

2007

Reading fluency through alternative text: Rereading with an interactive sing-to-read program embedded within middle school music classroom

Marie Biggs
University of South Florida

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

Scholar Commons Citation

Biggs, Marie, "Reading fluency through alternative text: Rereading with an interactive sing-to-read program embedded within middle school music classroom" (2007). *Graduate Theses and Dissertations*.
<http://scholarcommons.usf.edu/etd/634>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Reading Fluency Through Alternative Text: Rereading With an Interactive
Sing-to-Read Program Embedded Within a Middle School Music
Classroom

by

Marie Cecile Biggs

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Childhood Education
College of Education
University of South Florida

Major Professor: Susan Homan, Ph.D.
James King, Ed.D.
Robert Dedrick, Ph.D.
Linda Evans, Ph.D.

Date of Approval:
October 26, 2007

Keywords: singing, rereading, fluency, early adolescents, alternative text,
middle school, embedded literacy

© Copyright 2007, Marie Cecile Biggs

DEDICATION

I would like to dedicate this dissertation to my husband, Allen, for his constant love, patience, and support. I know that it would have been very difficult for me to complete this scholarly journey without your continued faith in me. Thank you, Allen, for being such a wonderful husband and companion. I look forward to a long wonderful life with you.

I would also like to dedicate this dissertation to my children, Nicole and Coleen and my little guy Jaiden. Nicole, thank you so much for your love and support. Your words of encouragement will stay with me forever. Coleen and Jaiden thank you both for reminding me of the importance of my family. I love you all so very much.

ACKNOWLEDGMENTS

I have many people to thank for their continued support throughout this scholarly journey. I am most grateful to my committee members, Susan Homan, James King, Robert Dedrick, and Linda Evans for their expert guidance on this project. Many thanks go to Dr. Homan for the opportunities that she has afforded me over the last few years. To Drs. King and Dedrick thank you for your insights and suggestions that helped to clarify this manuscript. I would also like to extend a warm thank you to Dr. Evans for demonstrating through her actions that research is about the children. Thank you all.

In addition, I would also like to thank Susan Bennett ABD, for her continued support and friendship. I don't know what I would have done without all your help and support over the last few years. I am so very grateful to have you in my life. Thank you .

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	viii
ABSTRACT	ix
CHAPTER ONE: INTRODUCTION.....	1
Statement of the Problem.....	11
Purpose of the Study	12
Research Questions.....	13
Significance of the Study	14
Limitations of the Study.....	16
Definition of Terms.....	17
Organization of the Manuscript	19
CHAPTER TWO: REVIEW OF LITERATURE.....	20
History of the Middle School.....	21
History of the Junior High	21
Evolution of the Middle School.....	24
Early Adolescent Learner Development.....	26
The Middle School Teacher.....	32
History of Content Reading	35
Instructional Practices.....	38
Skills and Strategies	40
Comprehension	41
Textbooks.....	42
The Learner and Content Reading.....	43
Summary	44
Current Complexities of Literacy Learning.....	47
The Early Adolescent Literacy Learner	47
School Structure.....	50
Mandates and High-Stakes	51
Summary.....	56
Recommended Effective Strategy Practice for Early Adolescent	58
Sociocultural Influences.....	58

Peer interactions.....	59
Effective Instructional Strategies	61
Direct Explicit Comprehension Instruction	62
Strategy Instruction Embedded in the Content	64
Diverse Texts	66
Technology	67
Motivation and Engagement.....	70
Fluency.....	72
Repeated Reading	75
Singing	78
Current Study	81
Summary	82
CHAPTER THREE: METHOD	84
Purpose of the Study and Research Questions.....	84
Research Design.....	86
Quantitative Phase	87
Qualitative Phase	89
Mixing the Methods.....	90
Research Context.....	91
School Site	91
Music Classroom	94
Participants.....	97
Sample Design	97
Sample Characteristics.....	101
Ethical Considerations	107
Instruments.....	108
Qualitative Reading Inventory-4.....	108
Validity and Reliability.....	108
Scoring.....	113
Florida Comprehensive Assessment Test.....	118
Validity and Reliability.....	118
Reliability of the Data.....	121
Internal Consistency.....	121
Interrater Reliability.....	122
Procedures.....	123
Pre-Study Involvement	123
How the Students Use the Tune Into Reading.....	127
Control Group	129
Data Collection	130
Quasi-Experimental Design Data Collection.....	130
Interpretive Case Study Design Data Collection	131
Data Analysis.....	135
Quasi-Experimental Design Data Analysis	135
Question 1	136

Question 2	138
Interpretive Case Study Design Data Analysis	139
Credibility of the Data	141
Second Observer	142
Analysis Check	142
Concurrent Triangulation Strategy	143
Integration of the Data	144
Summary	144
CHAPTER FOUR: FINDINGS	145
Question One: Findings for the Treatment and Control Groups	145
Results	146
Fluency (WPM) at the Same Reading Level	150
Word Recognition (WR) at the Same Reading Level	151
Comprehension (COMP) at the Same Reading Level	152
Highest Reading Level	154
Summary of Findings Question 1	155
Question Two: Findings for the FCAT Level Groups	157
Fluency (WPM) for FCAT Levels 1-5 on	158
Word Recognition (WR) for FCAT Levels 1-5	163
Reading Comprehension (COMP) for FCAT Levels 1-5	167
Highest Reading Level on WPM for FCAT Levels 1-5	171
Highest Reading Level on WR for FCAT Levels 1-5	173
Highest Reading Level on COMP for FCAT Levels 1-5	174
Highest Reading Level on RL for FCAT Levels 1-5	175
Summary of Findings for Question 2	176
Question Three: Qualitative Findings for Peer Interactions	178
Overview	179
Participants	179
The Role of the Researcher	180
Theoretical Considerations	180
Sociocultural Theory	180
Peer Interactions	181
Assessments	183
FCAT Reading Level Scores	184
Fluency: Absence of Prosodic Reading	187
Students Self-Reports: Reading Disposition	188
Treatment and Control Data Analysis	190
Integration of the Data	194
Treatment Group	194
Description of the Computer Lab	194
Data Analysis	195
Group Dynamics	196
Safe Risk-Free Environment	196
Peer Observation	197

Peer Hierarchy	199
Peer Support.....	200
Motivation.....	201
Extrinsic Motivation	201
Autonomy	202
Intrinsic Motivation	203
Singability vs. Readability.....	204
Students' Perspectives of Alternative Text.....	205
Disequilibrium	207
Summary of Results for the Treatment Group.....	208
Control Group.....	211
Description of Classroom Routine.....	211
Data Analysis.....	211
Engagement.....	213
Alternative Approach to Singing	213
Extrinsic Motivation	214
Group Formats	215
Dominate and Vulnerable Peers.....	215
Peer Leaders.....	216
Reading Strategies	218
Fake Rereading	218
Disengaged.....	219
Summary of Results for Control Group.....	220
Cross Case Analysis.....	222
Social Systems	222
Peer Positions.....	222
Instructional Expectations.....	223
Alternative Approaches	224
Reading Strategies	224
Summary of Cross Case Findings.....	225
Chapter Summary	226
CHAPTER FIVE: DISCUSSION.....	229
Summary of the Study	229
Discussion: Conclusions and Implications	236
Addressing Early Adolescents Differing Fluency Development.....	236
Assessing Fluent Readers	238
The Role of Prosody in Reading Fluency.....	243
Sociocultural Interactions	250
Contributions of the Study.....	252
Recommendations for Practice.....	254
Fluency Instructions.....	254
Assessing Fluent Middle School Readers	256
Sociocultural Interactions and Influences on Instruction.....	256
Recommendations for Future Research.....	258

REFERENCES	260
APPENDICES	282
Appendix A. Qualitative Reading Inventory-4	283
Appendix B. Matrix for Peer Interactions.....	286
Appendix C. Construct Key	287
ABOUT THE AUTHOR	End Page

LIST OF TABLES

Table 1. Middle School Timeline	46
Table 2. Ethnic Enrollment at School Site.....	92
Table 3. Percentages of Students' Classification Variables.....	103
Table 4. Group Percent of Student FCAT Level Scores.....	105
Table 5. Summary of Descriptive Statistics from QRI-4 Pretests	106
Table 6. Correlation of the QRI-4 Scores and Standardized Tests Scores	112
Table 7. Determining Instructional Reading Level from the QRI-4.....	117
Table 8. Schedule of Pretests for the Treatment and Control Groups	131
Table 9. Summary of Descriptive Statistics at the Same Reading Level.....	147
Table 10. ANOVA Table for Treatment and Control at the Same Reading Level.....	149
Table 11. Changes in Instructional Reading Levels	155
Table 12. Interactions and Effect Sizes.....	156
Table 13. Descriptive Statistics on Fluency (WPM) for FCAT Levels 1-5.....	158
Table 14. ANOVA Table on WPM for FCAT Levels 1-5	160
Table 15. Descriptive Statistics on Word Recognition (WR) for FCAT Levels 1-5.....	163
Table 16. ANOVA Table on WR for FCAT Levels 1-5.....	165
Table 17. Descriptive Statistics on Reading Comprehension for FCAT Level 1-5.....	167
Table 18. ANOVA Table on COMP for FCAT Levels 1-5.....	169
Table 19. Descriptive Statistics on WPM at the Highest Reading Level	172

