1.430
Middle	-.595	.308	.063	-1.226	.0035
Low	-1.443*	.360	.000	-2.179	-.706

*the mean difference is significant at the .05 level.

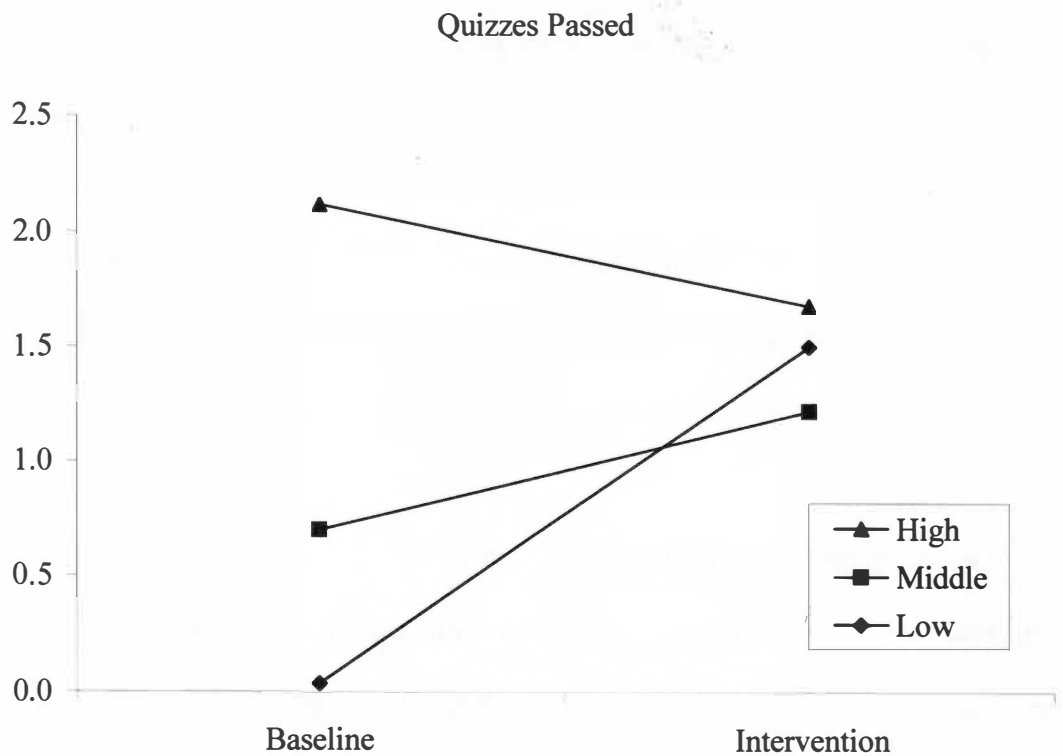


Figure 5. Quizzes passed by low, middle, and high groups.

hoc comparison using Tukey's test showed that the high group was significantly different from both the middle and low group (p <.05) while the low and middle group were not significantly different from each other (p>.05, Table 11).

The ANOVA revealed a main effect for the intervention which showed that daily quizzes taken were significantly higher, $F(1, 5) = 7.69, p < .05$, during the intervention phase relative to the baseline phase (Table 12). This main effect must be interpreted in light of the significant interactions found, $F(2, 29) = 4.1, p < .05$.

Table 13 and Figure 6 display the mean daily quizzes taken for the low, middle, and high groups during baseline and intervention phases. These data show an increase in quizzes taken for both the low and middle groups, but a decrease in quizzes taken for the

Table 11
Tukey's Post Hoc Tests, Multiple Comparisons

					95% Confidence Interval	
(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
high	low	1.29*	.372	.005	.37	2.21
	middle	0.99*	.378	.035	.06	1.93
middle	high	-0.99*	.378	.035	-1.93	-.06
	low	0.30	.318	.615	-.48	1.09
low	high	-1.29	.372	.005	-2.21	-.37
	middle	-0.30	.318	.615	-1.09	.48

*the mean difference is significant at the .05 level.

Table 12
Main Effect and Interaction Effects

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
		Main Effect of Reinforcement			
Reinforcement	6.676	1	6.676	7.687	.010
Error	25.185	29	.868		
		Main Effect of Group			
Group	7.118	2	3.559	4.098	.027
Error	25.185	29	.868		
		Interaction Effects			
Intercept	117.471	1	117.471	93.055	.000
Error	36.609	29	1.262		

Table 13
Average Quizzes Taken, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
High	7	2.224	0.696	2.114	-0.158
Middle	12	0.900	0.268	1.453	2.063
Low	13	0.009	0.145	1.660	11.386

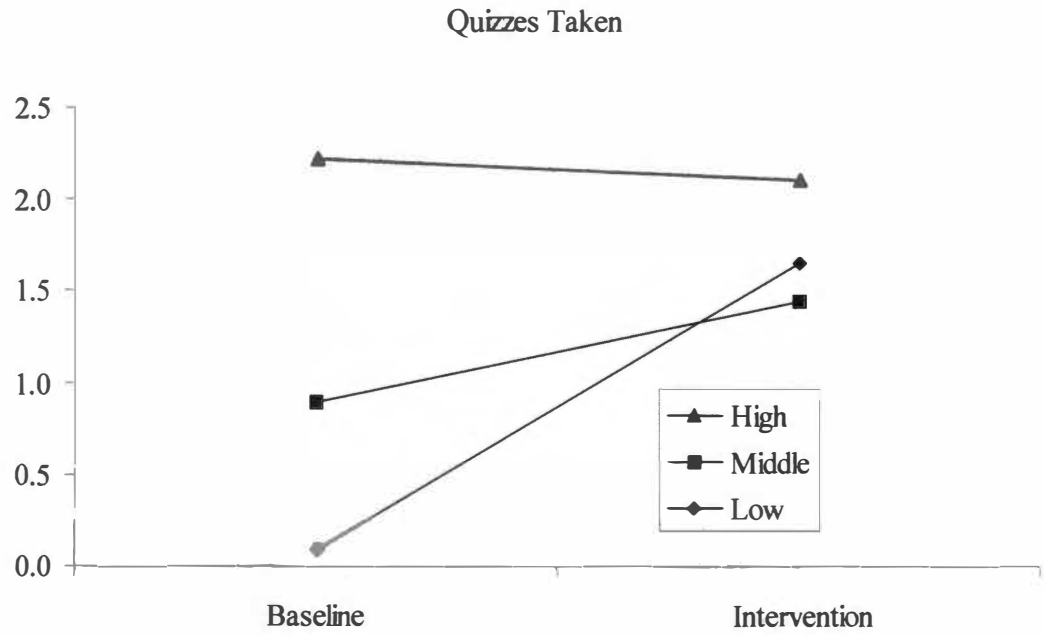


Figure 6. Quizzes taken by low, middle, and high groups.

high group. A large effect size from baseline to intervention was calculated for the low and middle group (low *ES* = 11.386, middle *ES* = 2.063). There was no effect for the high group (*ES* = -.158).

