A COMPARISON OF HIGH AND LOW FREQUENCY CRITERIA ON
READING AGILITY, RETENTION, ENDURANCE AND DIRECT
INSTRUCTION READING MASTERY CHECKOUT PERFORMANCE
OF ELEMENTARY STUDENTS ACADEMICALLY AT-RISK.

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of
Philosophy in the Graduate School of The Ohio State University

By

Barry D. Morgenstern, M.S.

The Ohio State University

2002

Dissertation Committee: Approved by

Professor John O. Cooper, Adviser

Professor Nancy Neef

Professor Timothy E. Heron

John O. Cooper

Adviser
College of Education
ABSTRACT

This experiment compared the effects of a low frequency criterion (i.e., 40 to 80 correct words per minute) and a high frequency criterion (i.e., 180 to 220 correct words per minute) on reading agility, retention, endurance, and Direct Instruction Reading Mastery checkout performances. The participants were three 3rd grade elementary students who were identified by their reading teachers as failing the checkouts in Direct Instruction Reading Mastery. The experimenter conducted daily 10 to 20-min sessions of repeated readings and paired reading with each student. The experimenter measured the number of words read correctly and incorrectly per minute during each session. Also, the experimenter recorded whether or not students passed the Reading Mastery checkouts associated with the stories read during the experiment and Reading Mastery checkouts not associated with those stories. The experiment did not show a functional relation between frequency criteria and retention, endurance, agility, or the Direct Instruction Reading Mastery checkout performances. The data may have implications for future research on developing fluent reading. In particular, the instructional method used to achieve fluent performance may be a critical variable.
ACKNOWLEDGMENTS

This dissertation would never have been completed without the help of many people. First, I would like to thank my wife Cheryl for the tremendous amount of support throughout the entire process, the late nights, the editing assistance, and all the other sacrifices she had to make in order for me to complete the dissertation.

My advisor and mentor, Dr. John O. Cooper gave very generously of his time and I used a lot of it. Drs. Nancy Neef and Timothy Heron provided extensive help with editing of the manuscript.

Dr. Matt Tincani and Dr. Priscilla Brame provided assistance with the conducting the study and the collection of the data. Ms. Hollee Bates and Ms. Jennifer Amos generously spent time listening to audiotapes for the purpose of collecting inter-observer agreement data and treatment integrity. Finally, special thanks to Ms. Christina Sutyak for management of the performance management system (Malott & Harrison, 2002) that finally helped me complete the document.
VITA

June 23, 1969.......................... Born- Bronx, NY

1995.................................. M.S. Behavior Analysis and Therapy,
Southern Illinois University

2000-Present.......................... Director of Education, The Bay School
Santa Cruz, CA

1999-2000.............................. Graduate Assistant,
The Ohio State University

Fields of Study

Major Field: Education
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Vita</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Chapters</td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Literature Review</td>
<td>8</td>
</tr>
<tr>
<td>3. Method</td>
<td>33</td>
</tr>
<tr>
<td>4. Results</td>
<td>49</td>
</tr>
<tr>
<td>5. Discussion</td>
<td>62</td>
</tr>
<tr>
<td>References</td>
<td>73</td>
</tr>
<tr>
<td>Appendix</td>
<td>80</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Last session frequency of correct and incorrect words read per minute during low and high frequency criterion, retention (R), and number of days to the retention test for each student and story numbers from Reading Mastery</td>
<td>51</td>
</tr>
<tr>
<td>2 Frequency of words read correctly at the completion of the low frequency (LF) condition, High Frequency (HF) condition and Endurance Assessment</td>
<td>52</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Moment-to-moment changes in number of correct words per minute for student 1 during low frequency criteria, an assessment of reading endurance.</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>Moment-to-moment changes in number of correct words per minute for student 1 during high frequency criteria, an assessment of reading endurance.</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>Moment-to-moment changes in number of correct words per minute for student 3 during low and high criteria, an assessment of reading endurance.</td>
<td>55</td>
</tr>
<tr>
<td>4</td>
<td>Number of correct and incorrect words read per minute with celeration courses for student 1 across all stories for low and high frequency conditions.</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>Number of correct and incorrect words read per minute for student 2 across all stories for low and high frequency conditions.</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>Number of correct and incorrect words read per minute with celeration courses for student 3 across all stories for low and high frequency conditions.</td>
<td>60</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Learning to read skillfully represents a critical outcome for all elementary school students. The ability to read influences the performance of students in virtually all-academic subjects and other everyday life experiences. For example, effective reading skills make it possible for students to read the newspaper, understand instructions on medications and use the Internet. Also, most jobs require at least some level of basic reading ability.

The educational literatures, including theoretical, conceptual, and experimental, have therefore devoted considerable attention to developing reading with elementary school students. In addition, school districts, legislators, and teacher educators have devoted extensive efforts to the problem of teaching reading to elementary school students. Despite this attention, many students still have difficulty learning to read. About 40% of the U.S. population have reading skill deficits sufficient to hinder their enjoyment of reading and about one in five children do not develop phonemic awareness. Over time, they fall behind in reading and all academic subjects (Grossen, 1997b).
Direct Instruction represents one measurably effective curriculum for teaching elementary students to read. In the largest educational experiment ever conducted, Direct Instruction had dramatically superior results in all academic areas and self-esteem as compared to every other instructional model tested (Engelmann, Becker, Carnine, & Gersten, 1988; Grossen, 1996). The other models of instruction included Behavior Analysis, Parent Education, Southwest Lab, Open Education, TEEM, Responsive Education, Cognitive Curriculum and Bank Street. Some of these models remained very popular despite Direct Instruction showing dramatically superior results. Direct Instruction had a lasting impact (Gersten, Keating, & Becker, 1988) and many replications have supported the original findings (White, 1988). Even with measurably effective instruction, however, some children with severe reading deficits continue to experience difficulties in learning to read and will likely need specialized instruction.

Precision Teaching provides a procedure that may assist these students with severe reading deficits who have difficulty progressing through the Direct Instruction materials. Precision Teaching does not represent a method of instruction or curriculum instead it seeks to improve the use of any method or curriculum. Some research demonstrates that precision teaching improves academic skills such as reading (Beck & Clement, 1991). Precision teachers monitor performance frequencies and display those frequencies on Standard Celeration Charts (Pennypacker, Koenig, & Lindsley, 1972). Making curriculum decisions based on student performances displayed on standard charts defines the critical component of Precision Teaching (Lindsley, 1995; West, Young, & Spooner, 1995).
In Precision Teaching, frequency aims refer to a specific spread of frequencies that define the long-range goal of instruction. Some research supports the finding that high frequency aims have positive effects on student academic behavior. High aims do not mean just going fast or racing. They mean a level of performance that produces certain outcomes including: (a) retaining the skill without practice for greater time than students who have not achieved the same level of automaticity (retention), (b) performing the skill for longer durations of time (endurance), (c) responding well during periods of distractions (endurance) and (d) applying the skill in new situations (application) (Binder, 1996). Precision teachers report that oral reading at a frequency of 180 to 220 per minute may produces REAPS (Retention, Endurance, Application, Performance Standards). In other words, when a students reach a performance standard of 180 to 220 per minute, precision teachers predict they will likely retain the material without practice (retention), perform for long periods of time (endurance) and apply the material to new situations (application).

Direct Instruction and Precision Teaching require a high level of accuracy. Direct Instruction, however, uses lower frequency criteria than most precision teachers recommend. For example, the Direct Instruction Reading Mastery II (Engelmann & Bruner, 1995) curriculum use frequency criteria in the range of 40 to 70 correct words per minute for oral reading (the criterion change as the students move through the curriculum). Most precision teachers recommend a frequency aim in the range of 180 to 220 per minute from the beginning to insure desired outcomes (Binder, 1996). Both Direct Instruction and Precision Teaching require a high level of accuracy.
One way to combine Direct Instruction with Precision Teaching involves charting data from stories read in the *Direct Instruction Reading Mastery* program onto Standard Celeration Charts. It becomes clear that when the students reach a mastery criterion in Direct Instruction (i.e., 40 to 70 correct words per minute), they would not have met the frequency aims correlated with performance standards used in Precision Teaching (e.g., 180 to 220 words per minute).

Some precision teachers argue that the criterion for mastery in Direct Instruction does not meet the performance standards at which students start to achieve the outcomes described by Binder (1996), and therefore some students may have difficulty mastering the curriculum. Once students reach a frequency aim of 40 to 70 words per minute in the Direct Instruction program they progress to new lessons. Precision teachers using the Standard Celeration chart would probably suggest additional instruction on the same lesson.

Lindsley (1987) has suggested that combining Direct Instruction with Precision Teaching will improve reading outcomes for many students. He believed that Precision Teaching could augment instruction using Direct Instruction program in several specific areas, including mastery criteria or frequency aims. Lindsley implied that often the Direct Instruction mastery criteria might not provide sufficient practice for many students to achieve true mastery.
Because research demonstrates that Direct Instruction produces excellent outcomes for most students, raising frequency aims may make little difference for them. One reason why Direct Instruction achieves such successful results may relate to how slowly and carefully the curriculum introduces new material. Each new lesson includes many opportunities to practice skills used in previous lessons. For students academically at-risk, however, Direct Instruction may not provide enough practice for some students at-risk to move successfully through the curriculum. Specifically, Precision Teaching might offer help for students who do not progress at a sufficient rate to pass the frequency and accuracy checkouts in the Direct Instruction Reading Mastery program. If a high frequency criterion produces better application to new situations, students who receive additional practice to attain a higher frequency criterion may have the ability to move through the curriculum, even if they have not had that ability at a lower frequency criterion.

At least two other factors support the belief that frequency criterion may make an important difference to students academically at-risk. First, higher frequency criteria may improve students' ability to read for longer periods of time (Binder, Haughton, & Van Eyk, 1990; Nasir, 1999). The Direct Instruction frequency and accuracy checkouts use long timings (e.g., 3 minutes), which some students academically at-risk will find difficult. Second, higher frequency criteria may improve children's ability to learn to
learn. Researchers have known for a long time that learning one set of material improves the ability to learn another set of material (e.g., Thorndike, 1898). If students academically at-risk improve their ability to learn-to-learn with additional practice and higher frequencies, it would facilitate their movement through the effective curriculum.