Table 20. Descriptive Statistics on WR at an Highest Reading Level.....	173
Table 21. Descriptive Statistics on COMP at an Highest Reading Level	174
Table 22. Descriptive Statistics on RL at an Highest Reading Level.....	175
Table 23. Effect Sizes for WPM, WR, COMP, and RL by FCAT Level Groups	177
Table 24. Constructs from Treatment Group.....	195
Table 25. Themes from Treatment Group	196
Table 26. Constructs from Control Group.....	212
Table 27. Themes from Control Group.....	213

LIST OF FIGURES

Figure 1. Diagram of Mixed Methods Concurrent Design and Procedures	91
Figure 2. Floor Plan of the Music Classroom.....	97
Figure 3. Fry Readability Graph.....	126
Figure 4. Qualitative Observation Schedule.....	132
Figure 5. An Example of Matrix Used for Peer Interactions.....	135
Figure 6. Interaction Graph on Fluency (WPM) at the Same Reading Level.....	150
Figure 7. Interaction Graph on Comprehension (COMP)	153
Figure 8. Interaction Graph on WPM for FCAT level At in Reading	161
Figure 9. Interaction Graph on COMP by FCAT Levels.....	170

Reading Fluency Through Alternative Text: Rereading With an Interactive
Sing -to-Read Program Embedded Within a Middle School Music
Classroom

Marie Cecile Biggs

ABSTRACT

Singing exaggerates the language of reading. The students find their voices in the rhythm and bounce of language by using music as an alternative text. A concurrent mixed methods study was conducted to investigate the use of an interactive sing-to-read program *Tune Into Reading* (Electronic Learning Products, 2006) as an alternative text, embedded within a heterogeneous music classroom. Measured by the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006), the fluency, word recognition, comprehension, and instructional reading level of the treatment students were compared to their counterparts who sang as part of the regular music program. Concurrently, this investigation also provided a description of the peers' interactions during the literacy task assigned by the music teacher. The intent of this study was to address the following three research questions. First, what is the difference in reading outcomes for students who used the singing software verses the students who sang as part of their regular music curriculum? Second, are the reading outcomes different when the students were grouped by FCAT reading levels? Third, how do the peers interact during the literacy task of singing to read? The first two questions addressed the quantitative phase of this study to assess the collective differences on the dependent variables overtime and by group. The

qualitative phase in this study used an interpretive case study approach to describe peer interactions during the assigned literacy task.

The study findings suggest that rereading through singing, using the interactive singing program, *Tune Into Reading*, was more effective regardless of the reading levels for treatment students compared to control students. In addition, prosody appeared to have a direct connection to reading comprehension. Furthermore, the use of the interactive program provided opportunities for differentiated reading level achievement. Finally, group dynamics highly influenced the early adolescent's motivation, engagement, participation, and successful outcomes in reading fluency.

CHAPTER ONE: INTRODUCTION

Prominent in educational discourse is understanding and meeting the unique and differentiated needs of the early adolescent literacy learner. This is extremely important as these students prepare to meet the challenges of living in an informational age as fluent, active, and independent readers (Alvermann, 2001; Guthrie & Wigfield, 2000; Kamil, 2002). However, less than a decade ago, this population of learners suffered from scant attention to their literacy learning as “policy makers, curriculum developers, and school leaders rallied to address the literacy needs of students in grades K-3” (Elish-Piper & Tatum, 2006, p. 6). As a result, this placed the specialized literacy needs of the early adolescent at a disadvantage. The 2005 National Assessment of Educational Progress (NAEP) reading results have shown improvement in literacy achievement for the elementary level. These reading improvements however, have not necessarily translated to early adolescent literacy learners as once developmentally, cognitively, contextually, or instructionally assumed (Biancarosa & Snow, 2006). Specifically, the developmental stance that assumes instructional practice for this population of literacy learners all have made the cognitive shift from learning to read to reading to learn. However, as these students navigate their literacy learning across various content areas and through diverse and alternative texts, it should not be assumed they are fluent readers and comprehenders prepared to meet the challenges of the new millennium.

Early adolescence, typically defined as ages 10-14 (middle school years), is a time of transition and rapid change in the students’ emotional, social, physical, and cognitive

development (Cottle, 2001; Moje, Young, Readence, & Moore, 2000; Pikulski, 1991). Developmentalists, following the work of G. Stanley Hall (1908), continue the debate that early adolescents are neither children nor fully mature adolescents. Instead, they are caught in the developmental tensions of adolescence (Bean & Brodhagen, 1996). These tensions, which parallel the onset of this developmental stage, can become even more daunting when the early adolescent student enters the contextual environment of the middle school. At this level, more cognitive strategic demands in reading are placed upon the students to comprehend diverse texts (Alvermann & Phelps, 2005).

The complex process of comprehending text is the ultimate goal of reading. Alexander (1998) believes this is extremely difficult for early adolescents because their cognitive strategic processes in reading are very diverse and are under continual development. Even though early adolescents are situated within a particular developmental stage, their cognitive abilities in reading vary with the different literacy tasks presented. Jetton and Alexander (2000) suggest, early adolescent readers' use of text comprehension strategies range across a developmental continuum, and there is interplay of prior knowledge, experience, and strategic processes. Therefore, an adolescent reader may be a competent fluent reader in one literacy task and yet fall back and need support in another task. Ivey (1999), in her case study of three sixth grade students of varying reading abilities, found that middle school readers were complex and multidimensional in their reading. These complexities may become more pronounced as the middle school reader enters the context of middle school. This may affect their ability to read fluently within and across various content areas.

Content area teachers assume it is their responsibility to cover their subject matter in a timely, accurate, and effective manner (Alvermann & Moore, 1991). The cognitive shift from learning to read to reading to learn is assumed to occur before students leave elementary school. Dole, Duffy, Roehler, and Pearson (1991) found, this assumption supported through the pedagogical lens of the middle school content area teacher. Middle school content area teachers often incorrectly believe that by the time most of their students enter their classrooms they are fluent readers. Therefore, they may believe incorporating strategic approaches towards fluency in reading are not needed for this population of learners.

To further complicate this contextual dilemma, middle school content teachers have resisted the recommendation to incorporate literacy-related instruction into their curricula (Phelps, 2005). Biancarosa and Snow (2006) suggest that content area teachers should be encouraged to provide literacy skills and strategies that are embedded in their content area. By emphasizing the literacy practices that are specific to their subject area, they can maintain the integrity of the content while providing strategic literacy instruction to comprehend and to be fluent with the specific concepts being taught (Alfassi, 2004). However, Bulgren, Schumaker, Deshler, Lenz, and Marquis (2002) report, content teachers feel they do not have the time or experience to include explicit literacy instruction into an already crowded curricula. This may be a result of deeply embedded values, beliefs, and practices, and the need to conform to stringent standards imposed by the No Child Left Behind Act of 2002 (O'Brien, Stewart, & Moje, 1995).

The problem looms even greater in this era of standards-based reform - one that calls upon educators to meet these standards, to teach to these standards, and to have

these standards evaluated through annual high-stakes testing (French, 2003). The results of this yearly assessment can have a dramatic impact on the early adolescent literacy learner with the possibility of retention, class placement, and specifically, instructional practices provided to the students. Rothstein (2000) questions whether an annual test of a student's knowledge, at just one point in time, can provide an accurate assessment of this population of literacy learners. The score obtained from this high-stakes test place the early adolescents below, at, or above their classmates in reading, and it is assumed that the early adolescent students who may or may not have passed the test will receive the instructional strategies needed to prepare them to be fluent readers and comprehenders.

This narrow focus places the literacy needs of this population of learners at a disadvantage as they prepare to become productive citizens in our larger world (French, 2003; Sackes, 2000). Currently, this cognitive stance integrates developmental and contextual considerations and is supplemented with an appreciation of the socialcultural influences that shape instructional practices for these literacy learners (Phelps, 2005). Specifically, the social interactions (e.g. talk, peer modeling, or social reinforcement) of the early adolescent peer groups, that blends their diverse backgrounds and experiences during the literacy task (rereading through singing), occurring within the cultural environment of the classroom.

During early adolescence, the peer group becomes a prominent context for development (Brown, 1990).The school and classroom provides opportunities for peers to interact throughout the day. Ryan (2000) reports “peer interactions consume significantly more time in adolescence compared to childhood” (p. 107). These interactions with peers can concern both academic (e.g., achievement) and nonacademic matters (e.g.,

engagement, motivation, self-efficacy, and interest). Ryan (2000) suggests, there are generally three ways that early adolescents experience peer interactions within the context of middle school: through information exchange, modeling, and peer pressure.

Information exchange occurs when adolescents have a discussion with their peers (Berndt, 1999). In an experimental study with eighth-grade students, Berndt, Laychak, and Park (1990) found that when adolescents had to make an academic decision such as attend a rock concert or study for a test, they initially responded differently from one another. However, after discussing this dilemma with their peers, their answers were similar to their peers. This form of interaction could influence the early adolescent's choice to partake in the literacy task presented by the teacher if it was used effectively.

Modeling is another form of adolescent peer interaction. This interaction refers to individual changes in cognition, beliefs, or affect, which are a result of adolescents observing their peers (Ryan, 2000). Observing a specific behavior performed, or listening to a peer voice a certain belief, can induce an adolescent to adopt such behaviors or beliefs. Schunk and Zimmerman (1996) reported, peer modeling influences self-efficacy beliefs. In their study, they found that early adolescents who verbalized that they had difficulty with a task and then observed their peers have success with the same task then believed they could complete the task. The early adolescent, when faced with a literacy task, may have success by observing their peers. Peer pressure is a third way that the early adolescent interacts with their peers.