Tukey's post hoc, pairwise comparisons revealed that the phase mean differences for the low performing group was statistically significant ($p < .05$, Table 14). The phase mean differences for the middle and high performing groups was not statistically significant. These analyses suggest that the reinforcement program may have increased the number of quizzes taken for the low performing group, but had little impact on the number of quizzes taken for the middle and high performing groups.

Table 14
Tukey's Post Hoc Tests, Pairwise Comparisons

Group	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
				Lower Bound	Upper Bound
High	.110	.499	.827	-.910	1.130
Middle	-.553	.381	.157	-1.331	.226
Low	-1.571*	.366	.000	-2.319	-.823

*the mean difference is significant at the .05 level.

Book Level

The total sample was divided into three groups (low, middle, and high) according to baseline weekly average book level. A student was placed in the low group if his/her weekly average book level was a half of a standard deviation (SD = 1.02) below the mean (x = 3.29). All students averaging a 2.27 book level per week or less during baseline were placed in the low group (n=15). All students averaging a 4.33 book level per week or more, half of a standard deviation above the mean, were placed in the high group (n=4) and all those averaging between a 2.28 and a 4.32 book level per week were in the middle group (n=13).

A mixed model, between and within subjects ANOVA was calculated to evaluate the differences between groups and between baseline and intervention means for each group. There was a main effect for groups, $F(1, 5) = 6.821$, $p < .05$, which was expected as students were placed into groups based on their relative performance. Post hoc comparison using Tukey's test showed that all groups were significantly different from each other ($p < .05$, Table 15).

Table 15
Tukey's Post Hoc Tests, Multiple Comparisons

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
high	low	3.071*	.3847	.000	2.120	4.021
	middle	1.306*	.3909	.006	.341	2.272
middle	high	-1.306*	.3909	.006	-2.272	-.341
	low	1.764*	.2591	.000	1.125	2.404
low	high	-3.071*	.3847	.000	-4.021	-2.120
	middle	-1.764*	.2591	.000	-2.404	-1.125

*the mean difference is significant at the .05 level.

The ANOVA revealed no main effect for the intervention for weekly average book level, $F(1, 5) = 1.376$, $p > .05$ (Table 16). This main effect must be interpreted in light of the significant interactions found, $F(2, 29) = 6.821$, $p < .05$.

Table 17 displays the mean weekly book level for the low, middle, and high groups during baseline and intervention phases. These data show an increase in book level for the low group and a decrease in book level for the middle and high groups. A large effect size from baseline to intervention was calculated for the low group (ES = 1.722). The effect size for the middle group was small (ES = -0.432) and the effect size for the high group was moderate (ES = -0.689).

Tukey's post hoc, pairwise comparison revealed that the phase mean differences for the low performing group was statistically significant ($p < .05$, Table 18, Figure 7).

Table 16
Main Effect and Interaction Effects

Source	Type III Sum of Squares	*df*	Mean Square	*F*	Sig.
	Main Effect of Reinforcement				
Reinforcement	1.483	1	1.483	1.376	.250
Error	31.260	29	1.078		
	Main Effect of Group				
Group	14.705	2	7.352	6.821	.004
Error	31.260	29	1.078		
	Interaction Effects				
Intercept	438.188	1	438.188	468.774	.000
Error	27.108	29	.935		

Table 17
Average Book Level, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	*N*	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
High	4	4.725	.472	4.400	-0.689
Middle	13	3.369	.551	3.131	-0.432
Low	15	0.653	.964	2.313	1.722

Table 18
Tukey's Post Hoc Tests, Pairwise Comparisons

				95% Confidence Interval	
Group	Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
High	0.351	.734	.636	-1.150	1.853
Middle	0.230	.407	.577	-0.603	1.062
Low	-1.661*	.379	.000	-2.437	0.886

*the mean difference is significant at the .05 level.

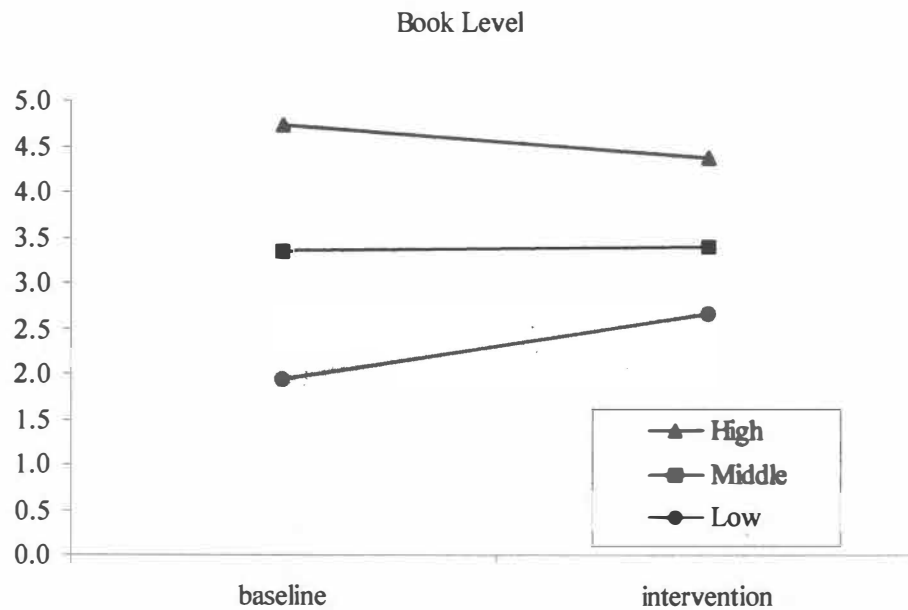


Figure 7. Book level by low, middle, and high groups.

The phase mean differences for the middle and high performing groups was not statistically significant. These analyses suggest that the reinforcement program may have increased the book level for the low performing group, but caused no significant changes in book level reading for the middle or high groups.

Percentage Correct

The total sample was divided into three groups (low, middle, and high) according to baseline weekly average percent correct. A student was placed in the low group if his/her weekly average percent correct was a half of a standard deviation (SD= 20.2) below the mean (x=75.0). All students averaging 64.9 percent correct per week or less during baseline were placed in the low group (n=12). All students averaging 85.2 percent correct per week or more, half of a standard deviation above the mean, were placed in the

high group (n=5) and all those averaging between 65.0 and 85.1 percent correct per week were in the middle group (n=15).

To evaluate the differences between groups and between baseline and intervention means for each group, a mixed model, between and within subject ANOVA was calculated. There was a main effect for the groups, $F(1, 5) = 29.54, p < .05$, which was anticipated as students were placed into groups based on their relative performance. Post hoc comparison using Tukey's test showed that the low group was significantly different from both the middle and high groups ($p < .05$) while the middle and high groups were not significantly different from each other ($p > .05$; Table 19).