The Standard Celeration Chart allows researchers to study the agility of learning. Celeration refers to the “basic unit of measurement of behavior change; change in the frequency per unit time” (Journal of Precision Teaching and Celeration, 1997, p. 123). By creating a measure of celeration that includes the full range of human behavior frequencies, researchers now have a standard way to directly measure the amount of learning per unit of time (e.g., per day, per week, per month, per year). Precision teachers measure agility by observing the changes of successive celerations (Lindsley, 1997). Researchers can compare celerations of the same student learning new material to obtain a direct measure of changes in learning agility of students. That is, if acceleration number values for correct responses multiply, or if deceleration number values for incorrect responses divide as the student progresses to new materials, agility in learning has improved. If a student’s agility improved, this would provide researchers with a standard measurement tactic to measure improvements in learning to learn.
The research procedures used in this dissertation provides for an analysis of the effects of different parameters for the frequency criteria used during story reading (i.e., 40 to 70 words per minute and 180 to 220 words per minute) on successful progress through the Direct Instruction Reading Mastery program. The procedures will define “successful progress” as passing the frequency and accuracy checkouts provided with the Direct Instruction Reading Mastery program. In addition, the study will evaluate the effects of low (i.e., 40 to 70 words per minute) and high frequency (180 to 220 words per minute) parameters on student academic retention, endurance, and agility.

Many procedures have research supporting their use to improve oral reading fluency (i.e., speed plus accuracy and quality). This dissertation uses the procedures of repeated oral reading (Samuels, 1979) and paired reading (Crow, 2001) to develop fluent reading. Probably any of the common methods to improve oral reading fluency could potentially answer the research questions. These two methods received attention because of their common use by precision teachers (repeated oral reading) and by Direct Instruction practitioners (paired reading). In repeated oral reading, the student reads the same passage many times. During paired reading, the student and teacher take turns reading words. The teacher reads one word followed by the student reading one word throughout the passage.
CHAPTER 2

LITERATURE REVIEW

This chapter will introduce the reader to why educators have such wide differences in what constitutes an acceptable performance during instruction. As previously reviewed in Chapter 1, Direct Instruction teachers use only 40 to 70 correct words per minute as an instructional goal for grade two oral reading, precision teachers use 180 to 220 correct words per minute as their instructional goal. It is also important to know what procedures teachers have used to produce those frequencies in reading and what research tells us about the use of those procedures. Finally, this dissertation discusses how the concept of fluency influences the criterion of what constitutes an acceptable performance.

Why do teachers select low or high frequencies?

Teachers who use the Direct Instruction teaching materials have excellent reasons for selecting a low frequency criterion at the start of reading instruction. Very simply, the low frequency criteria used by Direct Instruction teachers are effective with the vast majority of learners (Engelmann, Becker, Carnine, & Gersten, 1988; Gersten, Keating, & Becker, 1988; White, 1988). Direct Instruction teachers do not ignore fluency.
In fact, they consider fluency an important part of teaching reading; however, these teachers consider it important to develop high levels of accuracy first before any emphasis on fluency (Carnine, Silbert, & Kameenui, 1990). Direct Instruction teachers require higher and higher levels of fluency as student progress through the curriculum.

Lindsley (1987), while acknowledging the dramatic success of Direct Instruction programs, argues that higher frequency criteria would lead to improved outcomes. In contrast, he argues that teachers use high frequencies from the beginning. Precision teachers select higher frequencies because of the possibility of obtaining instructional outcomes of retention, endurance, and application of those skills (Binder, 1996). The frequency aim (i.e., goal criterion) for a particular skill, as applied to precision teaching, does not change as the learner moves through the curriculum; the difficulty level increases over time.

This researcher did not find any studies that directly compared the effectiveness of either of these two procedures (i.e., high frequency criterion from the start vs. the addition of high frequency criterion later). Although it remains unclear whether it would be better to aim for high levels of accuracy first and only later attempt to achieve a high frequency as recommended in Direct Instruction, or whether to aim for accuracy and high frequency right from the start of instruction as recommended in Precision Teaching, some published evidence from experimental and clinical literature does exist to support the use of high frequency from the beginning of instruction. First, many data reported by precision teachers and clinicians report dramatic success with the use of high frequency criterion from the beginning (e.g., Desjardins, 1981; Haughton, 1981; Johnson & Layng,
1992). Second, a large-scale implementation of precision teaching with an emphasis on high frequency criterions produced outstanding results in reading, spelling, and mathematics. On the Iowa Test of Basic Skills children in the high frequency instruction group out performed the control group by 20 percentile points in reading and 40 percentile points in math (Beck & Clement, 1991). Third some evidence exists (e.g., Binder, 1996) suggests that high frequency criterions improve the outcomes of retention, endurance, application, and student motivation. Also, high frequencies may improve student agility (i.e., how frequencies change per unit of time “count/per unit of time/per unit of time”, defined as celeration).

Retention

Some studies have found that when students obtain a high frequency criterion it improves retention as opposed using an accuracy criterion alone. For example, Olander, Collins, McArthur, Watts, & McDade (1986) studied retention of college students in a biology class eight months after the completion of the course. Both groups received the same instruction with the exception that one group was required to have a higher frequency criterion during the class. Results showed that after eight months on an unannounced retention test, students in the high frequency group scored significantly higher than students who had only to achieve high accuracy. In another example, Shirley and Pennypacker (1994) found that a high frequency criterion resulted in better retention than an accuracy criterion alone on retention of spelling words in two 8th grade boys diagnosed with learning disabilities.
Several studies in the literature on over-learning have shown that practice after reaching 100% accuracy will improve retention with a variety of different types of learners with different types of skills. For example, Chasey (1977) found that boys with mental retardation had better retention when they received additional practice on a fine motor skill beyond just an accuracy criterion. Chasey and Knowles, (1973) found that adults with mental retardation had better retention on a beanbag-throwing task with additional practice beyond 100% accuracy. Schendel and Hagman, (1982) found that army reserve trainees had better retention when they received additional practice beyond 100% accuracy. Also, latencies between trials get shorter, and these shorter latencies suggest frequencies were higher (Hall & Weneroth, 1972; Judd & Glasser, 1967).

Several issues related to experimental method make all of these studies difficult to interpret. First, it is not clear in these studies that high frequency is specifically what caused the improvements in retention. High frequency is always confounded with increased opportunities to practice. This researcher did not find any studies in the research literature that attempted to isolate the effects of increased opportunities to practice and increases in frequency. Second, although practice beyond 100% accuracy improves retention, some drop off does occur often. Students, however, will re-learn material quicker when they practice beyond 100% accuracy then when they practice only to 100% accuracy (e.g., Schendel & Hagman, 1982).
The application of these findings to reading instruction is not necessarily clear. Direct Instruction programs clearly have children practice well beyond 100% accuracy, but do not require the extremely high frequencies, which would require a lot more practice. Direct Instruction slowly builds on a daily basis. There is very little need to retain without practice, because the students do practice everyday. Approximately 90% of the lesson comes from previously learned lessons, so students do not have to retain without practice to progress successfully through the program. It is difficult to determine if the retention comes from the extensive practice or the higher frequencies they are able to produce after practice.

It seems at least possible that higher frequencies might improve reading instruction as students take long breaks (e.g., summer, Christmas) and may experience setbacks due to poor performance on retention. On the other hand, it is not clear that children in Direct Instruction programs experience difficulties with retention. They receive many opportunities to practice well beyond 100% accuracy, which might assist their retention. They certainly do not receive enough practice to achieve the high frequency aims suggested by precision teachers.

**Endurance**

Endurance refers to the ability of students to respond over long periods of time or when confronted with distraction. At least three types of evidence suggest a relationship between high frequencies and improved endurance. First, two experimental studies suggest high frequencies may improve endurance. Nasir (1999) studied the effects of repeated reading on academic endurance. The students in this study practiced for short
durations only. All students showed improvements on the long duration, despite having no practice at the long duration. Of the 6 students, 2 showed no reduction in frequency even when timed for a long duration. The other 4 students showed only slight reductions in frequency when timed for a long duration. These results suggest that meeting the high frequency aim only once improved performance without specific practice at the long duration. A second experiment (Kim, Carr, & Templeton, 2001) studied the effects of high frequencies on performance over long durations with undergraduate students reading Hindi characters. These researchers found that high frequency aims seemed to produce the ability to respond over long durations and in the face of distractions. These two experimental studies both suggest that the high frequencies produced improvements in endurance; however, neither used a controlled experimental design that could demonstrate functional relationships and rule out threats to internal validity.

Second, Binder (1996) offers some evidence in two sets of unpublished pilot data that suggest a relationship between high frequencies and endurance. Samples of written digits were collected from 75 students for different timing intervals (i.e., 15 s, 30 s, 1 min, 2 min, 4 min, 8 min, and 16 min). Students with low frequencies of writing digits showed dramatically reduced performances over long durations as compared to students who showed high frequencies. Those students had minimal disruptions in performance even over long durations. In a second unpublished study, Binder (1996) reports adults performed two sets of tasks, one with which they could achieve high frequencies and one with which they could not for 3 minute timings. The tests in which the adults could achieve high frequencies included reading numbers and saying answers to simple
addition problems. The tasks that the adults could not achieve high frequencies included reading Hebrew characters or saying numbers in response to the names of Hebrew characters. Distractions interfered with performance only when they had not achieved high frequencies even though performances were highly accurate for both sets of tasks. These correlations at least suggest the possibility that an individual's level of fluency is strongly related to his or her endurance performance.

A third kind of data demonstrates that the length of the timing interval has a powerful influence over the performance of individuals in a variety of tasks. Binder, Haughton, and Van Eyk (1995) found that reducing the timing interval produced improvements in correct responding and reductions in errors of children conducting fine motor tasks. Sroka (1990) found large differences in comparing the first 30 seconds of reading during a one-minute timing to the second 30 seconds. It seems likely that by adapting the strategy used by Sroka (1990) and timing students for a long period of time and measuring moment-to-moment changes in responding may provide an ideal way for researchers to study endurance. Unfortunately, it is very difficult to obtain cumulative records of the performance of elementary children on academic tasks.

Although evidence is clearly limited, the combination of these studies (Binder, 1996; Binder, Haughton, & Eyk, 1996; Kim et al, 2001, Nasir, 1999; Sroka, 1990) suggest the importance that high frequencies may play in the ability to respond at long durations and in the face of distractions. Endurance deserves further study.
The application of these findings to reading instruction is again not completely clear. The teacher at the beginning of reading instruction probably does not require students to perform over long durations and does not allow very distracting stimuli in the classroom, however, there are probably always some distracting stimuli during instruction. If achieving these high frequencies did produce the ability to perform over long durations and respond in the face of distractions, it seems that improving academic endurance would benefit students in beginning reading instruction. Many teachers report that children who have difficulty with instruction also have difficulties paying attention. If high frequencies produce improvements in attention span, it seems likely to benefit learners in early reading instruction.