Peer pressure takes on the role of social reinforcement (Ryan, 2000). Brown, Lohr, and Eicher (1986) found that beliefs and behaviors that are discouraged by the groups are not likely to be displayed, whereas beliefs and behaviors that are positively

received by the group are more likely to surface. Therefore, participation in the literacy tasks that the peer group positively received through this social interaction, could have a positive effect on the group's beliefs and decisions to participate by the group members.

The field of reading has moved far beyond the view that literacy is the ability to read and write across various content areas alone (Bean, 2000). Instead, the concept of content reading has been broadened to reflect the integration of communication processes (reading, writing, talking, listening, and viewing) as the students engage in text-related learning (Alfassi, 2004; Lenz & Deshler, 2003; Vacca & Vacca, 2002). There is the assumption that the linear textbook is necessary for teaching and learning the content specifics (Wade & Moje, 2000). It is this assumption that influences instructional delivery and perceptions of fluent, active, and independent readers (Alvermann, 2002). However, Phelps (2005) reports, alternative texts that focus on new literacies through digital media have had a great influence on the early adolescent's instructional practices.

The computer offers students more control in terms of support, pace, and active processing of text (Kamil, 2002). The use of technology as an alternative text links real world experiences and interests, and provides opportunities for alternative text reading with the early adolescent literacy learner. Leu (2000) reports on the positive effects for middle school readers when print and visual texts (e.g., hypermedia, the internet, and interactive CD-ROMS) are utilized. Reading diverse texts across and within various content areas can be further complicated if early adolescent students do not have the background knowledge, experiences, and strategies for reading a variety of texts fluently.

Fluency is a necessary aspect of successful reading as it allows readers to read with speed, accuracy, and proper expression (National Reading Panel, 2000; Rasinski,

2004). For years, teachers thought, if students could learn to decode words accurately, they would be successful in reading printed text. The assumption is often made that early adolescents are at a satisfactory level of fluency in reading. However, according to Alvermann and Phelps (2005), this is not always the case, specifically with content area materials. While it is true that accuracy in students' ability to decode words is important for fluency, decoding needs to be automatic. However, automatic decoding for fluent reading is not sufficient. Rasinski (2004) points out the need to connect accuracy and automaticity to reading prosody.

Reading prosody is the point where fluency connects fluent decoding directly to comprehension (Rasinski, 2004). The prosody components of reading fluency address the use of phrasing and expression (Dowhower, 1987, 1991; Schreiber, 1980, 1987, 1991; Schreiber & Read, 1980). When readers adjust appropriate volume, tone, emphasis, phrasing, and other elements while reading aloud, they are providing evidence of comprehending text (Rasinski, 2004). In this sense, fluency is a multifaceted event with reading comprehension as the goal.

Through guided and repeated reading, both prosody and decoding (automaticity and accuracy) in word recognition are developed. Samuels (1979) defines repeated reading as a fluency-building strategy that consists of timed rereading of a short passage several times (at least 3 times), checking for accuracy (word recognition), automaticity (words per minute) and with prosody (expression). The steps for an effective fluency instructional model are: (a) provide a model for students expressive fluent reading, (b) give the students a passage to read (approximately 150 words) 3 times at their instructional reading level (word recognition with 90-95 % accuracy), and (c) have the

students orally read the passage assessing for accuracy , automaticity, and expression (Rasinski, 2004).

The National Reading Panel (2000) found sufficient evidence that guided oral reading done through repeated reading will have a positive impact on fluency and comprehension across a range of grades levels and in a variety of general and special education classrooms. Rasinski (2004) contends that reading fluency is a “bridge between two major components of reading- wording decoding and comprehension. At one end of the bridge, fluency connects to accuracy and automaticity in decoding. At the other end, fluency connects to comprehension through prosody, or expressive interpretation” (p. 1).

Repeated reading is most authentic when the practiced material is eventually performed orally, such as plays, poetry recitation, or in this study singing lyrics to songs (Rasinki, 2004; Stayter & Allington, 1991). This form of repeated exposure through singing assists the reader with fluency through prosodic reading. The reader uses appropriate volume, rhythm, pitch, tone and phrasing (prosody), while singing the song lyrics, and therefore, they giving evidence of actively constructing meaning from the passage (Rasinki, 2004). Singing as an alternative text can build reading fluency and comprehension and can be naturally embedded within the music content classroom.

Butzlaff (2000) contends that there are similar characteristics with singing instruction and the reading process: (a) music text and written text involve formal written notations that are read left to right, (b) the sensitivity to phonological distinctions and word recognition requires a sensitivity to pitch and tonal distinctions in both reading and singing, (c) when students learn the lyrics to songs they are engaging in reading, and (d) learning song lyrics is often repetitive, so that rereading of text occurs through singing.

Hall, Boone, Grashel, and Watkins (1997) suggest that students should sing independently, on pitch, and with rhythm. While most singing in the music classroom is done in groups, minimal time is spent with students singing individually, making it difficult to assist each student to develop these specific faculties. Along with singing independently, Levinowitz (1989) suggests that students would sing songs more accurately with copies of individual text than without. However, singing in the music classroom is usually performed as a whole group with one song and one group text. Usually the text is displayed on an overhead or chart, regardless of the variety of instructional reading levels of the student body. Currently, the use of an individual computer program could address these concerns.

Individualized computer assisted training in the music classroom is a recent additional tool teachers can employ for students to learn to sing and acquire songs individually. In a study analyzing 150 empirical articles on computer applications in music learning, Webster (2002) reported generally positive results with singing performance and pitch accuracy; however, studies on song acquisition with software for students in the middle school setting are sparse, especially studies relating singing to reading. One report on the computer program *Carry-A-Tune* (Educational Learning Products, 2004) is in publication to date. This was a pilot study to examine the use of the sing-to-read software program with remedial reading middle school students (Biggs, Homan, Dedrick, Minick, & Rasinski, in press).

Carry-A-Tune is an individual computer product, originally developed to improve singing. The program uses a vocal range analyzer that tracks the singer's pitch and rhythm, comparing it to the correct pitch of the song. Each student uses a microphoned

headset linked to the computer to sing along repeatedly and to record their singing. There is a great need to investigate the effects of this and other computer singing programs, especially if the potential exists that they could be as helpful to music teachers as it seems to be for reading instructors and their students. The current study investigated the use of an individualized interactive sing-to-read program *Tune Into Reading* (TIR) (Electronic Learning Products, 2006), adapted from *Carry-A-Tune*, as an alternative text embedded in the middle school music classroom curriculum.

Tune Into Reading, not unlike its predecessor *Carry-A-Tune*, has several unique features that can be used to meet the specialized needs of this population of literacy learners. In both programs each student uses an individual soundproof microphoned headset for listening, singing, and recording. This provides real time pitch recognition and feedback to the user. The inclusion of pitch recognition is important because Lamb and Gregory (1993) found that pitch discrimination is significantly correlated (.77) with reading ability. In *Tune Into Reading* as was the case with *Carry-A-Tune*, the scoring mechanism (pitch accuracy scores 0-100) accommodates each individual's vocal range, and contains a portfolio sign-in menu that aligns with the custom vocal range of each participant. However, *Tune Into Reading* generates reports that print pitch scores for the individual student and/or the class, whereas *Carry-A-Tune* did not. In addition unlike *Carry-A-Tune*, *Tune Into Reading* provides individual folders for each participant. As soon as the participant signs into the program and clicks on the My Lesson folder they have access to the songs that are at their instructional reading level. Also, while both programs had songs analyzed for readability levels, *Tune Into Reading* has over 200 hundred songs, whereas *Carry-A-Tune* had only 24 songs. The songs range from first to

tenth grade readability levels. This wide range of available reading levels will provide opportunities for the students to build fluency through repeated reading by singing songs at their individual instructional reading level.

The literature on reading fluency often focuses on the beginning reader's initial stage of literacy acquisition or on the older adolescent reader who has difficulty learning to read. This focus has placed reading fluency in a deficit view, rather than creating a direct link to comprehension (Clay, 1985). Stayter and Allington (1991) suggest that "we have failed to consider some of the broader ramifications of an emphasis on fluency, especially with older and more developed readers" (pp.143-144). Especially when fluency instruction could support both the struggling and more developed reader's, as they transition to the context of middle school, navigating their literacy learning, across various content areas and though diverse and sometimes difficult texts.

Statement of the Problem

This study examined how the use of sing-to-read program *Tune Into Reading* as an alternative text might support literacy learning of early adolescents and thereby, improve their fluency(word per minute), word recognition (accuracy in oral reading), comprehension (implicit and explicit questions after reading), and instructional reading level (combined scores of accuracy and comprehension).

A majority of early adolescents need opportunities and instructional support to read varied and diverse materials in order to build their experiences, fluency, and range as readers (Kuhn & Stahl, 2003). Literacy learning should take into account developmental issues, as well as thoughtful and critical literacy expressions that embrace the multiple literacies that these students bring to school within and across various content areas

(Kamil, 2000). The problem is that little is known about this population of literacy learners and about how to provide literacy instruction that will address this change while, at the same time, providing support for their social and academic needs (Alvermann & Phelps, 2005).

In order to gain a perspective on the impact that these assumptions have on middle school readers, it is appropriate to examine these students within a music classroom, to investigate singing as a form of repeated reading to improve fluency. This study investigated a population of middle school students who are in a music classroom as part of their assigned yearly elective cycle. Examining this sample will provide better insights into the area lacking in the available literature – the possibility of providing effective literacy instruction through alternative text embedded in music content area instruction.