The ANOVA revealed a main effect for the intervention which showed that weekly average percent correct was significantly higher, $F(1, 5) = 9.34, p < .05$, during

Table 19
Tukey's Post Hoc Tests, Multiple Comparisons

					95% Confidence Interval	
(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
high	low	53.855*	7.3185	.000	35.781	71.929
	middle	16.841	7.1000	.062	-.693	34.376
middle	high	-16.841	7.1000	.062	-34.376	.693
	low	37.014*	5.3250	.000	23.863	50.165
low	high	-53.855*	7.3185	.000	-71.929	-35.781
	middle	-37.014*	5.3250	.000	-50.165	-23.863

*the mean difference is significant at the .05 level.

the intervention phase relative to the baseline phase (Table 20). This main effect must be interpreted in light of the significant interactions found, $F(2, 29) = 29.54$, $p < .05$.

Table 21 displays the mean weekly percent correct for the low, middle, and high groups during baseline and intervention phases. These data show an increase in percent correct for the low group, but a decrease in percent correct for the high and middle groups. A large effect size from baseline to intervention was calculated for the low group (*ES* = 3.75). The effect size for the high and middle group was largely negative (high *ES* = -2.3, middle *ES* = -0.84).

Tukey's post hoc, pairwise comparison revealed that the phase mean differences for the low performing group was statistically significant ($p < .05$, Table 22, Figure 8).

Table 20
Main Effect and Interaction Effects

Source	Type III Sum of Squares	*Df*	Mean Square	*F*	Sig.
	Main Effect of Reinforcement				
Reinforcement	2480.885	1	2480.885	9.341	.005
Error	7702.093	29	265.589		
	Main Effect of Group				
Group	15690.924	2	7845.462	29.540	.000
Error	7702.093	29	265.589		
	Interaction Effects				
Intercept	230098.443	1	230098.443	608.605	.000
Error	10964.189	29	378.075		

Table 21

Average Percent Correct, Means, Standard Deviations, and Effect Sizes ($N = 32$)

Group	N	Baseline Mean	Baseline standard deviation	Intervention mean	Effect Size
High	5	94.929	94.93	85.980	-2.297
Middle	15	76.987	8.065	70.240	-0.837
Low	12	7.915	15.288	65.284	3.753

Table 22

Tukey's Post Hoc Tests, Pairwise Comparisons

				95% Confidence Interval	
Group	Mean Difference	Std. Error	Sig.	Lower Bound	Upper Bound
High	8.949	10.307	.392	-12.131	30.029
Middle	6.747	5.951	.266	-5.423	18.918
Low	-57.369*	6.653	.000	-70.977	-43.762

*the mean difference is significant at the .05 level.

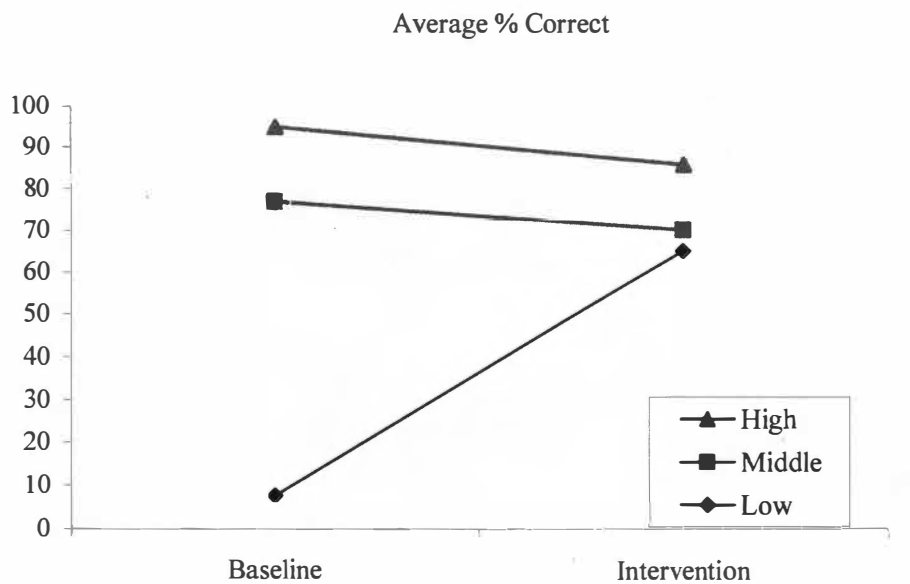


Figure 8. Percent correct by low, middle, and high groups.

The phase mean differences for the middle and high performing groups were not statistically significant. These analyses suggest that the reinforcement program may have increased the percent correct for the low performing group, but had no significant impact on the performance of the middle and high performing groups.

Summary of Analysis by Groups

While the classwide data revealed few consistent effects, when the students were divided into group based on their baseline performance some consistent findings emerged. Specifically, across all dependent variables, a significant group by treatment interaction was found. Post hoc comparisons showed that the group reward program caused a significant increase in student performance across all dependent variables for the low performing group and effect size analysis showed that these differences were large. For the middle and high performing groups, no significant changes in performance were found across any of the dependent variables. These data suggest that the intervention may have increased the amount and difficulty level of material read and the level of comprehension for the lowest performing students. However, this conclusion must be tempered by possible threats to internal validity, in particular, regression to the mean.

Teacher and Student Acceptability

The teachers and students completed treatment acceptability scales after data collection was complete. Scales were completed the last day of data collection just prior to the weekly drawing for their group contingent reward. The acceptability scales were completed prior to the reward drawing to reduce the effect that that weeks results (earned a reward or did not earn a reward) would influence the acceptability rating. The teacher acceptability form (see Table 1 in Methods) consisted of 16 questions with Likert scale

responses ranging from 1 (strongly disagree) to 6 (strongly agree). A 6 indicates a highly acceptable rating and a 1 indicates a very unacceptable rating.

The student acceptability form (see Table 2 in Methods) consisted of 10 questions with Likert scale responses. The responses ranged from 1 (strongly disagree) to 4 (stronger agree). A 4 indicates a highly acceptable rating and a 1 indicated a very unacceptable rating. 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 answered questions regarding any items on the acceptability scale.

Teacher Acceptability

The teacher's responses to the acceptability form are in Table 23. The teacher's average score across all of the items and all the 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 A/R reading program." The respondent underlined the words, "most teachers," and marked slightly disagree. These responses suggest a strong level of teacher acceptability.