Application

How Precision Teachers use the concept of application needs some clarification. Precision Teachers argue that behaviors that reach a high frequency are more likely to be used in novel situations (i.e., applied). This concept must be distinguished from how behavior analysts use the technical term “generalization.” In behavior analysis, the technical term generalization should be reserved for stimulus and response generalization (Johnston, 1979). Those generalization processes are well studied in the literature. When Precision Teachers use the concept of application they usually refer to something very different than generalization. In Precision Teaching, application almost always refers to component-composite analysis. This means that teaching some component of a skill to high frequencies will allow the learner to apply that skill to a new situation. For example, say a teacher is attempting to teach a learner to write answers to math facts. Fluent
performance on writing answers to math facts would be 80 to 100 digits per minute. Some students will have extreme difficulty meeting this performance standard. One possible reason involves the student's lack of fluency on a composite, or tool skill. In this case, writing digits 0 to 9 should be at about 180 to 200 digits per minute. If the student writes digits very slowly it seems obvious that he or she will not be able to meet the performance standard for writing answers to math facts (i.e., 80 to 100 digits per minute). Instruction on writing digits until fluent will often have a dramatic effect on writing answers to math facts (i.e., application of the skill).

Very limited experimental evidence exists which suggests that high frequencies have any relationship to application. Binder (1996) reviewed many studies on this issue, but almost all were unpublished reports, discussion articles, or documents nearly impossible to obtain. Still the evidence reviewed by Binder (1996) suggests the need for further study. This researcher located only one published research study that clearly showed a functional relationship between high frequencies and application. Young, West, Howard, and Whitney (1986) found that students with developmental disabilities did not apply their dressing skills to a new situation (novel caregiver) until they reached a high frequency criterion. When the students had only achieved high accuracy, they did not apply their skills.

Other evidence suggests the possibility that high frequencies may produce application to new situations. For example, Jones and Christensen (1999) reported a strong relationship between automaticity of handwriting and the student’s ability to generate written text. They further found that improving a student’s handwriting had a
substantial impact on the student’s ability to generate written text. In a similar manner, Berninger, Yates, Cartwright, Rutberg, Remy, and Abbot (1992) found that automaticity in such skills as writing alphabet letters, and other basic skills predicted the student’s ability in composition. Manis, Doi, and Bhadha (2000) found that second grade student’s skill in naming letters strongly correlated with their ability to decode. Gray, Mulhern, and Neil (2000) found that automaticity of spelling French words predicted how well eleventh grade students did on tests of French language acquisition. Fowler (1993) reported that students have difficulty with comprehension at oral reading frequencies below 200 per minute. None of these studies, however, clearly demonstrated a functional relationship between students obtaining high frequencies and their ability to transfer skills to new situations. This evidence in combination with the studies reported by Binder (1996) and the experimental study by Young et al. (1986), suggest the need for further study of the effects of high frequencies on application.

If high frequency aims produce improvements in a student’s ability to apply skills to new situations, it could have important implications in the area of reading instruction. Effective reading programs like Direct Instruction currently require only low frequency criteria, which may make it more difficult for students to apply their skills to novel situations. This may increase the time required to move the student through different concepts and require more examples than would otherwise be needed to teach the same concepts.
Motivation

Although not specifically mentioned in discussions of REAPs (i.e., retention, endurance, application, performance standards) there are some evidence that high frequencies may improve the student’s motivation. Carbone (2000) argued that a student’s level of fluency could function as an establishing operation (Michael, 1993) that changes the effectiveness of instructional materials as a reinforcer (i.e., the student is motivated to escape the instruction) and increases other behaviors with a history of reinforcement with respect to escape from instruction. Specifically, if a student is not fluent with a particular set of instructional materials, that will serve to decrease the reinforcing value of the instruction as a reinforcer and increase the value of escape as a reinforcer. Once a student becomes fluent with a particular set of materials, this will increase the value of the opportunity to respond to the materials as a reinforcer and decrease the value of escape. Carbone also argues that fluent performances becomes automatically reinforcing (Vaughan & Michael, 1982) that may partially explain a student’s “internal motivation.” In a similar manner, Stanovich (1986) argued that lack of skill in basic reading skills lead to avoidance of reading and the opportunity to improve. It seems likely that fluency in basic skills would prevent this motivational problem.

Kann (1983) provides a general summary of repeated reading and describes how fluency produces improvements in attention and comprehension. Kann also argues that it changes reading from a difficult activity to something fun. Besides showing improvements in reading, children also learn to like reading. He argues this is particularly important for disabled readers who have often experienced failure.
Several studies have found that improvements in fluency led to improved student motivation. For example, Nelso, Leonesio, Shimamura, Landwehr and Narens (1982) found that undergraduates who received overlearning trials improved their "feeling of knowing" or confidence as measured by questionnaire.

Other evidence suggests that motivation might improve with higher frequencies. For example, because evidence suggests that students have poor comprehension at low levels of fluency (Fowler, 1993), it seems reasonable to assume they will get minimal enjoyment out of reading. As comprehension improves with higher frequencies, the natural contingencies of reinforcement may take over.

Agility

Divigi (1976) argued that almost all studies on learning sets used trials-to-criterion as the primary measure. Divigi suggests that it does not give enough information about how the learning set is formed. Specifically, a way to describe progressive changes produced by the training is needed.

Lindsley (1997) described a way to describe the progressive changes produced by the training. He believes the study of changes in celeration (agility) -- how frequencies change per unit of time -- gives precision teachers a unique opportunity to study what Harlow (1949) called learning-to-learn. Harlow and most other researchers studying the phenomena of learning to learn have attempted to do so without standard measurement units. The study of changes in celeration (agility), a standard and absolute unit of measurement, will allow more sensitive measurements of these phenomena than the number of trials-to-criterion.
Learning-to-learn has been demonstrated in a wide variety of tasks. For example, Friedman and Pasnak (1973) demonstrated that learning-to-learn occurs with children who were blind learning Piagetian tasks. Brown and Kane (1988) found learning-to-learn occurs with three- and four-year-old children in solving analogy problems.

In his original work on studying learning-to-learn, Harlow (1949) defined a “learning set” as learning how to solve a particular kind of problem. Harlow found that both monkeys and typically developing children showed similar data in learning-to-learn experiments and they eventually solved many novel training problems on the first try.

Probably the first person to study learning-to-learn was Thorndike (1898). He studied how long it took cats to escape from puzzle boxes. Thorndike found that as cats learned to escape from puzzle boxes, their ability to solve new problems increased dramatically. Often they solved new problems very quickly.

Children academically at-risk continue to fall further behind their peers as they pass through elementary school (Grossen, 1997b). This makes the study of variables that affect learning-to-learn of children academically at-risk essential. Procedures are needed that will teach effective skills to children who are academically at-risk as well as facilitate improvements in these children’s agility of learning so they can learn new material faster. It is critical to determine how frequency affects the learning-to-learn phenomena. If frequency is related to the learning to learn phenomena, this would make the argument for high frequency instruction even more critical. Students who do not achieve high frequencies will fall even further behind because they will not be increasing their ability
to learn new material. On the other hand, if learning-to-learn occurs even at low frequencies, then high frequency instruction might be detrimental to student learning. If students spend additional time learning to high frequencies, it might waste their time when they could have been moving on to new material and learning that new material at a faster pace.

*The Relationship between Frequency and the Desired Outcomes*

The relationship between high frequencies and the desired outcomes of retention, endurance, application, high student motivation and agility is unclear. Most researchers have not distinguished between correlational and functional relationships. It remains unclear whether high frequencies simply predict, or are functionally related to these outcomes.

This distinction is critical for several reasons. First, it may be that children become fluent at different skills simply by engaging in extra practice. It may be that all the effects of high frequencies come simply from having more opportunity to practice those skills and are therefore correlated with high frequencies.

Even if high frequency is a causal factor in the production of REAPs and other outcomes described by Precision Teachers, it seems clear that frequency is not the only variable related to those outcomes. Many other variables have received extensive study in the literature related to these outcomes. Some of these variables included reinforcement, the amount of practice, spacing of the practice, pre-requisite skills, and degree of similarity of task among many others.
Instructional Methods to Produce High Frequencies

Direct Instruction. Direct Instruction does not use high frequency goals initially. It starts by requiring very low frequencies and only slowly increasing the frequency required. Direct Instruction programs require high frequencies over the course of several years, although the frequencies never approach those required by Precision Teachers. The main issue is not whether high frequency instruction is needed. Most educators agree that high frequencies are needed during reading instruction at some point in the curriculum. The issue instead involves deciding at what point in the curriculum teachers should introduce high frequency aims and to what levels.

Direct Instruction is based on research about learning to read. For example, children who do not develop phonemic awareness have extreme difficulty with reading (Watkins, 1997). Wagner and Torgeson (1987) found lack of phonemic awareness causes early reading failure. Research has demonstrated that several types of phonemic awareness tasks have a positive effect on reading acquisition for pre-readers. Watkins (1997) lists the following skills: “rhyming, discriminating sounds that are different, blending spoken sounds into words, word-to-word matching, isolating sounds in words, counting phonemes, segmenting spoken words into sounds, and deleting sounds from words.” Direct Instruction curricula in reading include practice in all of these skills.
Direct Instruction uses a variety of instructional methods, similar to other behavioral education models. Some of these include: teach for mastery, reinforcement, regular measurement of progress, break tasks down, brisk pacing, correction procedures and teach pre-requisite skills (Kinder & Carnine, 1991). According to Kinder and Carnine (1991), however, Direct Instruction also uses four principal instructional design methods that many other behaviorally oriented instructional programs do not always use. Those include:

1. Explicitly teaching the rule or strategy needed and providing guided practice in the use of the rules.

2. Careful selection of examples and non-examples to promote generality.

3. Careful sequencing of examples with minimal differences between examples.

4. Making the process covert by fading leading questions that temporarily make the covert processes overt (e.g., sounding out words).

Over the course of several years the Direct Instruction materials produce high academic frequencies, but they do not require those high frequencies from the beginning. There is extensive amount of evidence that the combined strategies used in Direct Instruction curricula produced outstanding results. In the largest educational experiment ever (Grossen, 1996), Direct Instruction and seven other methods of instruction were compared on a variety of outcome measures for reading, math, spelling, language, and affective measures such as self-esteem. Direct Instruction produced dramatically superior results (Engelmann, et al. 1988). These results had long lasting impact on the academic accomplishments of these students, and the longer they were in the regular curriculum the
less the lasting benefits. For example, students who received the Direct Instruction program in Kindergarten through 3rd grade were more likely to be accepted to college (Gersten, Keating, & Becker, 1988).

Weisberg (1988) evaluated the effectiveness of Direct Instruction in preschools. Children who had Direct Instruction in preschool had better performances on the Wide Range Achievement Test at the end of first grade. White (1988) reviewed the literature on Direct Instruction for special education students. A comprehensive literature search was conducted and 25 studies were located. Of the 25 studies reviewed, 53% of the comparisons favored Direct Instruction. No comparison favored an alternative approach. These studies clearly showed Direct Instruction often is effective with students with a variety of disabilities. In addition to this research, many of the individual components of Direct Instruction have been evaluated and demonstrated necessary elements of the instructional process (Grossen, 1996).