Purpose of the Study

The purpose of this study was to investigate the use of an interactive sing-to-read program *Tune Into Reading* as an alternative text, embedded within a heterogeneous music classroom. This investigation also provided a description of the peers' interactions during the literacy task assigned by the music teacher. This study used a concurrent mixed methods design. The intent of the study was to address the following research questions:

Quantitative Research Questions

1. To what extent is the reading performance of word recognition, fluency, comprehension, and instructional reading level, as measured by the QRI-4, of

students using the *Tune Into Reading* program different from their regular music curriculum counterparts?

2. To what extent does the *Tune Into Reading* program differently impact the reading scores of students who are “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) reading scores?

Qualitative Reading Question

1. How do middle school readers interact with their peers, within the context of their music classroom?

The first quantitative research question addressed the readers’ use of the interactive sing-to-read program *Tune Into Reading* as an alternative text. Prior to the treatment, I administered a pretest using the QRI-4. Scores from the pretest were used to ensure that the students in the control and experimental groups were not different in their performance in fluency (measured by words per minute), word recognition (measured by oral reading accuracy), comprehension (measured by implicit and explicit questions after the reading), and instructional reading level (measured by combined scores of accuracy and comprehension) before implementation. After the implementation of the interactive sing- to- read program, *Tune Into Reading*, I administered a posttest using the QRI-4 and compared the posttest scores with the pretest scores to determine if students in the experimental group had gained significantly over their counterparts in the control group.

The second quantitative research question investigated whether there is an interaction effect of the repeated reading methods using the sing-to-read program, *Tune Into Reading*, as an alternative text on the reading performance of the students when they were grouped as “below, at, or above” grade level as determined by the Florida

Comprehensive Assessment Test (FCAT) 2006 in reading. The results in reading achievement level scores (achievement levels 1-5), according to the state of Florida Department of Education, are reported as: (a) students who scored a Level 1 or 2, are considered below proficiency in meeting grade level benchmarks, (b) students scoring a Level 3 are considered at grade level, and (c) students who scored at a Level 4 or 5 are considered above grade level (FCAT Briefing Book, 2005).

Concurrently, the qualitative observations were used to probe for significant themes by describing aspects of peer interactions (e.g., peer talk, peer modeling, and peer social reinforcement) among students who are singing using the interactive program *Tune Into Reading*, versus the peer interactions among students who are singing in the traditional music class.

Significance of the Study

Currently, although there appear to be emerging themes and important information being investigated about the contextual conditions, developmental needs and instructional practices, concerning reading in the content areas, the knowledge base for early adolescent literacy learners is still limited (Alvermann, 2002; Bean, 2000; Kamil, 2002; Moore, 1996). The National Reading Panel (2000) identified fluency as one of the five critical components of reading (Pikulski & Chard, 2005). Fluency in reading, however, is often thought of as a deficit, remedial tool for word accuracy and automaticity, rather than a direct link to comprehension (Clay, 1985; Stayter & Allington, 1991; Rasinki, 2004). Repeated reading is the methodology that is most appropriate to develop fluency and comprehension so that early adolescents can navigate their literacy learning strategically across various content areas. However, little is known about

repeated reading to build fluency with early adolescent literacy learners of varying reading abilities. Even though the assumption is often made that many early adolescents are at a satisfactory level of fluency in reading, this is not always the case, and specifically, it is not the case with content area materials (Alvermann & Phelps, 2005). The standards based reform movement with high-stakes testing has also contributed to the assumption that a middle school student is a fluent reader. The current study will add to the knowledge base important information pertaining to fluency instruction through repeated reading for a range of literacy learners.

Along with this cognitive stance and its overlap with the development stage and contextual conditions is an appreciation of the socialcultural influences that shape instructional practices for this population of literacy learner (Phelps, 2005). Specifically, the social interactions (e.g., talk, modeling, and social reinforcement) of the peer group, blending each member's diverse background and experiences during the literacy task (repeated reading through singing), and occur within the cultural environment of the classroom. Through observations and descriptions of peer interactions, more information will be provided to the field concerning these interactions during specific the literacy task presented.

In 2004, to help address this population of literacy learners, a panel of five nationally known educational researchers met with representatives of the Carnegie Corporation of New York and the Alliance for Excellent Education. The focus was to draw up a set of recommendations on how to meet the needs of adolescent literacy learners while propelling the field forward (Biancarosa & Snow, 2006). A list of 15

elements were reported and then divided into two sections: instructional improvements and infrastructure improvements.

The instructional elements consisted of: a) direct, explicit comprehension instruction, b) effective literacy instruction embedded in content, c) motivation and self-directed learning, d) text based collaborative learning, e) strategic tutoring, f) diverse texts, g) intensive writing, h) technology, and i) formative on-going teacher assessments (Biancarosa & Snow, 2006). Researchers were urged to re-conceptualize how they perform research with early adolescent literacy learners. Investigations should combine different elements so that important information about the early adolescent can be determined. This study utilized five of these elements. It investigated early adolescent literacy gains when instruction is embedded in the music content area. Also, the interactive sing-to-read program *Tune Into Reading* is delivered through a technological format with a diverse and interesting text, which may be motivating and engaging (Guthrie & Wigfield 2002). Most important, explicit comprehension instruction through rereading to enhance comprehension was addressed.

Limitations of the Study

The following list is provided to acknowledge and clarify the limitations of this study that impact the generalizability of the findings:

1. Random sampling of individual students was not an option in this study, and therefore possessed a threat to the external validity. This limited the generalizability of the findings. To address this threat, random assignment by classes were made. In addition, analysis was conducted to match sample

characteristics including demographics, and reading performance prior to analysis and treatment.

2. Complete FCAT level scores in reading were unavailable for all participants (four students were missing scores from the treatment group and four from the control. The researcher acknowledges that missing data might have limited the findings for question two. However, there was an equal distribution of percentages in each group stratified as *Below* grade level (FCAT level 1 and 2), *At* grade level, (FCAT level 3), and *Above* grade level (FCAT level 4 and 5).
3. The study duration was only seven weeks, had it been longer it might of netted different results.
4. The characteristics of the samples were predominantly low SES White males in eighth grade. This limits the findings for other sample characteristics.

Definition of Terms

The following is a list of the terms and operational definitions that will be used throughout the study:

1. *Active Reader*: Readers who engage in an active search for meaning using multiple strategies as they monitor their understanding of what they have read (Pearson & Fielding, 1991).
2. *Alternative Texts*: Various textual formats that are used to supplement the linear text or replace the textbook in the content areas. Most often they are digital in nature. In this study the alternative texts refer to the interactive sing to read program (Alvermann & Phelps, 2005).

3. *Early Adolescent Literacy Learner*: typically defined as ages 10-14 years (middle school), is a time of transition and rapid change in the students' emotional, social, physical, and cognitive development (Cottle, 2001; Moje, Young, Readence, & Moore, 2000; Pikulski, 1991).
4. *Embedded Literacy in the Content*: Literacy embedded in the content addresses two directions for instructional implementation (Biancarosa & Snow, 2006). First, within the Language Arts classroom these principles are not discrete skills or techniques instead the emphasis should be how to teach the strategy or skill using other content-area materials. Second, content area teachers should encourage literacy skills and strategies that emphasize the reading and writing practices that are specific to their subject area (Alfassi, 2004).
5. *Fluent Reader*: A reader who reads with accuracy, automatic recall, and voice expression, volume and pitch (Rasinski, 2004)
6. *Independent Reader*: A reader who requires less in the way of structured learning support (Pearson & Gallagher, 1983).
7. *Literacy Tasks*: Assigned task related to reading and writing given to the students by the teacher.
8. *Sociocultural Influences in Literacy*: A sociocultural approach to literacy instruction is multidisciplinary and occupies the fields of history, anthropology, linguistics, psychology, and sociology. Sociocultural approaches emphasize the interdependence of social and individual processes in the construction of knowledge. When viewing literacy development from a sociocultural approach, literacy arises from the child's participation in social activities in which there are real reasons to use written language

(Englert & Palinscar, 1991). In this study the social interactions (e.g., talk, modeling, and social-reinforcement) of the peer group, blending diverse background and experiences during the literacy task (rereading through singing), occurring within cultural environment of the classroom.

Organization of the Manuscript

This manuscript has been organized into five chapters. Chapter One identified the problem and places it in the context for the study. The research questions, limitations, and definitions are also included. Chapter Two reviewed the literature relevant to the research questions. Research strands include (a) Historical Review of the Middle School Movement: The Context, the Learner, and Reading Instruction (b) Current Contextual Conditions: Influence of Standards and Mandates with Literacy Development (c) Effective Practice and Instructional Delivery for the Early Adolescent Literacy Learner. Chapter Three presented the methods that were used to conduct this study. It outlined the research questions, research context, and the participants. In addition it described the design of the study; including ethical considerations, instruments, and procedures. The final sections explained reliability measures and the manner in which the data was collected, analyzed, and interpreted. Chapter Four summarized the findings of the study. The descriptive statistics and findings derived from the data analysis are reported. Chapter Five presented the conclusions of the study, the resulting implications of the study results, and the recommendations for classroom and future research.