Student Acceptability

The student responses to the acceptability form averaged by class are presented in Table 24. The student's average score across all items and all classes was 3.43. Across all classes, the statement receiving the highest acceptability rating was, "I'd like to continue the A/R 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" ($x = 3.01$). The lower score on that statement may be a direct reflection of the

Table 23
Teacher Responses on the Acceptability Scale

		Teacher		
		A	B	C
1.	This would be an acceptable intervention for a class with an academic problem.	5	5	4
2.	Most teachers would find this intervention appropriate for academic problems in addition to the A/R reading program.	5	3	4
3.	The intervention should prove effective in changing the class's academic behavior.	5	5	4
4.	I would suggest the use of this intervention to other teachers.	6	6	4
5.	The class's A/R reading is severe enough to warrant use of this intervention.	5	5	5
6.	I would be willing to use this intervention in the classroom setting.	6	6	5
7.	The intervention would *not* result in negative side-effects for the class.	4	6	5
8.	The intervention would be appropriate for a variety of children.	5	6	5
9.	The intervention is consistent with those I have used in the classroom setting before.	5	6	5
10.	The intervention is a fair way to handle the class's academic problems.	4	6	5
11.	The intervention is reasonable for the A/r reading program.	5	6	5
12.	I like the procedures used in the intervention.	6	5	5
13.	The intervention is a good way to handle this class's academic problems.	5	6	5
14.	Overall, the intervention would be beneficial for the class.	6	6	5
15.	The intervention would produce a lasting improvement in the class's academic behavior.	5	5	5
16.	Soon after using the intervention, the teacher would notice a positive change in the academic problem.	5	5	5
	Overall Average	5.13	5.40	4.75

Table 24
Student Responses on the Acceptability Scale

		Class Average		
		A	B	C
1.	The reward system is good for the A/R reading program.	3.62	3.58	3.71
2.	I like the A/R reward system.	3.46	3.50	3.86
3.	I would read more for the A/R reward system.	3.54	3.25	3.43
4.	Most kids would read more for the A/R reward system.	2.92	3.50	3.31
5.	The reward system is good for all the kids.	3.54	3.58	3.57
6.	I'd like to continue the A/r reading reward system.	3.77	3.50	3.93
7.	The A/R reward program would be OK for other school work.	3.00	3.17	3.08
8.	Most kids would find the reward system OK for other school work.	3.15	2.92	3.21
9.	The reward system is fair for the whole class.	3.31	3.25	3.77
10.	The reward system would only have good results.	2.75	2.92	3.36
	Overall Total Average	3.62	3.32	3.52

side effects of a group contingency. There were many weeks when no reward was earned. Those weeks may have influenced the responses on that statement. These acceptability rating form responses from the students indicates a strong level of acceptability for the group contingency intervention.

CHAPTER 4: DISCUSSION

Summary

Previous researchers have shown that group contingencies with randomized components can enhance academic performance (Popkin & Skinner, 2003; Sharp & Skinner, 2004). The current study was designed to extend this line of research by investigating the effects of a group contingency on the reading behavior of 4th grade students across three intact classes.

Across all three classrooms, visual analysis of time-series graphs provide some support for an increase in A/R quizzes taken and passed immediately after the intervention was applied. However, this increase was not maintained throughout the intervention phase. These results were similar to the findings of Shapiro and Goldberg (1986) with spelling performance of 6th grade students. Shapiro and Goldberg found an immediate increase in spelling scores that did not maintain throughout the intervention. Furthermore, only one class (Class A) in the current study showed a clear increase in percent of comprehension questions answered correctly after the group contingency was applied. Finally, across all three classes there were no clear changes in reading level of books selected after the intervention was applied. Because the class-wide data analyses provide no clear or consistent evidence of a treatment effect, the current results provide little support for concluding that the group contingency enhanced reading performance.

When students were divided into three groups (high, middle, and low performers) based on relative scores for each dependent variable, statistical analysis of mean phase scores revealed some significant interactions. The low and middle scoring groups showed a significant increase in quizzes taken and passed after the group contingency was

applied *relative* to the high group. The low group showed a significant increase in percent correct on quizzes and book level *relative* to the middle and high groups. These interactions suggest that the group-oriented contingency enhanced the performance of the low group across all dependent variables *relative* to the high performers. These results are similar to the findings of Shapiro and Goldberg (1986) who also found that the group contingency enhanced the performance of the lowest performing 6th grade spelling students relative to the high performers.

Post-hoc analysis showed significant increases in reading performance for the low group across all four dependent variables, but no significant changes in performance for the middle and high groups. Thus, although the class-wide data revealed no consistent treatment effects, when students were broken intro groups, statistical analysis suggests that the treatment may have enhanced the performance of the lowest performing students across all dependent variables. The statistically significant increases in low performing students' performance were supported by effect size analyses which showed large increases in quizzes taken, quizzes passed, percent correct, and book level read for the low performers. Although some other effect size calculations showed moderate or large effects, because these differences were not statistically significant, interpretation of these effect sizes is questionable as it may have been caused by chance. While these analyses suggest that the group contingency had a consistent and large effect on the low performing students, these statistical findings must be interpreted with caution because some threats to internal validity confounded these results.

Limitations and Future Research

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 the interdependent group contingency may have caused the low performing students to increase their reading performance (i.e., percent correct, quizzes passed). Additionally, for the low performers these increases in comprehension occurred as they increased the grade-level of the material they were reading. While these statistically significant findings suggest that the intervention may have caused desirable outcomes with the low performing students, they may be compromised by regression to the mean.

Regression to the mean is a well-documented phenomenon that would suggest that the low performing baseline students' scores would increase and the high performing baseline students would decrease without an intervention. The current results suggest that this pattern emerged once the intervention was applied, but that pattern was not occurring during baseline without the intervention (see appendix Figure 9). Therefore, these results show that the group contingency enhanced the performance of the low performing group. Future researchers should address the possible presence of regression to the mean. For example, a multiple-baseline design where phase changes are made based on the performance of only the low performing students would control for this threat to internal validity.

Even if regression to the mean is controlled, the pattern of poor readers choosing more difficult material would still be difficult to interpret because reading materials may be limited. Specifically, these 4th grade students may have begun to exhaust available

reading material written for lower level readers. Although the A/R program and the school library had many reading material options for low level readers, because these 4th grade students and many other students in the school (e.g., 1st, 2nd, and 3rd 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 to read more difficult material. Along with the limited availability of reading materials, the low performing students may have been influenced by floor effects. It's possible that these students could not read at a much lower level than their current level and therefore, the only change could be an increase in book level.

A ceiling effect may be influencing the performance of the highest performing students. For the book level and percentage correct dependent variables, the highest performing students may have been limited in increasing their performance by ceiling effects. Because the highest performing students were averaging 95% correct during baseline, the most likely change would be a decrease.

Visual analysis of class-wide data suggest an immediate increase in quizzes taken and passed across all three classes. However, this increase was not maintained. Because the initial increases were promising and consistent with earlier research on group contingencies (e.g., Popkin & Skinner, 2004), future researchers should attempt to identify what caused the failure to maintain.

Visual analysis of the class-wide data shows a cyclic pattern with an increase followed by a decrease followed by an increase in target behavior. This pattern is apparent in Class A and Class C. Post-reinforcement pause may have contributed to the possible existence of a cycle. With post reinforcement pause, there is an increase in

target behavior followed by a decrease (e.g., a pause) followed by an increase again (Felton & Lyon, 1966). In the current study, a post reinforcement pause may be influencing the target behaviors. If post reinforcement pause is influencing the target behaviors, researchers may be able to reduce this influence by providing a thicker reinforcement schedule, an increased rate of reinforcement, or randomizing reinforcement intervals.