Repeated Reading. Repeated reading represents one well-researched method to improve reading frequency (Samuels, 1979). Some Precision Teachers use that method. Repeated reading does not represent a comprehensive reading curriculum like Direct Instruction, but one instructional method to improve reading performance. In repeated reading, the teacher encourages the student to read the same material many times for short practice times to build the accuracy and frequency at which he or she can read the passage (Samuels, 1979).
Many studies have reported the effectiveness of repeated reading in producing higher frequencies of words read during instruction (Carroll, McCormick, & Cooper, 1991; Carver & Hoffman, 1981; Daly & Guldswog, 1992; Diviaio & Ellis, 1985; Downhower, 1987; Herman, 1985; Polk & Miller, 1994; Rashotte & Torgessen, 1985; Sweeney, Omness, Janusa, & Cooper, 1992; Teigen, Malanga, & Sweeney, 2001). These studies had widely varied outcomes with the amount of generality they reported from none (Polk & Miller, 1994) to excellent (Downhower, 1987; Herman, 1985).

Rashotte and Torgessen (1985) investigated one possible explanation for the differences in generality of reading speed. They found that the degree of generality depended on the degree of similarity between the passages. Passages during early reading instruction with Direct Instruction do have a large amount of similarity. It seems warranted to investigate whether adding a repeated reading component to Direct Instruction will aid in producing retention, endurance and application of skills in children academically at-risk who are not progressing through the Direct Instruction curriculum.

Despite the effectiveness of repeated reading to produce high frequencies, none of the studies reviewed clearly evaluated the effectiveness of different frequency criteria on the outcomes described by Binder (1996). Because this technique has well documented effectiveness, it seems ideal to combine it with Direct Instruction to determine whether improved outcomes for retention, endurance, and application will occur.
Combining Direct Instruction with Precision Teaching

The effects of combining high frequency instruction and/or Precision Teaching with Direct Instruction have not been studied extensively in the literature. At least one study has attempted to combine Direct Instruction with Precision Teaching. Legault, Maloney, and Giroux (2001) found that adding precision teaching to the Direct Instruction Corrective Reading Series improved fluency for all students. It remains unclear how students would have performed without the Precision Teaching component added to the instruction. Although it seems clear from this study that precision teaching did aid in producing high frequency reading, the students' levels of retention, endurance or application remain unknown. In addition, it is possible that Precision Teaching as used in this study could delay the progress of students through the curriculum if students spent extra time to reach a high frequency that was not needed to progress successfully through the Direct Instruction curriculum. Further study combining these instructional techniques seems warranted.

There are many possibilities as to how researchers might combine Direct Instruction with fluency-based instruction. For example, this dissertation and many other precision teachers have simply had students build fluency on the stories at the end of each Direct Instruction lesson. Another possibility might be to accept the tested fluency aims used in the Direct Instruction curriculum and build fluency only on basic skills (e.g., saying letter sounds, saying blends). Research is needed on what aspects of the different possible performances would lead to the most benefit for students.
In addition, many school programs have reported that the combination of Direct Instruction with Precision Teaching has been effective. Morningside Academy (Johnston & Layng, 1994) and Quinte educational services (Maloney, 1998) are two schools in particular that have reported outstanding results. Morningside Academy even gives a money back guarantee for students who are two years behind if they do not gain two years in reading and math in one year. Morningside Academy has not been required to give a refund in many years, suggesting that the procedures they use produce results. Some might argue that the guarantee does not represent a meaningful measure, because it might be possible to gain two years on standardized tests by answering only one or two additional questions correctly.

*How to define fluency?*

Binder (1996) reports that most people understand fluency without any need for an explanation of the concept. He gives several definitions or descriptions of the concept of fluency. Essentially he argues most people think of fluency as the combination of accuracy plus speed. He also gives an empirical definition of fluency based on its measured effects. Specifically, a behavior is said to be fluent if a learner achieves a certain frequency of performance and has the ability to (a) retain the information without practice, (b) endure for long periods of time even in the face of distraction, and (c) apply the learning to new situations. These outcomes define “true mastery” for Binder and many precision teachers. If the three forms of behavior occur the behavior is said to be fluent, whereas if they do not occur the behavior is said to be not fluent. Several teachers and researchers have suggested performance standards for different skills that have been
correlated with the occurrences of retention, endurance, and application (REAPS). For example, for writing digits 0 to 9 the suggested performance standard is 160-180 digits per minute. Despite extensive discussion of these performance standards, (Binder, 1996; Johnson & Layng, 1992; Lindsley, 1996), their empirical bases are not evident in the experimental literature. As reviewed previously, some minimal experimental evidence exists that high frequencies may predict or produce retention, endurance and application.

This functional definition of fluency requires clarification in several areas in order for it to become a useful concept in behavior analysis. First, it needs clarification in regards to a clear definition of when the high frequency has been achieved. The only published discussion of this point this researcher could find was in Johnson and Layng, (1992). They discussed this point briefly arguing simply that achieving a high frequency only one time will not always predict or produce REAPS. It is essential that researchers further clarify exactly what it means to achieve the performance standards. For example, what if a student achieves the performance standards two times, or five times? Do the performances have to be consecutive? Others have argued it is not just the frequency but also the consistency with which the learner can achieve the frequency (personal communication, John O. Cooper). Without a clear definition of the criterion performance it becomes difficult, if not impossible to separate the effects of high frequencies from the effects of receiving extra time to practice a skill.
Second, the functional definition of fluency needs clarification on what would constitute criterion performance on retention, endurance, and application. For example, is a skill fluent if a learner retains and endures, but does not apply? Or if they retain at 85%? How long do they have to retain?

Third, writers who have discuss fluency must distinguish between correlations and demonstrating functional relationships. This represents an important issue. If high frequencies only predict these phenomena, then research on discovering the critical variables that are functionally related to the phenomena is needed. If high frequencies are functionally related to retention, endurance, and application, then procedures for developing those high frequencies are important to discover.

**Evidence for Fluency**

Fluency and Precision Teaching have generated a lot of interest among behavior analysts. For example, Eshlemann (1990) reports there were 884 precision teaching presentations and about 316 publications during the years 1964-1990. A quick glance at the Association for Behavior Analysis program guide in the last few years will confirm the enormous growth of interest in fluency among behavior analysts. Unfortunately, all of this attention has yielded very little in the way of publications of actual research with data to support the claims made by those who support fluency-based instruction. Sidman (1994) argues, "If you have not published your research, then as far as the rest of the world is concerned, the research has not been done (p. 120)." He gives two reasons for this position. First, science is a social enterprise. Failure to share the results, even if the highest quality scientific methods have been used mean that the results will have little
impact on the behavior of others. Second, even if the results are shared in other ways (e.g., conference presentations) without having textual representations many of the details are lost. It becomes impossible to compare work of different scientists or evaluate another’s work. Most scientists will not change their behavior without publications that demonstrate the usefulness of new procedures (Grossen, 1997a).

Lindsley (1992) has argued against publications in precision teaching for two reasons. First, it becomes difficult to change terms that have become embedded in the text after publications, and with experience, some of these terms and methods might not represent current best practices. In addition, Lindsley (1994) reported that most precision teachers were not in academia and did not need to publish in order to obtain promotions. Often precision teachers speak of the “Oral Tradition” as a way to communicate among professionals rather than formal scientific publications.

Although some limited published studies with data exist, it remains unclear whether teaching to high frequencies does predict or produce the outcomes described by Binder (1996). One of the most common applications of fluency-based research is to the field of education. High quality research on the effectiveness of these procedures is essential for the field of education and behavior analysis.

Summary

Some children even in science-based academic curricula (i.e., Direct Instruction Curricula) will still have difficulty learning to read. There is some evidence that fluency-based instruction may produce improved outcomes in retention, endurance, application, increased motivation and agility. There is a need for researchers to
investigate fluency given the increased interest, but limited scientific evidence. Currently, it remains unclear whether adding a fluency component to the Direct Instruction Curriculum will produce superior results. In fact, it could potentially produce poorer results if children take longer to move through the curriculum in order to achieve extremely high frequencies at each step in the curriculum.

Purpose

This research specifically sought to determine the effects of low and high oral reading frequencies on successful progress through the Direct Instruction Reading Mastery curriculum. Of particular interest was whether students would retain more without practice, endure for longer periods of time, apply what they had learned to new material (i.e., the frequency and accuracy checkouts), and show improvements in agility.

The study used typical second graders in a Direct Instruction Reading Mastery program. The students were taught to the low frequency criterion suggested by the Direct Instruction Reading Mastery program, and to a high frequency criterion suggested by Precision Teachers. The experimenter measured the effects on retention, endurance, and application of skills as well as changes in agility.

The experiment sought to control for the effects due to practice by using the same subjects and educational material in both the low frequency and high frequency conditions. This allowed for direct comparisons between low and high frequency criteria on the outcomes.
Research Questions

1. What effects will repeated oral reading to a frequency of 40 to 80 correct words per minute and 180 to 220 correct words per minute have on passing the frequency and accuracy checkouts that accompany Direct Instruction Reading Mastery?

2. What effects will repeated oral reading to a frequency of 40 to 80 correct words per minute and 180 to 220 correct words per minute have on students' retention of stories that accompany Direct Instruction Reading Mastery curriculum?

3. What effects will repeated oral reading to a frequency of 40 to 80 correct words per minute and 180 to 220 words per minute have on students' endurance (i.e., ability to perform for long periods of time) on reading passages that accompany Direct Instruction Reading Mastery?

4. What effects will repeated oral reading to a frequency of 40 to 80 correct words per minute and 180 to 220 correct words per minute have on the acceleration multiples (agility) of reading passages that accompany Direct Instruction Reading Mastery?

5. What is the relation between the frequency criterion and acceleration multiples (agility) to the criterion and passing the frequency and accuracy checkouts?
CHAPTER 3.

METHOD

This chapter describes the participants, setting, measurement procedures, experimental design, dependent and independent variables. In addition, it provides a detailed description of the procedures used in this study.

This research attempted to assist children who failed frequency and accuracy checkouts from the Direct Instruction Reading Mastery program. These checkouts provide teachers with data on how well students progress through the program.

Participants

The participating school had a Title One program. This federally funded program provides money to support children who are behind academically in reading or mathematics in elementary grades. The Title One reading teacher at the school recommended nine students for participation in the study. The experimenter reviewed attendance records and eliminated two students from participation because of a history of extremely poor attendance. One student did not return his permission form. For a student to qualify for participation in this research she had to fail a minimum of four of her last five frequency and accuracy checkouts. The school did not have records of these children's frequency and accuracy checkouts. Therefore, the experimenter tested the remaining six children and two of these students did not qualify for the study because they passed more than
four of the last five frequency and accuracy checkouts at that time. The remaining four students qualified for the experiment. During the course of the experiment, one of these four students had excessive absences, and the experimenter dropped her from the study. The remaining three students participated in the study. All three students were female African-Americans nine years of age. All three students participated in-group instruction in *Direct Instruction Reading Mastery II* (Engelmann & Bruner, 1995). All three students were general education students, but during the course of the study one student received a referral for an evaluation for special education and qualified.