CHAPTER TWO: LITERATURE REVIEW

The National Assessment of Educational Progress (NAEP) (2005) reports over 73 % of eighth grade students perform below or at a basic level in their reading achievement. Consistent with NAEP results Biancarosa and Snow (2006), in their report to the Carnegie Corporation, contend that over 70% of adolescents struggle with their reading in some manner and therefore require instruction that is differentiated and strategic. This is alarming as few gain the literacy knowledge needed to successfully engage in higher-level problem solving required for an information transforming economy (Donahue, Voelkl, Campbell, & Mazzeo, 1999). In addition, although emerging themes appear and important information in connection to reading in the content classrooms, while addressing the contextual conditions, developmental needs, and instructional practices, the knowledge base for early adolescent literacy learners is limited (Alvermann, 2002; Bean, 2000; Kamil, 2002; Moore, 1996).

How do we prepare our early adolescents to be fluent, active, and independent readers, who meet the literacy demands and challenges of living in an informational age? Although this issue poses current complexities for adolescent literacy learners, the dilemma of how best to meet the unique needs of the early adolescent, typically defined as ages 10-14 (middle school years), has been a historical debate for over 100 years. In order to understand the gaps in the literature and how best to currently meet the literacy needs of the early adolescent, I will provide a brief review of the historical background.

This historical context will be helpful to inform current practices. Therefore, this literature review will chronologically address the history of the middle school movement, its overlap with the unique developmental needs of its learners and teachers, and the parallels historically with reading in the content areas. This will be followed by a review of the complexity of current practices of content reading embedded in the middle school content areas and the influences of mandates and standards. The final section of this review examines studies and thoughts about effective strategic practices to meet the needs of the higher literacy demands for the future.

Brief Historical Review of the U.S. Middle School Movement:
The Context, the Learner, and the Parallel of Content Area Reading Instruction

To understand the challenges in today's middle school, teaching and instruction can not be separated from the social and institutional context in which it occurs. To gain a perspective about the context it is important to understand its history (Brodhagen & Bean, 1996). This historical lens allows us to gain an understanding of the current contextual, developmental, and instructional conditions afforded to this population of literacy learners, and suggest how the field should move to address their needs.

The History of the Junior High School

The prominent configuration of education in the 1900's consisted of eight years of primary school and four years of secondary school. Instructional focus for the early adolescent (grades seven and eight) consisted of a review of the first six years of schooling (Brimm, 1969). There were claims that the early adolescent's time was wasted in school with this narrow focus, which resulted in political and societal pressures to reconfigure the elementary schools (Cuban, 1992).

Educational researchers (Beane, 2001; Brough, 1995; Cuban, 1992; Spring, 1986; Van Til, Vars, & Lounsbury, 1961) agreed that the suggested reconfiguration of the schools came from societal, political, and academic pressures: (a) influx of immigrations and burdened enrollment at the elementary level, (b) industrialized period, (training a workforce) and high drop-out rate by 8th grade-which resulted in a workforce of unskilled workers, and (c) preparation for the academic rigor of high school and college.

Along with the societal and political issues that impacted the reconfiguration of the elementary schools, there was also a developmental movement taking hold. The National Education Association (NEA) was one of the groups taking a developmental stance for school reconfiguration. The NEA (1899) argued for a reconfiguration of the elementary schools and a need to start secondary school at 7th grade rather than 9th grade. In their position statement they argued:

[T]he transition from elementary to the secondary period may be natural and easy by changing gradually from the one-teacher regimen to the system of special teachers, thus avoiding the violent shock now commonly felt on entering high school. (p.10)

This reform effort was led by an NEA committee member Charles Elliot, then president of Harvard College. In his position statement he argued that a better college preparation could be achieved for the early adolescents by extending the secondary school programs downward (Brimm, 1969). However, in 1917 the Smith Hughes Act spearheaded by the National Society for the Promotion of Industrial Education, proposed curricula programs that focused on improving the workforce of skilled laborers, specifically agriculture (Brimm, 1969). This two-track system met the societal and

political pressures to train a workforce, while providing rigorous academics for the college bound students earlier, and to ease overcrowding conditions at the elementary level.

The developmental position taken by the NEA was consistent with the work of the influential psychologist G. Stanley Hall (1908), who argued for years that early adolescents were in a unique stage of development and they should be separated in the context of their schooling from their predecessors and successors. Hall (1908) contended that early adolescents, if placed in the elementary school, would have a negative influence on the younger children, and if placed in the secondary school, would be negatively influenced by older adolescents (Beane & Brodhagen, 1996). It was therefore recommended by The Committee on the Economy of Time and the Commission on the Reorganization of Secondary Education (1918) that the reorganization of schooling for the adolescent be divided into junior and senior high levels for secondary school (Juvonen, Nhuan Le, Kaganoff, Augustine, & Constant, 2004).

Although NEA developmental position contributed to the new configuration of the junior high school, educational historians report that the motivation for this new institutionalized structure was created for multiple purposes. Beane (2001) and Cuban (1992) contend societal and political pressures had the strongest influence on the reorganization of the junior high school as a result of the converging interests of humanists, societal efficiency advocates, and stage-related developmentalists.

Specifically, the issues were related to overcrowding of the elementary school and tracking of students for the vocational path (workforce) or academic path

(college). Lounsbury (1984) contends this period from 1890-1920 was a struggle between academics and vocations. This tracking path translated to the instructional practices and curriculum delivery afforded to the early adolescent. The students directed towards the vocational path received a very different instructional program of survey academic courses and life skills, as compared to their counterparts on the academic path who were afforded coursework with academic rigor (Beane & Brodhagen, 1996). However, in spite of the needs of rapidly change society and the premise that the junior high school would help facilitate this change, only one-third of the students made it to 9th grade, from the early 1900's to the late 1950's (Van Til, et al., 1961).

The Evolution of the Middle School

Even though the junior high school reconfiguration was not a success (e.g., due to the large numbers of students dropping out of school), enrollment at the elementary level continued to increase. Therefore more junior high schools were built, specifically for space purposes (Alexander & George, 1981). The 1950's brought about discussion not only pertaining to the uniqueness of the students but also how the instructional programs for this population should be matched to their needs. In their analysis of the literature on instructional practice for the early adolescent, Gruhn and Douglas (1956) synthesized the following goals for the junior high school:

- integration of skills, interests, and attitudes
- exploration of interests and abilities
- differentiation of educational opportunities based on student background, interest, and aptitudes

- socialization experiences that promote adjustment, guidance in decision making
- articulation that assists youths in making the transition from an educational program designed for preadolescents to a program designed for adolescents.

(p.12)

However, after many theoretical discussions about the unique needs and instructional programs that should be developed for the early adolescent at the junior high level, the translation of theory to organizational and instructional practices was very similar to that of the senior high school in the 1960's. Bough's (1995) research reports, that there was "an emphasis on content rather than exploration, departmentalized rather than integration, and adherence to a rigid schedule" (p. 38). The junior high's curriculum and organization assumed similar characteristics as the senior high school. Brimm (1969) contends "The very name, "junior high school," was pointed to as a serious obstacle in the development of a special program for the early adolescent" (p. 8). These challenges created obstacles for the reformers to meet their goals of: (a) schooling that addresses the unique developmental needs for the early adolescent students, and (b) preparation for their future, whether it would be work or college.

A growing concern and dissatisfaction existed during this time period for the contextual conditions afforded to the early adolescent. While the secondary school enrollment dropped, the elementary level of school expanded. The 1960's brought another wave of political talk to change the junior high to middle school (Cuban, 1962). The goal was to match instructional practices to meet the needs of these young learners.

However, Alexander (1968), in his survey research, with a stratified random sample of 110 reorganized middle school principals, found that 58 % of the respondents reported that middle schools were developed to eliminate overcrowding of the elementary school, while 42 % said programs were needed to meet the developmental needs of the early adolescent. Ten years later Brooks and Edwards (1978) conducted a replication of Alexander's study and found that 42 % of the principals suggested the same reason to eliminate overcrowding, whereas 58 % reported to have a program designed to meet the developmental needs of the early adolescent. Cuban (1992) contends it is evident "that the mix of stated motives echoes the variety of reasons given by promoters of junior high schools at the turn of the century" (p. 243). Specifically, the reconfiguration of the school context because of over crowding conditions at the elementary level, and developmental needs of the early adolescent learner.

The Early Adolescent Learners Developmental Needs

Research on the developmental characteristics of the early adolescent learner was crucial to the reconfiguration of the junior high and later to the middle school. Hall's (1908) work in his book, *Adolescence: Its Psychology and Its Relations to Physiology, Anthropology, Sociology, Sex, Crime, Religion and Education*, portrayed early adolescence as a period of turmoil and stress. He contends this period is a result of the biological and psychological changes that occur. He argued schooling for these students should be separated from their predecessors and successors because they had unique developmental needs brought on by puberty:

This child is driven from his paradise and must enter upon a long viaticum of ascent, must conquer a higher kingdom of man for himself, break out a new sphere, and evolve a more modern story to his psychophysical nature. (p. 71)

Tanner's (1962) research was not unlike Hall's, showing a decline in the average age of puberty for the early adolescent. He found early adolescents were experiencing puberty earlier, approximately 4 months earlier each decade from the 1900-1960's. These results were used to help justify the reconfiguration that resulted in the move of 6th graders to the middle school and 9th graders to the high school level. Eichhorn (1966) coined the term "transescence" as the developmental stage of early adolescents. He defines it as:

The stage of development, which begins prior to the onset of puberty and extends through the early stages of adolescence. Since puberty does not occur for all precisely at the same chronological age in human development, the transescent designation is based on the many physical, social, emotional, and intellectual changes in body chemistry that appear prior to the time, which the body gains a practical degree of stabilization over these complex pubescent changes. (pp. 3-4)

Eichhorn (1973), Havighurst, (1972), and Tanner (1962), all argued that students should be grouped according to their developmental stages and not their chronological age. However, cognitive developmentalists did not take this stance.