Future researchers should consider conducting studies to determine if a thicker reinforcement schedule could reduce the effects of post reinforcement pause and enhance student responding by including a systematic procedure designed to enhance encouragement, teacher praise, and student praise. Teacher and peers may be more likely to encourage and praise reading behavior if there were some sort of mechanism to signal success. For example, every time a student passed a quiz, he/she could ring a bell. Additionally, the teacher could read the names aloud of students who passed quizzes at the end of each day. Such feedback procedures may encourage teachers and peers to deliver social reinforcement at higher rates. This higher rate of social reinforcement may help students maintain behaviors that are reinforced less frequently with the group-oriented contingency (weekly reward). Another suggestion for thickening the reward schedule is to supplement the group contingency with an independent reward (e.g., a sticker for every quiz passed). The use of a cumulative reinforcement schedule (e.g., each week the number of quizzes passed continue to accumulate toward a reward whether the class earned the reward that week or not) may also enhance performance. The current study's use of a randomly selected criteria did not smooth out responding indicating that

if post reinforcement pause is effecting the target behavior, the randomly selected criteria did not reduce the post reinforcement pause impact.

The appearance of a cycle in the data may also be influenced by the length of the books chosen for A/R reading. Longer books typically take more time to complete. The initial increase in A/R quizzes taken and passed may have been the result of finishing books that were started during the baseline phase. Once those are completed, it takes time to start and finish a new book. The time required to read a new book may appear as a pause in A/R data. To control for this natural cycle, future researchers may want to change the unit of time used for the reward system. Rewarding and graphing data every two or three weeks (rather than weekly), may allow additional time for the students to finish reading at least one book. A change in unit of time would also require an increase in the length of the study. It would take twice as many weeks to collect the same number of data points.

Increasing the number of data points collected may also help investigate the cyclic pattern of data. The more data points, the more likely the data will reveal a cycle (Richard, Taylor, Ramasamy, & Richards, 1999).

In addition to the length of the study and the number of data points, the size of the sample group also influences the results of the current study. 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 versus 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). The group size for each class may have impacted the effectiveness of the interdependent group contingency on the low performing students. These students may have improved their reading performance even more if they were grouped into a smaller group. Future researchers can address this limitation by breaking the classes into smaller groups (i.e., group size of 4). Small groups should be heterogeneous in skill with careful attention to equating each group in skill level. For example, Shapiro and Goldberg (1990) split low performing and high performing students into separate groups ensuring that each group baseline average was equivalent.

The results of this study should also be interpreted in light of the treatment integrity. Several variables influenced treatment integrity across classes. Before the intervention, the teachers implemented the A/R reading program in their classroom in different ways and each teacher had differing levels of enthusiasm and encouragement for independent silent reading behaviors. Informal observation suggests that Teacher A was most anxious to start the intervention and appeared to provide the most verbal encouragement to students (e.g., encouraged them to read and provided praise when the quizzes were passed). Teacher A's overall score on the acceptability rating scale was 5.13 (on a 6-point Likert scale). More specifically, Teacher A marked "strongly agree" on the statements, "I would suggest the use of this intervention to other teachers", "I would be willing to use this intervention in the classroom setting", "I like the procedures used in the intervention", and "overall, the intervention would be good for the class." These responses indicate a high level of acceptability from Teacher A supporting the informal observation of teacher enthusiasm.

Teacher acceptability rating indicate that the teachers in this study would support the implementation of this intervention in their classroom. Most teachers are concerned about the lowest performing students in their classroom particularly in light of the No Child Left Behind initiative. The current results indicate that the interdependent group contingency utilized in this study was effective in increasing the reading behaviors of the lowest performing students. The students that did not take any A/R quizzes during baseline, increased their quiz taking rate during the intervention. The benefit to the teacher of a low achieving student increasing his/her reading behaviors out weights the cost involved in implementing the intervention. The total cost is minimal including the purchase of any rewards or supplies and time away from instruction. The rewards can be altered to fit any budget of time or money, so that the intervention is easily accommodated in the regular education classroom.

Another difference was observed across classrooms. Some teachers were strict with reading levels and access to the school-wide reward system. For example, in Class C, the teacher would not allow students to read lower level books. However, the other two teachers allowed students to change book levels. For example, one student in Class B who read at grade level during baseline chose to read lower level books at the beginning of the intervention phase so that she could take more quizzes to help the class earn the reward. However, after a couple of weeks, she said that she got bored with the easy books and went back to reading grade level books. While these practices were also in place during the baseline phases, future researchers should investigate the effects of manipulating the A/R program to determine if the degree of book level restriction impacts student performance across high, middle, and low performing students.

The interpretation of Class C's data is limited by the increasing baseline trend. The baseline data for Class C was increasing prior to the implementation of the intervention. This increasing trend impedes the ability to interpret the immediate increase in A/R quizzes taken and passed or the increase after the implementation of the intervention may have been due to another variable.

The use of prorated data to control for 4-day versus 5-day school weeks, provides further limitations on the interpretation of the current studies data. The students may or may not have taken 20% more quizzes on the missed day. The prorated data doesn't accommodate for other variables that accompany shortened school weeks. When weeks are shortened, the school year is not lengthened and therefore teachers are required to cover the same amount of curriculum in fewer days. Teachers may increase the amount of instruction and individual work to make up the curriculum lost. This may lead to less free time. The nature of this study depends on some free time for the students to do sustained silent reading. Although there is time designated during the school day for sustained silent reading, typically, students are also allowed to engage in sustained silent reading when individual seatwork is completed. However, when the curriculum is increased due to shorter school weeks, there is less free time in the day. Future researchers should control for the number of days in a school week by only including 4-day or 5-day weeks.

Summary

Class-wide data from the current study utilizing an interdependent group contingency applied to the A/R reading program in three 4th grade classrooms revealed no clear and consistent treatment effects. When the total sample was divided into three

ability groups by baseline averages, the data showed a clear and consistent treatment effect for the lowest performing students. These students increased their quizzes taken, quizzes passed, percent correct on the A/R quizzes, and reading book level. The middle and high performing students showed no statistically significant effect of the group contingency on any of the variables. These results should be interpreted in respect to the limitations and threats to internal validity.

When the sample was divided into ability groups, the results of the current study show that the low performing students increased their performance on the quizzes (increase in quizzes passed and increase in percent correct). This suggests that these students actually read the A/R books and comprehended what they read. Thus, the lowest performing group appeared to read more and read more carefully, applying more comprehension skills and effort. These increases in reading performance (e.g., quizzes taken, passed, and percent correct) for the lowest level group are even more impressive when considering the book level data. Book level data showed that the lowest performing group read significantly more difficult A/R books following the application of the group contingency. These data suggest that the group contingency may have caused the poor performing readers to read more, read more challenging material, and comprehend more of what they read. The threats to internal validity must be considered when interpreting these results.

For the middle performing group effect size analysis suggest that the group contingency results in a large increase in quizzes taken, a large increase in quizzes passed, a large decrease in percent correct, and a moderate decrease in book level. However, statistical analysis showed that none of these differences were statistically

significant at the $p < 0.05$. These analyses suggest no clear effect on reading performance for the middle group. While the group contingency may have caused the middle performers to take and pass more quizzes, because the passing level was 60%, this increase was also accompanied by a decrease in percent correct. Additionally, where as the group contingency may have caused the low performers to read more challenging books, it may have caused the moderate performers to read less challenging books.