*Setting*

A local urban charter school in Columbus, OH that provided instruction using both the *Direct Instruction Reading Mastery* program and Precision Teaching served as the setting for the research. During the 2000-2001 school year, the school had 558 children enrolled in kindergarten through fifth grade. Of the enrolled students, 99.7% of the students were African-American and 74% of the students qualified for free or reduced-price lunch.

The selected students received 10 to 20 minutes of daily individual instruction to supplement their regular scheduled reading programs. The experimenter coordinated with the classroom schedule of instruction so that the students did not miss important classroom activities. The teacher selected the best time for students to attend the individual instruction. The instruction took place in the Title One teacher’s classroom, the school library, the teacher’s lounge, and special education classroom depending on available space. Each of these locations had a desk with two chairs that allowed the
experimenter to sit next to the student and an outlet for the tape recorder. Approximately the first half of the experiment took place in the Title One teacher's classroom. In addition to the desk where the experiment took place, this room had another teacher's desk and two rectangular tables that seated between six to eight children at each table.

The school decided they needed the space for additional Title One students. The second half of the experiment took place in the teacher's lounge. In a similar manner, the experiment took place at a desk so the experimenter could sit next to the student. In addition, the location had an outlet for the tape recorder. The teacher's lounge had other desks and a few couches. On occasion, the school used the teacher's lounge for staff trainings or meetings. During those times the experiment took place in either the special education classroom or the library depending on the availability of those rooms.

Experimenters

The primary experimenter was a full-time student in the Ph.D. program in Special Education at The Ohio State University. He had a Masters of Science Degree in Behavior Analysis and Therapy and experience with many different types of educational settings such as regular education classrooms, special education classrooms and private school classrooms. He had special education teacher certification in New Jersey and certification as a Board Certified Behavior Analyst. He conducted 93% of the experimental sessions. Two other full-time Ph.D. students in the Special Education program at The Ohio State University served as experimenters when the primary experimenter could not conduct the experiment on a particular day. They had Masters Degrees in Special Education and experience and training in the use of Precision Teaching and Direct Instruction.
Definition and Measurement of Dependent Variable

This experiment used three dependent variables, the frequency and accuracy checkout performance from the Direct Instruction Reading Mastery Program,\(^1\) correct and incorrect words read per minute from stories in the Direct Instruction Reading Mastery program. At the end of each lesson the students read a story. Words read correctly and incorrectly served as two of the dependent variables. Every fifth lesson, the program selects a portion of one of these stories as a frequency and accuracy checkout. The student’s performance on these checkouts served as the third dependent variable. The experimenters audio taped all oral-reading assessments.

Definition of correct and incorrect words read. This study adapted the definition of correct and incorrect words from Nasir (1999). Correct words corresponded to the conventional word name. Incorrect words included 5 different categories.

1) Omissions or words skipped (e.g., the passage says “Walter was sad” and the student read “Walter sad.”

2) Insertions or words added to the original passage (e.g., the passage read, “Walter was sad” and the student read, “Walter was still sad.”

---

\(^1\) Readers familiar with the Direct Instruction Reading Mastery program may wonder about the use of the term “frequency and accuracy checkout.” The program calls these checkouts “Rate/accuracy checkouts.” Confusion exists between the differences in the terms frequency and rate. In behavior analysis most use rate to mean number of response per some unit of time, and frequency to mean number of responses per a standard unit of time. In statistics the term frequency has acquired other meanings (e.g., number of scores in a distribution). Most other physical sciences use the term frequency to mean number of response per unit of time. Precision teaching in an attempt to stay consistent with other physical sciences and to avoid confusion with other meanings of rate (e.g., how do you rate?) use frequency to mean number of responses per unit of time. This dissertation will avoid the use of the term rate, and will use frequency to mean number of responses per unit of time.
3) Substitutions or reading different words (e.g., the passage says, “mother” and the student says “mom.”)

4) Mispronunciations (e.g., the passage said “can” and the student read “cane.”)

5) Self-corrections (e.g., the passage read, “Walter was sad” and the student read “Walter was still sad, no wait, Walter was sad.”) Self-corrections were counted as errors to provide a conservative measure of accuracy.

*Measurement of oral reading.* The experimenter instructed the student to read as many words as possible during the timing period and if she reached the end of the passage to start over at the beginning of the story. While the student read, the experimenter followed along on a duplicate copy of the *Direct Instruction Reading Mastery* story. The experimenter recorded all errors on the duplicate copy of the story. At the end of the timing, the experimenter counted the total words read correct and incorrect and recorded it on the data sheet and praised the student for attending and participating. The experimenter then charted correct and incorrect words read on Standard Celeration Charts using the charting conventions of Pennypacker, Koenig, and Lindsley (1972).

*Frequency and accuracy checkouts.* The experimenter recorded correct and incorrect words read per minute in the same manner as described in measurement of oral reading with one exception. The experimenter told the student that if she completed the story, she could stop reading and didn’t have to start over again at the beginning of the story. In addition, the experimenter recorded each passage as pass/fail according to the frequency and accuracy criteria for the specific lesson as specified in the *Direct Instruction Reading Mastery.*
Inter-Observer Agreement, Reliability of Measurement, and Treatment Integrity

Inter-observer agreement. The experimenter trained two individuals for the purpose of collecting inter-observer agreement data. During observer training, the experimenter provided observers with a copy of the written response definitions and measurement procedures used in the study. The experimenter discussed these definitions and measurement procedures with the observer and clarified any questions. Finally, the experimenter and observer practiced scoring from audiotapes made for this purpose. During these sessions both the observer and the experimenter scored from audiotapes. The experimenter and observer discussed any disagreements and clarified response definitions. After the experimenter and each observer obtained at least 90% agreement for 3 consecutive stories without discussion or interpretation of response definitions they began collecting data for the experiment.

The observers counted the number of correct and incorrect words read during each reading to provide data for an assessment of inter-observer agreement (IOA). The observers recorded student performances during 30% of the days of the experiment and approximately 30% for each phase of the study. The experimenter selected IOA sessions on a semi-random basis by writing dates on slips of papers and blindly selecting sessions. The observers could not score some sessions from audiotapes due to poor tape quality or background noise. In those situations the experimenter selected another session from the remaining slips of paper. Observers measured inter-observer agreement regularly throughout the experiment. Additional inter-observer agreement occurred at the end of
the study. The experimenter calculated the percentage agreement using the formula number of agreements divided by the number of agreements plus the number of disagreements and multiplied by 100%. Data presented on Standard Celeration Charts represent the data collected by the experimenter in situ.

*Reliability of measurement.* To assess the reliability of measurement, the primary experimenter reported a percentage of agreement between his measurements of student performances in-situ with his measurements obtained from the audiotape recordings. Reliability assessment occurred in exactly the same manner as inter-observer agreement with the exception the primary experimenter listened to the audiotapes a second time. The experimenter listened to the audiotapes at least one week after conducting and experimental session.

*Treatment integrity.* Because the research questions in this study focused on the student's oral reading frequencies, the dependent variable measurement also served as a measure of treatment integrity. This experimenter assessed the integrity of the independent variable, i.e., oral reading frequencies, by measuring the number of words read correctly and incorrectly for each student. The experimenter used the same response definitions and procedures to calculate inter-observer agreement and the reliability of measurement as described in the section on Definition and Measurement of the Dependent Variable.
In addition, the experimenter developed a checklist (Appendix E) used to assess the integrity of the measurement procedures. Two observers independently observed the implementation from audiotapes on 30% of the days of the experiment and calculated inter-observer agreement as described in definition and measurement of the dependent variable.

Experimental Design

This study used a multiple-baseline design across subjects (Barlow & Hersen, 1984) to assess the effects of the low and high frequency conditions on endurance and agility. This design allows for the demonstration of functional relationships in which the experimenter cannot or will not reverse the effects of the independent variable (Barlow & Hersen, 1984). Students learned to read stories in the low frequency condition and the high frequency condition sequentially across two students. The third student participated in an A-B design, which did not allow for demonstrations of functional relationships (Barlow & Hersen, 1984), but could provide an additional replication of any effects observed in the multiple-baseline design.
In addition, the experiment used a variation of the alternating treatments design (Barlow & Hayes, 1978). All students received a baseline on frequency and accuracy checkouts. Then students received instruction in the lessons associated with one of those frequency and accuracy checkouts in the low and high frequency conditions. In addition, the experimenter continued to monitor performance on frequency and accuracy checkouts not associated with the lessons in which the students received instruction in an attempt to discover if a functional relationship existed between the frequency criterions and passing the frequency and accuracy checkouts.

Procedures

Materials

1. **Frequency and Accuracy Checkouts** from the Direct Instruction Reading Mastery


3. **Cassette Recorder and tapes**

4. **Standard Celeration Charts, Dpmin-10ESDaily .001-1,000 X 140 days, Behavior Research Company, Box 3351, Kansas City, KS 66103-3351**

4. **Timings Chart, Tpmin-3ECDaily Timings Chart, Behavior Research Company, Box 3351, Kansas City, KS 66103-3351**
**Motivation System**

Throughout all conditions students received points, which they could trade in one time per week for prizes from the “store.” The store consisted of small toys, school supplies, candy and other similar items. The students received points for participation and improvements in academic behavior.

**Selection of Lessons**

The experimenter started 25 lessons behind where the students received instruction in the regular class, based on the fact that they had failed the frequency and accuracy checkouts of those lessons. Student 2 had so much difficulty with this lesson, she started at the very beginning of Reading Mastery II, having failed all frequency and accuracy checkouts given to her.

**Selection of Timing Intervals**

For the frequency and accuracy checkouts, the Direct Instruction materials give an individual timing interval for each story, but for the other stories the experimenter selected a timing interval based on the length of the story and the frequency aim for that story. For example, if a story contains 180 words and the frequency aim was 180 words per minute the timing interval was 1 minute. If the frequency aim was 60 words per minute for the same passage, the timing interval was 3 minutes.
Conditions of the Study

Pre-Tests. Before starting instruction on each story the instructor conducted a pre-test using the measurement procedures described above.

Retention Tests. After the student met the frequency aim for each story, the experimenter conducted a retention test approximately 7 days after the student met aim for that story. The experimenter conducted the retention test in exactly the same manner as described above. If the experimenter could not conduct a test 7 days after due to student absence or school holidays, the experimenter conducted the test on the next possible day. The experimenter made attempts to match the number of days in both conditions. When the experimenter could not match the number of days, the low frequency condition had a shorter retention period than the high frequency condition.