The cognitive development of the early adolescent during the middle school movement was defined using a Piagetian framework (Beane & Brodhagen, 1996). The developmental theory of cognition proposed by Piaget (Inhelder & Piaget, 1958) was on

the emergence of formal logical structures and was not specifically related to the uniqueness of the adolescent. According to this cognitive framework the early adolescent was at the concrete operational stage, a formal operational stage of development, or between the two (Beane & Brodhagen, 1996). In particular, their thinking was shifting from concrete understanding to more abstract and higher-order reasoning. Like his predecessor (Hall, 1908), Piaget too was concerned with the developmental unique stage fit of the early adolescent, and should receive instruction appropriate with their developmental stage. However, Piaget's theory was deficient concern for a broader array of biological, emotional, social, and societal concerns engaged in other theorists' discussions (Beane & Brodhagen, 1996).

A new paradigm in the 1980's middle school reform movement impacted this population of learners. A call to society's lack of attention was brought to focus by Lipsitz's (1980) book *Growing up Forgotten*, which stated that the early adolescent was generally underserved and that education should address the "whole child." The focus should not only include an understanding of the development stage for the early adolescent, but also an understanding of the social relationships and affective conditions that influenced this population of learners.

Johnson, Markle, and Stingley's (1982) research investigated how peer acceptance was related to academic achievement. Greenberg, Siegel, and Leitch (1982) studied adolescents' attachments to their parents and peers using a newly developed psychometric instrument *Inventory of Adolescent Attachments* that measures self-esteem and life satisfaction of relationships with parents and peers. Using a hierarchical regression model with 213 early adolescents (ages 10-14) the researchers contend

attachments were more powerful with parents than peers in measures of well-being and self-esteem.

Another line of studies focused on the affective issues of the early adolescent literacy learners. Mager (1968) and Rosenshine (1980) provided data that suggested the students' attitude is directly related to learning and that school climate impacts students' attitude. The shift in focus moved to not only understanding the early adolescent's physical and psychological developmental needs but also how these needs matched the learning environment provided for this population.

Recognition of the need to understand the whole child was explored in Alexander and George's (1981) book *The Exemplary Middle School*. The authors contend that the reconfiguration of the middle school had very little to do with academic achievement of the early adolescent. Instead what should be used to guide student achievement are the characteristics of an exemplary middle school model. The researchers offer 12 characteristics for this model:

1. Statement of school philosophy and goals
2. System for planning and evaluating designed for middle school and involving all stakeholders (school administrators, teachers, parents, and students)
3. Curriculum plan that provides instruction that builds continuous progress and meets the differentiated needs of the population
4. Guidance and relationship with adults
5. Interdisciplinary planning, teaching, and evaluation
6. Flexible grouping for instruction
7. Block scheduling to provide flexible and efficient use of time

8. Varied programs
9. Instruction which utilizes a balanced variety of effective strategies to achieve continuous progress of each learner to meet instructional objectives
10. Strong leadership, and professional development
11. Plan for evaluation for both the students and the school
12. All stakeholders working to meet the needs of the early adolescent learner.

(pp.18-19)

Along with the line of research that addressed the developmental needs, contextual conditions, and a match to effective instructional practices, was a concern with the transition to a middle school during the onset of puberty. This was thought to be disruptive for the early adolescent.

Simmons and Blyth (1987) conducted a comparison study across two different school configurations of 7th grade students. One group of 7th graders transitioned at the beginning of their 7th grade year to a middle school and a second group of 7th graders remained in a K-8 school. Using a short-termed longitudinal design, indices of self-concept, social adjustment, school attitudes, as well as academic achievement, the researcher's assessed 160 adolescents both prior to and during the transition of middle school. They found the students who transitioned to the new school configuration had lower self-esteem, lower grades, and more negative attitude towards school. Eccles, Lord, and Midgley (1991), replicated this study by using the National Education Longitudinal Study (NELS: 88, 1988) data to compare 8th grader students who attended a K-8 school and 8th graders who were students in other school configurations (junior high/ middle school).

The researchers documented that the transition to junior high or middle school was marked by a general decline in the students' motivation, attitude about school, perception of ability, and academic achievement. The researchers proposed that it was not a good fit between the developmental needs of the adolescent and the environmental change.

In their study Eccles et al. (1991) argued that early adolescent were facing changes (social, emotional, physical, psychological, and cognitive) and the school environment provided did not fit their needs. Instead of providing for the developmental needs of the students (e.g., wanting more autonomy), they were given less choice and had more restrictions placed on them. As a result of the poor match between developmental needs and the transition into middle school, the students showed decreased motivation, self-esteem, and academic performance (Juvonen et al., 2004).

However, the effects of middle school transitions have varied across studies. While some researchers such as Simmons and Blyth (1987) and Eccles, Lord, and Midgley (1991), argued that, a negative effect existed with these transitions to the middle school for the adolescent, other researchers illustrated the adjustment had no adverse effects on these students (e.g., Crockett, Petersen, Garber, Schulenberg, & Ebata, 1989; Hirsch & Rapkin, 1987). Regardless of these alternative reports the Carnegie Council focused their recommendations on the negative developmental fit.

The Carnegie Council on Adolescent Developmental (1989) presented a powerful vision for the middle school and their learners in *Turning Points: Preparing American Youth for the 21st Century*. The Carnegie report (1989) concluded:

Middle grade schools-junior high, intermediate, or middle schools-are

potentially society's most powerful force to recapture millions of youth adrift. Yet too often they exacerbate the problems the youth face. A volatile mismatch exists between the organization and curriculum of the middle grades schools, and the intellectual, emotional, and interpersonal needs of young adolescents. (p.32)

In this report the Carnegie Council (1989) identified five overarching goals the early adolescent student should attain on leaving the middle school. They should be: (1) an intellectual caring person, (2) a person en route to a lifetime of meaningful work, (3) a good citizen, (4) a caring individual, and (5) a healthy person. In order to achieve these goals the council made eight recommendations:

- dividing large middle schools into smaller communities of learning
- students should all be taught a core of common knowledge
- ensure success for all students
- empower teachers and administrators
- prepare teachers to teach the middle grades
- improve academic performance through better health and fitness
- connect schools with communities

These students need an understanding of their unique developmental needs and instructional practice to match their needs. It was these tensions that complicated the lives of the middle school content area teacher (Brodhagen & Bean, 1996).

The Middle School Teacher

The contextual debate on whether the middle school should be more like the elementary classroom that emphasized a (child centered approach to teaching) or the high

school that emphasized (disciplinary rigor) placed the middle school teacher in the tensions of the contextual configurations and developmental needs of the early adolescent (Brodhagen & Bean, 1996). Along with general contextual and developmental issues for the middle school teacher a concern was what should be taught and how instruction should be delivered. Specifically, this left the middle school teacher in a state of ambiguity, questioning whether they were content specific professionals, child centered developmentalists, or somewhere in between.

In his study of organizational design and instructional features McPartland (1987) drew data from a sample of 433 schools in the Pennsylvania Education Quality Assessment. The purpose of the study was to examine effects of instruction that was accomplished through a self-contained classroom setting and instruction that was departmentalized, while looking at: (a) student-teacher relationship and (b) quality of subject matter instruction. McPartland concluded self-contained classrooms were conducive to student-teacher relationships however, departmentalization instruction provided higher quality instruction. He recommended a balance of instructional features that combine both a personal relationship with students and mastery of the teacher content, to benefit the early adolescent learner.

Becker's (1987) research investigated whether different grade level configurations (elementary or middle) affect learning for the students with different abilities and especially socio-economic levels. From a sample of 8,000 sixth graders in Pennsylvania he determined that the elementary school setting and instructional approach benefited students from a lower social economic backgrounds because of different experiences and background knowledge related to school instruction.

Researchers of curriculum instruction, Lounsbury and Vars (1978), Hodgkinson (1986), and Slavin (1988) identified improvement in learning when cooperative techniques were in place. Mollified (1988) stressed the need to balance learning needs for the early adolescent, and to provide professional development for teachers.

Mac Iver and Epstein (1993) researched Elementary and Middle Schools (CREMS), through the Johns Hopkins Research Center. They conducted a survey with principals of 2,400 schools in the United States which included seventh-grade students. A total of 1,753 (73 %) provided information about their schools. 1,344 returned the surveys by mail, and 409 completed surveys by telephone. The telephone interviews were conducted through a random subsample of all nonrespondents to the mail survey. The researchers used multiple regression analyses to identify significant consequences of instructional practices by middle school teachers for their students.

The focus of the study was to investigate instructional delivery (strategic approach in reading), teacher- student relationships, organizational instructional formats (interdisciplinary teams, or departmentalization), and remediation for students. They found responsive practice (strong teacher to student relationships), and support for students who struggle (extra period during the school day) are most beneficial. In addition, instructional organization through interdisciplinary teams was shown to be more responsive to the needs of the early adolescent rather than departmentalized organization. However, even with the compliance with responsive practice, most middle grade instructional delivery emphasized drill and practice and infrequently used interactive instructional approaches or cooperative learning. This practice especially impacted

strategic processes that would be used to understand content material, particularly in reading and comprehending content subject materials.