For the high performing group, effect size analysis suggests that the group contingency had no effect on quizzes taken or book level, a moderate decrease in quizzes passed, and a large decrease in percent correct. While these data suggest that the group contingency may have decreased comprehension in the high performing group, none of these differences were statistically significant at even the $p < .380$ level.

LIST OF REFERENCES

LIST OF REFERENCES

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., 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.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In R. R. Kratochwill & J. R. Levin, Eds., *Single-case research design and analysis* (pp. 187-212). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

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.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65*, 145-153.

Crouch. P. L., Gresham, F. M., & Wright, W. R. (1985). Interdependent and independent group contingencies with immediate and delayed reinforcement for controlling classroom behavior. *Journal of School Psychology, 23*, 177-187.

Davey, B., Alexander, M., Edmonson, C., Stenhoff, D., & West, R. P. (2001). The effects of a discriminative stimulus, paired with individual and group reward contingencies, on the decibel levels in an elementary school lunch room (Report No. RIEOCT2001). San Diego, CA: Growing Partnerships for Rural Special Education. (ERIC Document Reproduction Service No. ED453023).

Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with Attention-Deficit/Hyperactivity Disorder. *Psychology in the Schools, 37*, 135-147.

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., & Gomez, A. (1994). Choice making to promote adaptive behavior for students with emotional and behavioral challenges. *Journal of Applied Behavior Analysis, 27*, 505-518.

Durkin (1993). *Teaching them to read* (6th ed.). Boston, MA: Allyn & Bacon.

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.

Felton, M., & Lyon, D. O. (1966). The post-reinforcement pause. *Journal of the Experimental Analysis of Behavior, 9*, 131-134.

Gagne, E. D., Yekovich, D., W., & Yekovich, F. R. (1993). *The cognitive psychology of school learning* (2nd ed.). New York: Addison Wesley Longman, Inc.

Green, B., Hodgens, J., & Luke, A. (1997). Debating literacy in Australia: History lessons and popular f(r)ictions. *Australian Journal of Language and Literacy, 20*, 6-24.

Gresham, F. M., & Gresham, G. N. (1982). Interdependent, dependent, and independent group contingencies for controlling disruptive behavior. *Journal of Special Education, 16*, 101-110.

Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Own, M., et al. (1971). The teacher as observer and experimenter in the modification of disrupting and talking-out behaviors. *Journal of Applied Behavior Analysis, 4*, 141-149.

Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*, 523-533.

Kern, L., Bambara, L., & Fogt, J. (2002). Class0wide 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.

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.

Madsen, C. H., Becker, W. C., Thomas, D. R., Koser, L., & Plager, E. (1968). An analysis of the reinforcing function of "sit down" commands. In R. K. Parker (Ed.), *Reading in educational psychology* (pp. 265-278). Boston: Allyn & Bacon.

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.

Patterson, G. K., Jones, R., Whittier, J., & Wright, M. A. (1965). A behavior modification technique for the hyperactive child. *Behavioral Research and Therapy, 2*, 217-226.

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.

Reading Renaissance I: Using Learning Information Systems to Create World-Class Readers (1998). Madison, WI: The Institute for Academic Excellence.

Report of the National Reading Panel (Dec. 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, D.C.: National Institute of Health.

Richards, S. B., Taylor, R. L., Ramasamy, R., & Richards, R. Y. (1999). *Single subject research: Applications in educational and clinical settings*. Belmont, CA: Wadsworth Group/Thomson Learning.

Romeo, F. F. (1998). The negative effects of using a group contingency system of classroom management. *Journal of Instructional Psychology, 25*, 130-134.

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.

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,*

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., 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.

Slavin, R. E. (1987). Cooperative learning: Where behavioral and humanistic approaches to classroom motivation meet. *The Elementary School Journal, 88*, 29-37.

Smith, M. A., & Misra, A. (1994). Behavior management: Using group contingencies with students with learning disabilities. *LD Forum, 20*, 17-20.

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.

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.

Stuckey, J. E. (1991). *The violence of literacy.* Portsmouth, NH: Boynton/Cook.

Theodore, L. A., Bray, M. A., Kehle, T. J., & Jenson, W. R. (2001). Randomization of group contingencies and reinforcers to reduce classroom disruptive behavior. *Journal of School Psychology, 39*, 267-277.

Wilkinson, L. (1999). An introduction to the explicit teaching of reading. In J. Hancock (Ed.), *The explicit teaching of reading* (pp. 1-12). Newark, DE: IRA.

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.

Winett, 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.