Baseline

The experimenter collected data on the frequency and accuracy checkouts. The experimenter used the frequency and accuracy checkout material from the Direct Instruction Reading Mastery program and the timing interval specified in the particular frequency and accuracy checkout. The experimenter used the same measurement procedures described above with one exception. If a student completed the story before the end of the timing interval, she did not have to return to the beginning of the story and start again. The students received a score of pass or fail for each frequency and accuracy checkout.
Repeated Reading to a Low Frequency Criterion

Researchers and teachers refer to the procedure used in this study for developing oral reading to the fluency as “Repeated Reading” (Samuels, 1979). The tactic for repeated reading has the student practice reading the same text each session until the reading performance meets the established instructional aim. Orally reading 40 to 80 correct words per minute establishes the instructional aim for the low frequency criterion.

To achieve this instructional aim easily, students practiced only a small part of the passage for short durations of time, e.g., 5 to 10 seconds. Precision teachers (Binder, 1996) call these short duration practices “sprints.”

Sprint practice. The experimenter established a sprint frequency aim each session based on the student’s oral reading performances during the pre-test of the story. Specifically, the experimenter set an individual sprint aim at 20 words per minute higher than the performance on the pre-test. For example, if during the pre-test the student read 50 words correct, the experimenter set the first frequency aim for sprints at 70 correct per minute, or 12 correct words per 10 seconds (rounded from 11.6 words per 10 seconds. The experimenter set a countdown timer for 10 seconds and said “start when you feel ready” or something equivalent and started the timer when the student started reading. The student stopped reading with the signal from the countdown timer indicating the elapse of 10 seconds. The experimenter and the student then counted the number of
words read correctly and incorrectly and charted the sprint performance on timings charts (These charts do not show celerations- data charted by session/trial not charted in real calendar time). If the student met the sprint aim, the student started another sprint practice beginning immediately after the last word she read in the previous sprint. The experimenter charted each trial on a new “timing” line.

If during any of the sprints, the student did not meet her session aim, the student repeated that same text up to 4 times. After each attempt, the experimenter and the student charted the sprint performance on a timings chart. Each time the student repeated the same sprint the experimenter charted the data on the same “timing” line. If the student did not meet their sprint aim after 4 attempts the experimenter and student moved on in the text to the next sprint practice.

If the student met all sprint session aims on the first timing during a session, the experimenter moved the sprint aim to 20 words per minute higher for the following session. The experimenter increased the sprint aim more than 20 words per minute if the student progressed at a fast pace and the present performance would have met or almost met that criterion.

The experimenter delivered praise for attending, increased correct frequencies of oral reading, and other appropriate behaviors. Also, the experimenter provided error correction on the student’s performances. The experimenter used a model/lead/test procedure for correcting errors. The experimenter told the student the word (model), had the student say it with the experimenter (lead), and asked the student to say the word by herself (test). The student practiced reading difficult words and phrases between sprints if
the experimenter had used the error correction procedure several times and the student continued to make errors. The student practiced reading the difficult word or phrase until she read it three times correctly without errors. If the student exceeded the frequency criterion for the low frequency condition, but failed to meet the accuracy criterion the experimenter no longer conducted repeated reading, but only practiced reading the difficult words or phrases until the student met the accuracy criterion.

*Paired reading.* At the end of sprint practice each day, the experimenter practiced reading the selected passage with the student using the paired reading technique 1 to 3 times. In the paired reading technique, the experimenter read one word, followed by the student reading the next word and continuing throughout the passage (Crow, 2001). At the end of the paired reading techniques, the experimenter timed the student in the manner described below.

*Assessment of oral reading.* The experimenter determined the length of the timing period for the assessment by the number of words in the passage the student read. The assessment used the same text as the student used for sprint practice and paired reading. The experimenter ensured that the student has sufficient text to read at a pace of 200 words per minute by using the number of words in the text to establish the duration of the timing. For example, if a student reads a passage with 100 words the experimenter set the timing duration for that passage at 30 seconds.

The experimenter told the student to orally read as many words as possible and to stop reading at the signal from the countdown timer. At the end of this assessment, the experimenter and the student counted the number of words read correctly and incorrectly.
The experimenter gave praise for participating and attending, but did not provide specific feedback on reading performance. The student received, however, general feedback on her reading performance (e.g., “you read great, you tried hard, great job!”) The experimenter charted the performance on Standard Celeration Charts according to the conventions described by Pennypacker, Koenig, and Lindsley, (1972).

Repeated readings, paired reading, and assessments continued until the student reached instructional aim of 40 to 80 words per minute for each story from the Direct Instruction Reading Mastery program that corresponded to the frequency and accuracy checkout assigned by those instructional materials. After meeting aim on all corresponding stories, the student took the frequency and accuracy checkout associated with those stories. The procedures were the same as those described for baseline and the next higher level frequency and accuracy checkout for which the student did not receive any instruction.

Repeated Reading to High Frequency Criterion

The procedures for the condition using a high frequency criterion followed the procedures described for the low frequency criterion exactly with one exception. The high frequency condition used a criterion of 180 to 220 correct words per minute rather than the 40 to 80 correct words per minute used in the low frequency criterion. Lessons three and four of program had word lists for students to practice instead of connected
reading passages. Therefore, it required separate instructional aims for the high frequency condition as opposed to all other lessons. This experiment used an instructional aim of 90 to 130 per minute for reading the stories with word lists as opposed to 180 to 220 for reading passages.

_Assessment of Academic Endurance_

Binder (1996) defined endurance as “the ability to perform for extended periods of time as a function of initial performance frequency (p. 175).” For example, if a student reads a passage at 180 words per minute during a one-minute timing and then the student receives a 3 minute timing on the exact same passage can the student still perform at the same frequency?

This experiment measured endurance after the student met aim in the high frequency condition. At that time, the experimenter conducted a longer timing equal in length to the timing for the same story conducted during the low frequency condition. In addition, for students 1 and 3, the experimenter listened to audiotapes of the sessions where the student met aim in the low frequency condition, and the long duration timing. The experimenter then recorded correct words read per minute every 10 seconds and charted the data on Standard Celeration Charts. This allowed for analysis of moment-to-moment responding during the timing periods.
CHAPTER 4

RESULTS

This chapter describes results of measurements obtained for the dependent variables, independent variables, inter-observer agreement, reliability, and treatment integrity. Refer to Chapter 3 for descriptions of the procedures used for collecting all of the above.

Dependent Variables

Frequency and Accuracy Checkout Performances. The students failed all frequency and accuracy checkouts during baseline and during the low frequency condition. Student 1 failed the frequency and accuracy checkout associated with the lessons studied during the experiment. She did pass the frequency and accuracy checkout on the next set of lessons not studied during this experiment. Although she improved substantially on both frequency and accuracy checkouts, she improved more on the set of lessons in which she did not receive instruction during the experiment. Student 2 failed all frequency and accuracy checkouts during the high frequency condition. Student 3 did not have time to complete the experiment due to the school year finishing and did not complete the frequency and accuracy checkouts in the high frequency condition.
Retention. Results for retention are presented in Table 1, and link to research question two. Due to school holidays and excessive student absences the experimenter could not match the number of days in the low and high frequency conditions most of the time. The high frequency condition always had a longer number of days when this occurred. Overall, the data do not reflect a clear pattern of superiority for either condition.

Endurance. The effects of low and high frequencies on endurance were evaluated with Students 1 and 3 in the context of a multiple baseline design across students, and link to research question three. Figures 1 and 2 show moment-to-moment changes in responding during low and high frequency conditions respectively. The high frequency condition does seem to have a very slight reduction in bounce compared to the low frequency condition. Figure 3 shows the moment-to-moment changes in responding for student 3. This student had inconsistent findings for different stories during the low frequency condition and only had time to complete one story in the high frequency condition before the end of the school year.

Table 2 shows data collected in the pre-test, last performance and post-test for endurance. All students showed improvements in endurance after practice at high frequencies. All of the students could perform for longer periods of time without direct practice at longer timing periods. Performance for student 2 dropped considerably during the endurance test in both conditions, but clearly had improved performance in the high frequency condition.
Table 1

<table>
<thead>
<tr>
<th>Student</th>
<th>Story Number</th>
<th>Frequency Criterion</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Last R Days</td>
<td>Last R Days</td>
</tr>
<tr>
<td>1</td>
<td>41</td>
<td></td>
<td>85/1 77/2 11</td>
<td>207/1 158/2 12</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td></td>
<td>68/1 70/2 9</td>
<td>198/1 125/2 20</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td></td>
<td>56/1 69/1 8</td>
<td>200/1 139/0 8</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td>76/1 76/1 7</td>
<td>239/0 179/7 12</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>67/1 94/2 10</td>
<td>200/0 158/14 11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>63/0 30/12 9</td>
<td>180/0 78/6 10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
<td>66/0 62/3 6</td>
<td>120/0 78/0 11</td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td></td>
<td>81/1 55/4 7</td>
<td>206/0 227/0 7</td>
</tr>
</tbody>
</table>

Last session frequency of correct and incorrect words read per minute during low and high frequency criteria, Retention (R), for each student and story number from Reading Mastery.
Table 2

<table>
<thead>
<tr>
<th>Student</th>
<th>Story Number</th>
<th>LF</th>
<th>HF</th>
<th>Endurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41</td>
<td>85</td>
<td>207</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>68</td>
<td>239</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>74</td>
<td>210</td>
<td>238</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>56</td>
<td>200</td>
<td>147</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>76</td>
<td>239</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>67</td>
<td>202</td>
<td>145</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>63</td>
<td>180</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>66</td>
<td>123</td>
<td>83</td>
</tr>
<tr>
<td>3</td>
<td>41</td>
<td>81</td>
<td>206</td>
<td>213</td>
</tr>
</tbody>
</table>

*Frequency of words read correctly at the completion of the low frequency (LF) condition, high frequency (HF) condition, and Endurance assessment.*
Figure 1: Moment-to-moment changes in number of correct words per minute for student 1 during low frequency criteria, an assessment of reading endurance.
Figure 2: Moment-to-moment changes in number of correct words per minute for student 1 during high frequency criteria, an assessment of reading endurance.
Figure 3: Moment-to-moment changes in number of correct words per minute for student 3 during low and high criteria, an assessment of reading endurance.
Agility. The data presentation in the agility section link to research questions four and five. Figure 4 shows correct and incorrect words read per minute and celerations for student 1. During the first story in the low frequency condition (i.e., story 41) corrects accelerated at X1.15 per minute, per week. The next two stories accelerated at X2 and X3 respectively. The experimenter could not quantify the acceleration for last story number 44 (i.e., the last story) in the low frequency condition since the student learned story 44 after the collection of just one data point, however, clearly it would have a value higher than X3 achieved in the previous story. In order to measure agility, the experimenter needed specific measures of celeration. To obtain a quantitative measure of celeration requires at least 2 data points separated by some time interval (Pennypacker, et. al., 1972). While possible to calculate celeration with only 2 data points it clearly has little value without more data. When calculating celeration with only 2 data points values might change largely due to chance. Celerations remained stable throughout the high frequency condition. The number of correct words read per minute, per week spread between X1.1 and X1.3.