Historical Parallels of Reading in the Content Areas

Historically, content area reading origins transpired, as a result of a readers' need for strategies when they engage in certain subject areas, with many different types of texts for different purposes. Specifically, content area reading instruction is designed to deliver those strategies, so that students develop reading-to-learn strategies across and within various content areas (Moore, Readence, & Rickleman, 1983).

In a historical review Moore et al. (1986) presented an historical overview of this field by presenting the origins of content reading and a discussion on how best to deliver instruction. The researchers assert that the historical review is of “public discourses...tracing the prominent opinions and research findings, that were reported in journal articles, conference proceedings, and textbooks” (Moore et al., 1983, p.420).

Instruction in general during the early 1900's consisted of rote learning. Students were responsible for memorizing and then reciting information back to show evidence of learning. This changed, however, with the turn of the century as new goals for reading instruction were influenced by humanists, developmentalists, and scientific determinists (Moore et al., 1983).

Humanists were concerned for the development of the whole child, and the schools were charged with ensuring that learning should be meaningful and a student should be an independent thinker. The Progressive Movement was derived from the humanistic stance, contributed to meaningful reading (Moore et al., 1983).

Dewey (1910) and James (1923) two compelling factors in the progressive movement contributed to moving the education field forward for meaningful reading. Dewey (1910) criticized rote learning and argued that learning should be connected to a child's experiences, interest, and problem solving abilities. In his classic work, *How We Think* (1910) Dewey presented the theoretical development of reflective thought and how that should transfer to practice. James's (1923) work was concerned with the child knowing factual information, but not being able to make inferences about the information read. It was this meaningful, inferential learning and independent thinking that carried clear implications for the reading process (Moore et al., 1983).

Developmentalism also became influential to content area reading history. Identifying the needs of the early adolescent through child study, psychologists (e.g., Gesell, 1915; Hall, 1908) informed the reading field pertaining to growth and development patterns of the early adolescent. Reading educators (e.g., Gray, 1939) translated this practice to reading. Gray's (1939) research focused on reading strategies for purposeful reading. Gray noted:

instead of assuming that pupils enter the higher grades with fully developed and adequate reading habits, an essential step on the part of all teachers is to ascertain the level at which their pupils can read with ease and understanding... This may be different for each student, but is necessary for teachers to identify the developmental level of each student. (p.7)

His findings contend reading strategy instruction should be for all students beyond the elementary grades, and instruction should be differentiated to meet the literacy needs of all students.

Along with the progressive movement and developmentalism, the Scientific Determinists called for scientific empirical support of reading. Scientific Determinists looked for one absolute truth about the reading using the process of the scientific methods. However, there was debate between social efficacy groups and reading researchers about interpretations. Social efficacy groups sought to identify the most effective ways to measure students' academic ability in reading. They argued for the use of standardized testing measurements as a way to determine the students' reading achievement (Callahan, 1962). Test developers Binet (1904), Rice (1913), and Thorndike (1917) investigated standardized testing instruments measuring reading comprehension. In this way tests could be administered and scored under a consistent set of procedures, and this would make it possible to compare results across individuals and schools. These instruments measured reading comprehension without the benefit of direct instruction. Resnick and Resnick (1977) contend that students need support in comprehending text, and assessing comprehension without explicit instruction does not accurately measure what the student understands.

Huey (1908) and Thorndike (1917) examples of pioneers helped lay the foundation in reading theory and practice. Huey's (1908) form of inquiry in the literacy field explored the psychological influences of reading comprehension particularly how children's personal background literacy experiences influence their reading development. Discussions between the researchers consisted of oral language acquired both in and out of school, playing with sounds and words, and developing schema about the complex process of reading. Thorndike (1917) explored and investigated the complexity of comprehending text, cognition in reading and how internalizing reading moves through

questioning from oral to silent guided reading. In his classic work *Reading as Reasoning: A Study of Mistakes in Paragraph Reading*, Thorndike (1917) conducted a quantitative study of 200 sixth grade students to understand the reasoning process in reading. He found readers need a predetermined purpose for reading. Thorndike argued for oral reading to be replaced with silent reading, and to have students ask themselves questions while reading, answer questions after reading, and summarize material that they read. Thorndike (1917) concluded that “Perhaps it is in their outside reading of stories and in their study of geography, history, and the like, that many school children learn to read” (p. 282).

Studies also looked at the correlation between student academic achievement and reading. Smith (1919) compared subject matter achievement in math with grades in English and found high correlation among the measures. He concluded reading ability was related to school achievement. Along with Smith’s work, Wagner’s (1938) work measured reading skills in nine areas of subject matter achievement for ninth grade students. She found ability in composite reading comprehension was related strongly to composite ninth-grade achievement in other content areas.

Instructional Practices

Historically along with the conceptual understanding of content area reading there were issues relating to how instructional delivery should occur. In essence there were two forms of instructional formats existed: direct, skill-centered instruction and functional, content center instruction (Herber, 1970). Direct instruction strategies to understand content occurs when teachers identify a set of skills and present them to students regardless of the content tasks. Reading researchers and educators Gray (1919), Gates

(1935) and McKee (1934) gave arguments that by providing systematic reading instruction students were sure to receive instruction in all skills that were deemed important. This placed content specific learning second to reading skills. Another argument for direct instruction was secondary-school educators only assumed that they were responsible for content specifics and by requiring reading skills taught across disciplines, students would acquire the process for understanding the content.

Functional content centered instruction occurs when content teachers identify reading skills that are a prerequisite for comprehending content material. These skills are then presented along with the subject matter to be taught. This format of instructional delivery was endorsed by early progressivists (e.g., Parker, 1894; Thorne-Thomsen, 1901). These researchers claimed reading would be enhanced through the study of various content subjects. Therefore, reading and specific skills and strategies should be embedded in the content instruction.

Moore et al. (1983) contend there were two main historical arguments for content-centered instruction: motivation and transfer. Motivation assumed the affective aspects of reading; if students were interested and understood the purpose for reading the content material, they would improve their reading. Motivation in reading can be defined as the cluster of personal goals, values, and beliefs that an individual possesses and applies in a literacy situation (Guthrie & Wigfield, 2000). Transfer of reading skills and strategies concerned the ability to use specific skills learned in one content area and transfer it to another. In *Teaching Reading in the Content Areas* Herber (1970) addressed determining whether early adolescent would be best served by reading instruction in separate reading periods or during the presentation of content material. Herber (1970) believed functional

instruction is the preferred method, where content teachers address reading abilities while teaching the content specific subject matter. This contention received empirical support from a series of investigations (e.g., Herber & Barron, 1973; Herber & Riley, 1979; Herber & Sanders, 1969; Herber & Vacca, 1977).

Skills and Strategies Related To Specific Content Areas

Along with general instructional procedures related to content area reading, a historical question arose whether there should be content specific reading skills and strategies or generic reading skills and strategies taught across all content areas remains of interest. Judd and Buswell's (1922) studies involved an eye movement analysis over seven different content areas. They found different types of text materials require different strategies. They measured the number of eye fixations per line, duration of fixations, and the number of regressions that differed according to text being read. They recommended that across various content areas there should be different reading skills and strategies to access the content material being read.

Vocabulary frequency counts, and difficulty with subject matter technical and vocabulary within various content areas, were also of concern historically. Thorndike's (1921) work sought to scientifically measure the frequency of vocabulary uses from sources that students would have to read. Lists of words were generated and these words were tested for accuracy and automaticity by timed tests (15 minutes) (Dolch, 1928). Along with the number of commonly occurring high frequency words, studies were conducted on technical vocabulary within various content areas.

Pressley (1923) collected 200 school texts of various subjects, and then subjectively chose words she felt were words that appeared frequently and that required

understanding for the content area. Although the methodology was flawed because of subjectiveness of word choice, other studies in content specific fields: gave evidence supporting for the need to understand technical vocabulary within various content areas to comprehend the subject matter: history Barr and Gifford (1929), math Buswell and John (1931), and science Curtis (1938) (Moore et al., 1983).

Comprehension

Another line of research investigated how the early adolescent comprehends various content materials. Ritter and Lofland (1924) studied the correlation between comprehension questions that were answered after reading content specific expository passages (e.g., science passage from text) and comprehension of general narrative reading (e.g., passage from book or language arts text) tasks. They found correlations varied among different grades and individuals. They interpreted the findings as meaning reading competencies were to be learned within the context of the content area to be comprehensive.

McAllister (1930, 1932) conducted a qualitative study assessing content reading materials and classroom tasks. He used observations of subjective analysis with students' written reports and interviews. McAllister concluded differences in students' comprehension within various content areas because of the type of reading activities and the support given to students to complete the tasks. However, he also suggested there should be considerations for more generic reading skills and strategies that could be taught and modified to meet the specific content material.

Generic treatment of reading skills and strategies is based on the premise that one common set of skills and strategies can be used in various content areas and be adjusted

to meet the needs of each content area (Moore et al., 1983). Strategies such as comprehension monitoring, fluency, and questioning can be used in all content specific classrooms and be modified to meet the specific content demands. This way the integrity of the content remains and content teachers are using reading skills and strategies to assist their students.