APPENDIX

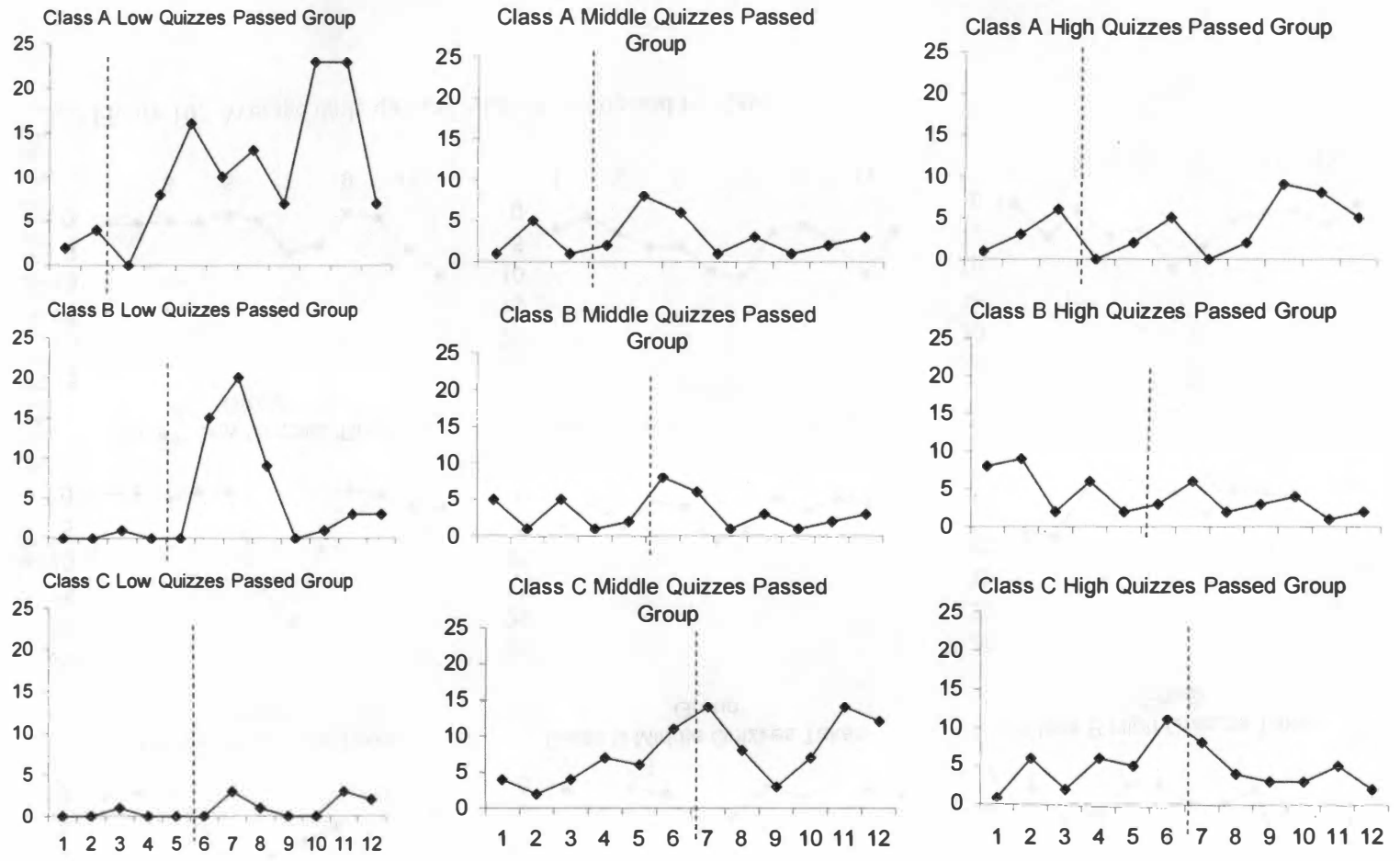


Figure 9. Average daily quizzes passed by group and by class

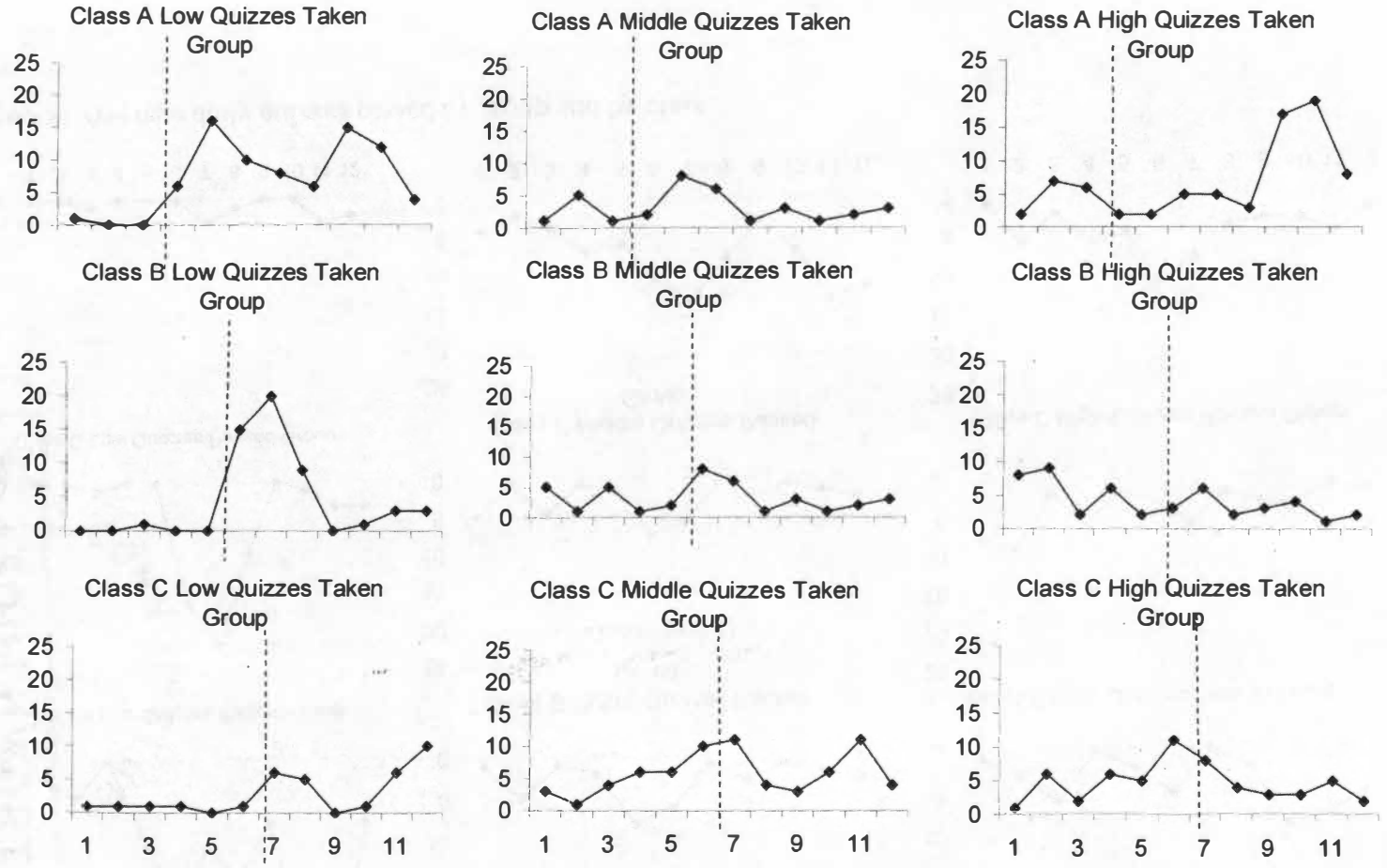


Figure 10. Average daily quizzes taken by group and by class.