Student 2 did not produce any usable celerations during the course of the experiment as shown in Figure 5. This is due to the fact that some of the celerations represent words read from stories and some of the celerations represent words read from story lists, making the celerations not comparable.
Figure 6 shows correct and incorrect words read per minute and celerations for Student 3. Number of correct words read per minute, per week during the low frequency condition accelerated with successive stories. Correct words accelerated X1.5 per minute, per week during the first story and X1.2 per minute, per week during the second story. The third and fourth stories accelerated at X30 and X7 respectively. Student 3 only completed one story in the high frequency condition and therefore no measure of agility is possible. Measurement of agility required at least two measurements of celeration in the same condition.

*Inter-Observer Agreement of Words Read Correct and Incorrect per Minute*

Inter-Observer agreement on the number of words read correctly and incorrectly for student 1 averaged 98.7% with a range of 94% to 100%. For Student 2 it averaged 96.9% with a range of 89% to 100%. For student 3 it averaged 98% with a range of 91% to 100%.

*Reliability of Words Read Correct and Incorrect per Minute*

Reliability on the number of words read correctly and incorrectly for student 1 averaged 98.5% with a range of 97% to 100%. For student 2 it averaged 97% with a range of 80% to 100%. For student 3 it averaged 97.5% with a range of 91% to 100%.

*Independent Variables*

*Treatment Integrity.* Treatment integrity refers to how consistently the experimenter implemented the independent variable. This dissertation had one independent variable called frequency, and presented at two levels. The two levels of the independent variable consisted of “low” frequency and “high” frequency. Low frequency
Figure 4: Number of correct and incorrect words read per minute with celeration courses for student 1 across all stories for low and high frequency conditions.
Figure 5: Number of correct and incorrect words read per minute for student 2 across all stories for low and high frequency conditions.
Figure 6: Number of correct and incorrect words read per minute with acceleration courses for student 3 across all stories for low and high frequency conditions.
was defined as the frequency recommended by the Direct Instruction Reading Mastery curriculum based on the individual lesson. This spread from 40 to 80 correct words per minute. This dissertation defined the high frequency condition as 180 to 220 correct words per minute as defined by precision teachers as the range of frequencies that usually predicts REAPs (i.e., retention, endurance, application, performance standards) (Binder, 1996). During the measurement of the dependent variables (i.e., frequency and accuracy checkouts, retention) students achieved a frequency in the appropriate range 100% of the time.

Accuracy of Implementation of Measurement Procedures

On 30% of the sessions observers scored the accuracy of implementation of the measurement procedures by listening to audiotapes of the sessions and scoring each item in Appendix E as correct or incorrect. The primary observer recorded the experimenter’s implementation of the measurement procedures as accurate a mean of 97% of the time (range 71% to 100%).

Inter-Observer Agreement of Measurement Procedures

Inter-observer agreement on the implementation of the measurement procedures was 99% (range 86 to 100%). In addition, the experimenter listened to the tapes and used a different timer to assess the measurement of the timing interval. The experimenter scored an agreement if the timing interval matched the original timing interval within 3 s, due to the difficulty of starting the timer at the exact moment. On 100% of the sessions assessed the timing matched within 3 seconds.
CHAPTER 5
DISCUSSION

The findings do not suggest any functional relationships between the frequency levels of correct responding and passing the frequency and accuracy checkouts, retention, endurance, or agility measures. The data, however, may have implications for future research on repeated reading, fluency, Direct Instruction and Precision Teaching.

_Intraverbal vs. Textual Control During Repeated Reading_

This dissertation intended to investigate the effects of high and low frequencies on different aspects of academic performance. The research methods did not consider how the students obtained those frequencies. Anecdotal evidence collected during this research, however, suggests that it may make a critical difference. Research suggests individuals often have little or no generality among Skinner's (1957) verbal operants (Chase, Johnson, & Sulzer-Azaroff, 1985; Lamarre & Holland, 1985; Partington & Bailey, 1993; Twyman, 1996; Watkins, Pack-Teixtera, & Howard, 1989). This dissertation and other research that has studied the effects of repeated reading have assumed that they actually investigated reading, or in terms of Skinner (1957) "textual behavior." Textual behavior is controlled by the written stimulus. During the course of the experiment, the students clearly showed that their behavior was often not controlled by textual stimuli. Some examples of errors that students made that demonstrated the
students' behaviors were not controlled by textual stimuli included, substituting the name of a character in the story for the pronouns he or she, reading mom for mother, reading words or whole sentences from other parts of the story. In addition, during the high frequency condition, one of the children said to the experimenter, "I know the whole story by heart." The experimenter asked her to read the story and she did quite well. She did better with the textual stimuli, but clearly had other sources of control over her responses.

Clearly textual stimuli could have only partially controlled the student's responses during the study. Most likely the student's verbal behavior during this study had intraverbal sources of control. Intraverbal behavior refers to verbal behavior controlled by a prior verbal response with no point-to-point correspondence between the stimulus and the response (Peterson, 1978; Skinner, 1957). Possibly the repeated practice of the same text weakens the control by textual stimuli and increases control by prior verbal stimuli (intraverbals).

Because research suggests that generally no transfer occurs from one type of verbal operant to another, investigating the type of control that occurs during repeated reading seems critical. It seems possible that the method of repeated reading may produce improvement primarily in intraverbal behavior and minimally in textual behavior or reading. This possibility deserves investigation. If repeated reading does not improve textual behavior as much as previous research suggests, that could have affected the
outcomes of this study. Specifically, if the fluency building procedures primarily improved intraverbal behavior, improvements in textual behavior (i.e., the frequency and accuracy checkouts) would not necessarily follow. Therefore, this dissertation cannot rule out the possibility that fluency building of textual behavior may improve student's movement through the Direct Instruction Reading Mastery curriculum.

Another source of evidence suggests distinguishing between textual behavior and intraverbal behavior may make a critical difference during reading instruction. Specifically, all students in this study engaged in "word guessing." When the students came to an unfamiliar word, they would fill in a word based on the context of the story. This almost certainly represents intraverbal behavior. The instruction provided during the dissertation did nothing to discourage this process, which seems likely to limit student's development of textual behavior or reading. Research suggests that students must unlearn this strategy to become effective readers (Stanovich & Stanovich, 1995). Other methods of building fluency seem more likely to produce direct improvements in textual behavior. Future research might use a different tactic to investigate the same research questions as this dissertation. For instance, researchers could teach all aspects of the lessons from Direct Instruction Reading Mastery until the students become accurate and then test if the students can pass the frequency and accuracy checkouts. Most will have the ability to pass these checkouts. Students who do not pass receive a fluency building component on the lessons (e.g., saying sounds, blending). If the students can pass the checkouts only after the additional fluency building component, then it becomes
clear some students academically at-risk need the additional instruction. It seems possible teaching to higher frequencies might not produce additional benefits. This dissertation could not answer that question.

Do high frequencies produce REAPs or predict REAP’s?

This research seems to confirm clinical data reported by Johnston and Lange (1992) that having a student reach a frequency aim only one time will not produce or predict retention, endurance, and application (i.e., REAPs) often reported in the literature. Researchers often debate whether high frequencies really produce REAPs or whether it is simply the number of practice opportunities. If achieving aim only one time does not produce REAPs, then achieving high frequencies alone does not produce or predict REAPs, and the more important variable might involve the number of practice opportunities.

Researchers do not always clearly distinguish between correlation and causation. Possibly high frequencies by themselves do not produce REAPs, but may assist in predicting when a student will achieve REAPs. It seems likely a better predictor of REAPs might involve the consistency at which a student can reach a high frequency. For example, a student who practices repeated oral reading until reaching a fluency aim of 180 to 220 per minute only one time probably will not demonstrate REAPs, as seen in
this study. If the student continues to practice for an additional week and reaches the instructional aim of 180 to 220 two times out of those five times she has received a lot of additional practice, but still might not demonstrate REAPs. They probably need to continue practicing until they can reach the instructional aim very reliably (e.g., they reach the instructional aim of 180 to 220 words per minute on the first try five out of five days.)

Retention

A confounding variable in the study of retention during this dissertation involved the length of time from the last test to the time the experimenter measured retention. Due to student absences and school holidays/days off, the experimenter could not match the number of days in the two conditions. Therefore, the experimenter chose to have the high frequency condition have the longer retention period. This dissertation did not show an effect for retention. From the data presented it cannot be determined whether this is due to no difference in the effects or due to the difference in retention periods.

A second problem with retention data collected in this study involves effects due to similarity. Because the stories the children read revolved around a common theme, they had many similarities. Research suggests that people will retain less information when learning similar material than material very distinct due to "interference" (Underwood, 1957). Future research on the effects of high and low frequencies on retention should match the time between conditions and avoid effects due to interference from similar material.
Endurance

One limitation of this experiment is that the data collection procedures had a built in bias against finding an effect for endurance. The duration of the timing period was determined by the frequency aim. If a story has 100 words in it and the frequency aim was 50 per minute, the students had 2 minutes to read the story. If the frequency aim was 200 per minute, the students had 30 seconds to read the story. After the students met aim in the low frequency condition, the experimenter listened to the tapes and scored the moment-to-moment changes in frequencies to study endurance. This session represented the best score the student had achieved up to this point in the low frequency condition. During the high frequency condition, after the student met aim on a story, she read it again at a longer duration. Then the experimenter compared this data with data already collected during the low frequency condition. Although the student had already reached aim in the high frequency condition, this data point could have been higher or lower depending on the student’s performance. In the low frequency condition, it was always the best score the student had achieved up to that point. Despite this bias, the study still found that the high frequency condition seems to improve endurance. All students had major improvements from their endurance in the low frequency condition even without any practice at the longer timing period. These data seem to replicate the findings of Nasir (1999), who also found that students who meet a high frequency aim one time, will have the ability to perform at the high frequency for longer durations with only some of the performance lost. The effect did not produce enough change to clearly demonstrate a
functional relationship. This dissertation could not avoid this bias while studying the effects of frequency on the frequency and accuracy checkouts. Future research might control for this by studying these topics in separate experiments and controlling for the confounding effects of the timing interval.

Agility

Although no published research exists, many unpublished dissertations have observed no relationship between frequency and celeration. (Ogden Lindsley, personal communication.) This dissertation replicated those findings. If no relationships exist between frequency and celeration no relationship can exist with agility (changes in the celerations over time). One interesting finding that future researcher should expand on and confirm involves the fact that students in this research seemed to show improvements in agility even at very low frequencies. This represents an important finding because if teachers can improve the celerations on new material students will move through the curriculum at a faster pace. In this research, students always experienced the low frequency condition first. Therefore, the students also worked extensively on accuracy during the low frequency condition. When the students entered the high frequency condition they already had high degrees of accuracy. One possibility that might explain large improvements in agility during low frequency and not during high frequency conditions involves improvement in the accuracy ratio. Improvements in celeration came primarily from improving accuracy during the low frequency condition. During the high frequency condition students had very minimal improvements in accuracy because they had already experienced the low frequency condition. It seems clear that accelerations
will differ if a student starts with material he or she has already learned to read accurately as opposed to material in which the student has not yet achieved accuracy. Future researchers should control for this problem and by comparing celerations only after the student has achieved a high degree of accuracy. A second possibly alternative would be to compare celerations in which the student needs to learn all material.

*Frequency and Accuracy Checkout Performance*

Although data reported clearly show that even after completing the high frequency condition students still failed most frequency and accuracy checkouts, students did achieve large gains in performance over baseline and the low frequency condition. The effects do not seem to have social significance because the students still could not keep up with the rest of their classmates even with the additional instruction. The fact that they showed major improvements after the high frequency condition does, however, suggest that intervention had some positive effects and that the students clearly showed generality to new material. The intervention used simply did not produce enough gains to allow the children to catch up with their peers in class. Future research might investigate many possible strategies to improve the results including interventions to improve attendance, increase the amount of instruction and more than one performance at aim.
Effects of combining Precision Teaching with Direct Instruction

It remains unclear whether the Precision Teaching components add any benefit to students as opposed to using Direct Instruction Reading Mastery alone. First, all of the students in this study had significant absences. They regularly missed reading instruction in class, which could easily limit the effectiveness of the Direct Instruction program. Second, the classroom had a first year teacher with limited training in Direct Instruction; this dissertation took no data on the quality of the implementation of the Direct Instruction program. The fact that the children failed frequency and accuracy checkouts 25 lessons or more behind where each child received instruction suggests significant problems with the implementation of the Direct Instruction. Finally, it seems questionable whether children in this study received instruction in the appropriate curriculum. Students in this study had reached 3rd grade and consistently fell behind their peers. It seems likely they might have done better if they had been placed in the "Corrective Reading" program specifically designed for remedial readers.

Functional Definition of Fluency

In Precision Teaching some authors have argued that fluency should have a functional definition (Binder, 1996). Precision teachers define fluency as the frequency that produces REAPs. If the particular performance standard produced REAPs, then we call that behavior fluent. If the performance standard did not produce REAPs, than we call that behavior not fluent.
Functional definitions have proved useful for other concepts in behavior analysis. For example, reinforcement has a functional definition. Reinforcement occurs when a stimulus is added or removed contingent on a behavior and increases the future probability of the occurrence of that behavior under similar stimulus conditions and establishing operations (Michael, 1993). If a stimulus is added or removed and it increases the future probability of that behavior, we call it reinforcement. If a stimulus is added or removed and it does not increase the behavior it does not meet the definition of reinforcement. One other component of the reinforcement definition that researchers do not explicitly state involves the fact that the adding or removing of the stimulus causes the increase in frequency and the stimulus change alone caused the change in frequency. This makes it a useful concept. The concept of reinforcement does not involve "theory." It simply describes an event that frequently occurs in the world. If a behavior analyst says, "The presentation of the reinforcer increased the behavior," he or she can demonstrate that experimentally. It does not represent a circular argument.

The functional definition of fluency, however, may represent a circular argument. First, it seems clear that while frequency may represent one component of producing REAPs, it does not represent the only component necessary to produce REAPs. Second, if increases in frequency do produce REAPs it remains very difficult to show experimentally. Most evidence comes from reports by teacher, clinicians etc. Very
limited experimental evidence demonstrates that high frequencies produce REAPs. It seems reasonable to conclude that a functional definition of fluency does not compare with other functional definitions in behavior analysis. This does not imply that fluency does not represent a useful concept, but simply that a functional definition may lead to research questions that will not assist in discovering more about this important area of behavior.

This dissertation used the functional definition of fluency in considering research questions. The research questions assume that frequency alone will make an important difference in retention. A better question might involve something to the effect of “What variables produce retention of reading passages?” The original question comes from having the functional definition of fluency as the model to work from. It forces thinking in terms of fluent or not fluent depending on whether a student achieves REAPs. The interesting questions involve retention, endurance, and application and what role frequency plays in their development.

Summary

This experiment studied the effects of academically at-risk student’s reading at high and low frequencies on their ability to pass the frequency and accuracy checkouts associated with Direct Instruction Reading Mastery in a public charter school. In addition, retention, endurance and agility of reading performances were evaluated. The results did not suggest any functional relationships. The data collected during this dissertation, however, may have implications for future research on repeated reading, fluency, Direct Instruction and Precision Teaching.
References


Grossen, B. (1997a). *What does it mean to be a research-based profession?* Eugene, OR: National Center for Improvement of the Tools of Educators, University of Oregon. Internet publication at [www.darkwing.uoregon.edu/~bgrossen/resprf.htm](http://www.darkwing.uoregon.edu/~bgrossen/resprf.htm)

Grossen, B. (1997b) *30 years of research: What we now know about how children learn to read*. Internet publication at [www.oregoneducation.org/reading/nichd.htm](http://www.oregoneducation.org/reading/nichd.htm)


Lindsley, O. R. (1997). *Do X2 then go for four*. Presentation at the 23rd Annual Convention of the Association for Behavior Analysis, May 24th, Chicago, IL.


Malott, R.W., & Harrison, H. (2002). *How to stick to your New Year’s resolutions or I’ll stop procrastinating when I get around to it*. Western Michigan University: Kalamazoo, MI.


Appendix A

Dear Parent:

My name is Barry Morgenstern. I am a doctoral candidate at The Ohio State University in the School of Physical Activity and Educational Services. I have written to invite your child to participate in my dissertation research. I selected students for possible participation based on recommendations from teachers of students who needed additional assistance in reading. Specifically, I asked your child’s teacher to recommend students who have had difficulty passing their tests as part of the Direct Instruction Reading Mastery program already used at your child’s school. I will provide individualized instruction to attempt to help students who have had difficulty passing the tests. The individualized instruction will use two different types of practice techniques. First, it will use repeated practice of reading material during short practice times. Second it will use a reading activity called “paired reading.” In this activity the instructor and the student take turns reading words in a story. The purpose of these activities is to assist your child in building fluent decoding (reading) skills and assist them in passing their mastery tests. Dr. Nancy A. Neef assistant professor in the College of Education will supervise the research. The school, CEO, principal and your child’s teacher have approved my proposal to complete my dissertation at your child’s school.
If you give your permission for your child to participate in this study he/she will receive individualized instruction 4 to 5 days a week for approximately 10 to 20 minutes each day for approximately 10 weeks. The individualized instruction will occur in addition to reading lessons he/she normally receives each day.

If you decide to give permission for your child to participate in this study, I will audiotape your child reading and store those tapes in a locked cabinet. After I have completed data analysis, I will destroy those tapes. Your child's name will not appear in any publication or documentation of this study and no data obtained will appear as part of any permanent record that could identify your child. You can withdraw your consent to participate at anytime and discontinue participation in this study without prejudice to your child.

If you give your permission for your child to participate in this study please sign the parent consent form attached to this letter and return it to your child's teacher. If you have any questions, please feel free to contact me at home at 781-9654 or by e-mail at BCMOR@aol.com. You could also direct questions about this study to Dr. Nancy A. Neef (The principal investigator), at The Ohio State University at 688-8107 or by e-mail at Neef.2@osu.edu.

Sincerely,

Barry D. Morgenstern  
Dr. Nancy A. Neef, Advisor  
Doctoral Candidate  
Assistant Professor
Appendix B

SCRIPT TO STUDENTS

Hello and nice to meet you. Please call me Barry. I go to school at The Ohio State University. I plan on teaching in your classroom. Your teacher said you might want to practice some reading. I came here today to find out if you would like to work with me.

If you do want to work with me, I would practice reading with you about 10-20 minutes a day. We would have fun trying to beat your best score and playing reading games.

If you would like to work with me just say “OK.” If you find that you don’t want to do it after awhile you can quit and you won’t get into any trouble with your teachers, your parents, or me. Do you have any questions? Thank you for talking with me today.
Appendix C

SCRIPT TO TEACHER

Hello and nice to meet you. Please call me Barry. I am a third year doctoral candidate at The Ohio State University in the College of Education. I plan to conduct a study at your school. I believe you may have some students in the Direct Instruction Reading Mastery program that have difficulty passing their mastery tests.

If so, I would like to provide some tutoring to those students on an individual basis for about 10-20 minutes a day. The study would last approximately 10 weeks. In addition, I will keep in close contact with you and share all of the accomplishments of each student.

I will accept students who consent to participate on a strictly voluntary basis. They can leave the study at anytime without penalties. I will keep all information confidential and no part of the data will become a part of the permanent record. The students’ must have failed at least 4 out of their last 5 mastery tests.

If you feel you have some student’s who would benefit and may want to participate, I will arrange to speak to the student’s and their parents. I will also provide permission slips to send home. Thank you for taking time to speak to me today. I look forward to hearing from you.
Appendix D

SCRIPT TO PRINCIPAL

Hello and nice to meet you. Please call me Barry. I am a third year doctoral candidate at The Ohio State University in the College of Education. I would like to conduct a study at your school on building oral reading fluency in children having difficulty in the Direct Instruction Reading Mastery program. I believe you may have some students in the Direct Instruction Reading Mastery program that have difficulty passing their mastery tests. I would like to know whether you think one of your teachers, and some of the students in the class would like to participate in this study.

If so, I would like to provide some tutoring to those students on an individual basis for about 10-20 minutes a day. The study would last approximately 10 weeks. In addition, I will keep in close contact with the child’s teacher and share all the accomplishments each student makes.

I will accept students who consent to participate on a strictly voluntary basis. They can leave the study at anytime without penalties. I will keep all information confidential and no part of the data will become a part of the permanent record. The students’ must have failed at least 4 out of their last 5 mastery tests.

If you feel you have some student’s who would benefit and may want to participate, I will arrange to speak to the teacher’s, student’s and their parents. I will also
provide permission slips to send home. Thank you for taking time to speak to me today.

I look forward to hearing from you.
Appendix E

Checklist to Assess Measurement Procedures

1. Ask the student to say their name on the tape for the first passage of the day.

2. Ask the student to say the date on the tape for the first passage of the day.

3. Ask the student to say the story number on the tape.

4. The experimenter gives the student the opportunity to preview the passage for one minute. "You have one-minute to look over the passage and then we will read it aloud."

5. The experimenter instructs the students to read as many words as possible and stop when they hear the count down timer.

6. Except during frequency and accuracy checkouts the experimenter reminds the student if they get to the end of the passage before the timer goes off to start again at the beginning of the story.

7. If the student stops at the end of the story the experimenter prompts the student to start again at the beginning of the story.

8. The experimenter provides praise for attending and participation, but no specific feedback on words read correctly or incorrectly.