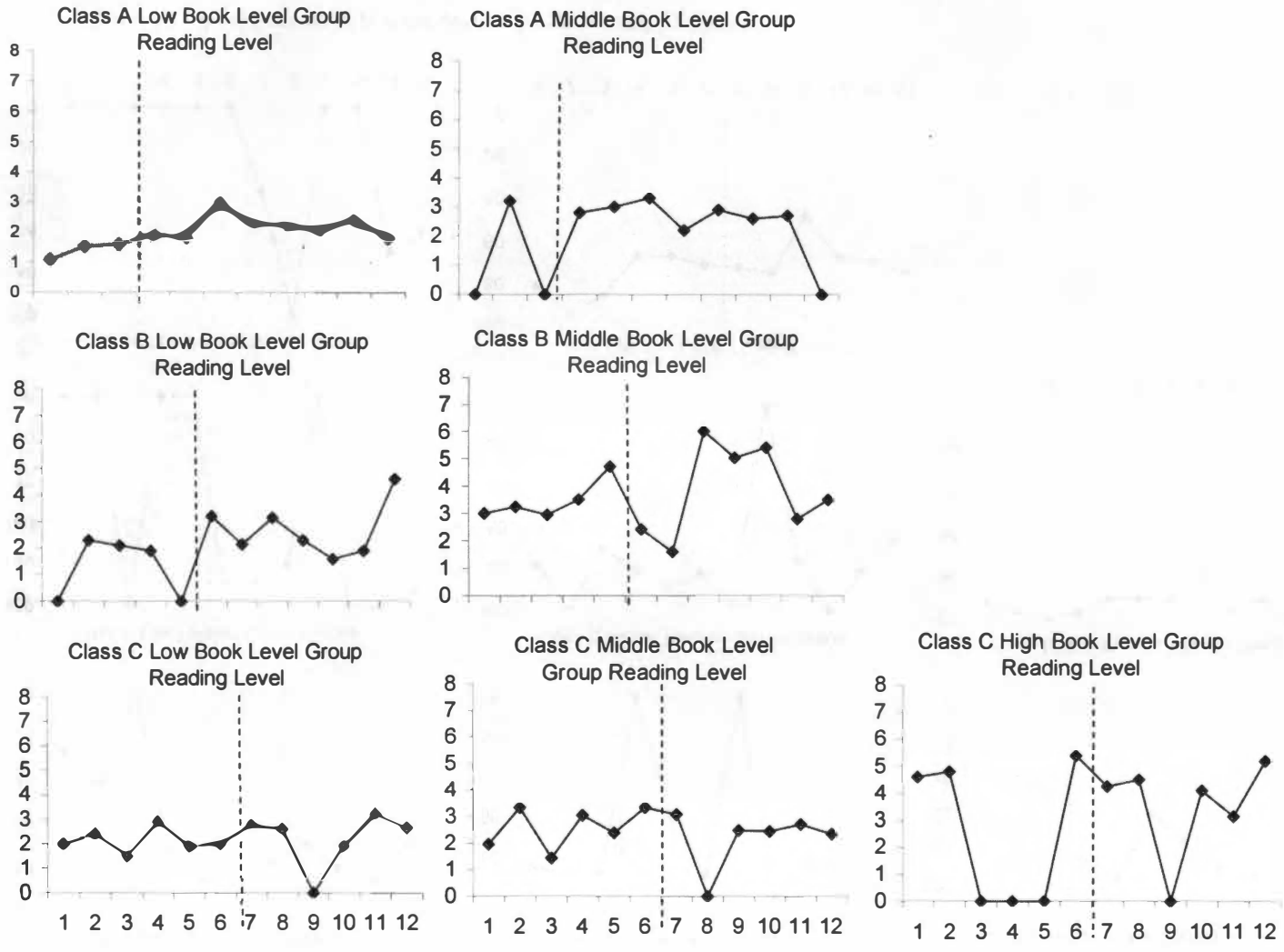


Figure 11. Weekly average book levels by group and by class

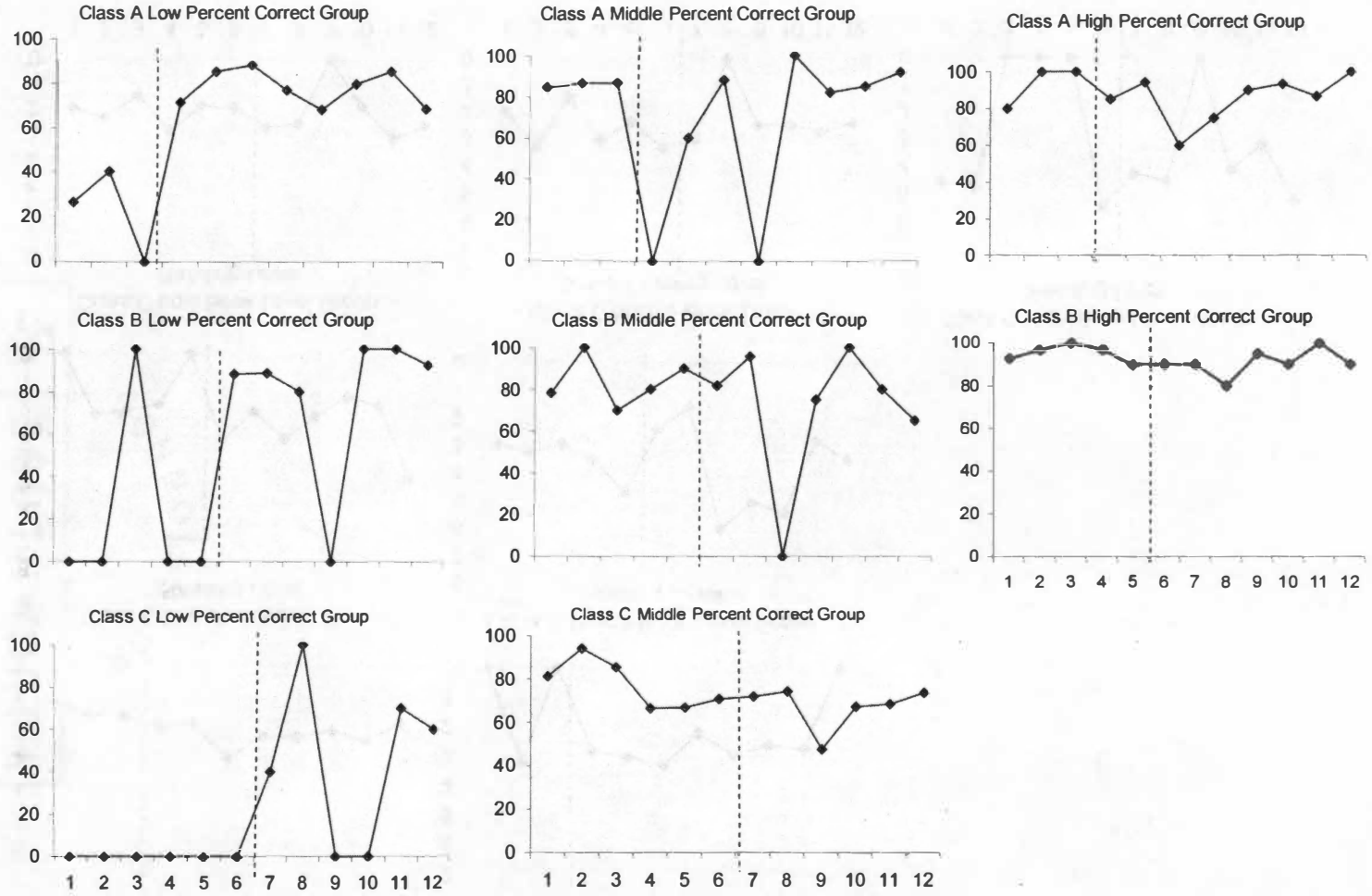


Figure 12. Weekly average percent correct by group and by class

VITA

Danielle Pappas grew up in Salem, Oregon and graduated from Sprague High School in 1990. In 1994, she received a Bachelor of Science in Behavioral Science from Northwest College in Kirkland, Washington. After years of working as an administrative assistant, Danielle enrolled in the University of Tennessee's school psychology program. Internship activities were completed with the Tennessee Internship Consortium and Monroe County Schools. In August 2006, Danielle Pappas will be receiving her Ph.D. in Education with a concentration in school psychology.

8792 6514 7
11/88/86 USA