Textbooks

In the early 1900's students' textbooks were McGuffey Readers, selections in prose and poetry for the reading classrooms. Furthermore, carried over from the 1800's were messages including religious and moral themes (Moore et al., 1983). Most of the texts were narrative in nature. This created problems, however, when the early adolescent had to read expository texts, because of the different strategies and skills needed to comprehend this textual format. Content area teachers often used a single textbook to teach content specifics. The content teachers had difficulty using the content textbooks to meet the needs of their students because of the difficult language the new text utilized. In addition, teachers were not trained effectively in pedagogy (Beane & Brodhagen, 1996). In the beginning of the 20th century the need to supplement the textbook was addressed.

In 1927, Good wrote *The Supplementary Reading Assignment*, which reported suggested practices to use supplemental books along with the classroom text. Kilpatrick (1919) and Whipple (1920) presented various units of study that used thematic approaches with supplemental materials. The use of supplemental materials in various content areas was difficult because of time, management, and cost (Moore et al., 1983). Also, the various levels and different needs of the readers in the classes added to the complexity of a single text for the classroom.

The Learner and Content Area Reading

Gray (1937), Kottmeyer (1944) and Witty (1948) contended reading instruction for the early adolescent in the secondary school setting during the early 1900's was a general remedial pull-out model, of large scale testing, and instruction in special classes. The problem was that content area teachers then had to provide reading instruction that would transfer to meet the needs of these learners in their classroom. These concerns for remedial programs helped focus content teachers on reading in the content area and the need for not only students who struggled but also average readers who needed support in reading content material (Moore et al., 1983). School testing demonstrated reading development did not stop in the elementary grades, and reading abilities were seen to have no upper levels. Therefore, differentiated reading instruction was determined to be important in the middle grades (Bond & Bond, 1941).

During the time of the Carnegie Report of 1989, which, urged a development fit for school instructional practices for the early adolescent learner, Alvermann and Moore's (1991) review of the secondary school reading practices gave insight into reading in the content areas. The dominant instructional activity in the secondary school content reading practice was a combination of lecture, textbook assignment, and classroom recitation (Holton, 1982). Dolan, Harrison, and Gardner (1979) noted that half of all classroom reading occurs in short bursts of less than 15 seconds in any one minute. Usually, according to the researchers, these reading bursts were combined with speaking, listening, or writing activities.

Textbooks or teacher lectures were the primary source of information. Conley (1986), and Mitman, Mergendoller, and St. Clair (1987) concluded this was due to a lack

of teachers' knowledge and self-confidence about reading practices, including how to integrate reading skills within content information. Wiley (1977) reported a single set of textbooks was used to relate content material over the course of the school year. The teacher could assess content material by asking questions requiring verbatim responses from the texts. The purpose was to control classroom discourse to specific factual information extracted from the single text.

Secondary teachers relied on a single textbook, due to concerns about time and resources (Alvermann & Moore, 1991). The time to prepare for projects and pulling supplemental material, due to the changing of classes with multiple students, and the need to cover content materials made relying on a single textbook easier (Dillion, 1983). Content coverage and price often determined adoption of specific textbooks. Unfortunately, the comprehensive coverage of content often left the early adolescent unable to comprehend the material needed (Broudy, 1975; Coser, Kadushin, & Powell, 1982). Pearson and Fielding (1991) contend that a sequence is necessary to be most beneficial in order to build reading comprehension. They suggest:

The optimal context for independent contextual practice may be one in which practice is preceded by instruction, it is carried out on appropriate materials, is monitored to insure students actually are engaged, and it followed by response of feedback to what is being read. (p. 850)

Summary

The last one hundred years have shifted the contextual focus on the early adolescent learner from the “wasted grades” in the elementary schools to discipline oriented practices of the high school. The developmental nature of the learner also shifted

from psychophysical nature to understanding the unique qualities of the early adolescent learner. The middle school and its philosophy brought attention to unique needs of this population of learners. There was a call for the developmental fit for the learner, within the school context, delivered through instructional practices concerning reading in the content area.

Unfortunately, until the 1980s research on the contextual issues of the middle school, the overlap of the needs of early adolescents, and how best to provide instructional practice, was of remarkably low quality (Johnston, 1984). The low quality was attributed to weak design and methodology, and as claimed by Wiles and Thompson (1975) in their analysis of the research on the middle school, research by proponents and opponents of the middle school movement merely studied and reported outcomes that confirmed their subjective positions. The late eighties and nineties witnessed many changes in the early adolescent literacy learner and their needs within the context of middle school however, they are still faced with many new and yet similar complexities as their historical predecessors. Table 1 provides a timeline of implications from the middle school movement and how this historically parallels the evolution of content reading.

Table 1

Middle School Timeline

Period	Middle School	Content Reading
1900's	National Education Association called for the reconfiguration of secondary school for 7th and 8 th grade, taking a developmental stance.	Progressive Movement Dewey (1908) James (1923) learning should be connected to child's experiences, and should provide opportunities for critical thinking.
1908	G. Stanley Hall- book on adolescence unique stage of development.	Research explored the psychological influences on reading comprehension. Huey (1908) and Thorndike (1917)
1917	Smith Hughes Act- beginning of vocational curricula programs.	Testing instruments to measure reading comprehension developed.
1920	The development of the Junior High School by the Commission on the Reorganization of Secondary Education.	Reading in the content area should provides direct systematic instruction. Secondary-school teachers assumed they were responsible for content.
1956	Gruhn & Douglas-dissatisfaction with instructional programs for the early adolescent. Need to match instruction to the students.	Dissemination on theories related to diversity of reading levels and needs for the early adolescent in the various content classes.
1960	Junior high has too many similarities to senior high. The evolution of the middle school.	Research on motivation and transfer of reading skills. Also, whether reading should be taught in a separate class or in the content class.
1981	Alexander & George's book <i>The Exemplary Middle School</i> .	Herber's work addressed functional instruction- addresses reading abilities with content materials.
1989	Carnegie Council concluded that understanding of the developmental needs of the early adolescent and a match of those needs to practice.	Pearson & Fielding work on reading comprehension found that practice with text is preceded by instruction.

Current Complexities of Literacy Learning

The Early Adolescent Literacy Learner

As previously discussed, historically the focus was to understand the unique stage of development for this population of learners (e.g., Hall 1908), along with matching this developmental stage to the context of school (Alexander & George, 1981). However, through this historical lens there was very little depth in understanding the unique and complex issues for the learner, particularly involving their literacy learning needs.

Alexander (1998) developed and tested her “ Model of Domain Learning” for adolescent readers. She suggests that the early adolescent readers range across a developmental continuum in their reading with various texts. She believes that reading development can be traced in the evolution and interplay of three fundamental factors: prior knowledge, interest, and strategic processing. Her research suggests that there are stages that the reader goes through in their reading:

1. Acclimation: occurs when the reader is on unfamiliar terrain and this requires considerable strategic effort.
2. Competence: occurs when the reader is starting to efficiently process and becomes more fluent in their reading.
3. Expertise: occurs when the reader is comprehensive, fluent, creative, and analytical.

Alexander (1998) cautions, however that stages are not grade or age specific, and that a reader may be competent or an expert fluent reader in one kind of literacy task, and in turn may drop back to acclimation during another literacy task. This specifically

occurs when there is a lack of prior knowledge, interests, or strategic processing. This complexity was supported in Ivey's (1999) case study with three sixth grade students of varying reading abilities.

In her multicase study, Ivey spent five months with three sixth grade students who had different levels of success with their reading. She found these students were complex and multidimensional as readers. All were motivated to read texts they found interesting and had self-selected. Their disposition to read was dependent on the instructional environment in which their reading occurred. Ivey also noted that care must be taken about generalizations or labels placed on readers. The student deemed a "struggling reader" was able to read fluently and comprehensively when texts were at her instructional level, and her listening comprehension was also strong. The "average reader" was unmotivated to read and therefore, although fluent with words, lacked strong comprehension strategies. However, when this student was able to self-select books of interest this changed. The "capable reader" would read whenever it was requested but was troubled by not understanding the purpose for reading some school sanctioned texts. Therefore, reading occurred only when required, or when this reader understood the purpose of the reading assignment. Ivey cautions that labels or categories given to classify readers offers limited information about who the reader is and the complexities of individual experiences.

Although Ivey's study supports her argument that the early adolescent is a complex and multi-dimensional reader, this study occurred in a single classroom with three students. However, O'Brien (2001) also cautions us not to be too facile in our assessment of adolescents' literacy abilities.

Drawing on his work with “at-risk” students in a high school literacy lab, he argues that the students’ full literacy competence is not apparent solely by the narrow structure of school-sanctioned literacy. Instead he contends that in his research the students displayed sophisticated literacy skills as they are combined with art, sound, and print in their multimedia productions. O’Brien argues that we must recognize these students who are labeled as “at risk” can be “artistic, creative, innovative, and daring at using a variety of popular media... [T]hey are skillful and creative at constructing and interpreting a range of media texts... using a variety of symbols and signs for conveying and communicating” (p.3).

Along with understanding the diversity of the early adolescent literacy learner, their cognitive abilities are also under continual development (Phelps, 2005). Kuhn, Black, Keselman, and Kaplan (2000) study addressed instructional practices in the content area of science. The researchers did not conduct a reading study per se but they contend middle school student’s cognitive development is aided by both direct instruction and by practice.

Kuhn et al., experimental study consisted of middle school students in a multimedia science experiment project of six weeks. They found the treatment students outperformed the control students on the project final assessment. The experimental group received explicit direct instructions and practice on how to complete the tasks, whereas the control groups were given instructions on how to complete the tasks. The experimental group outperformed the counterparts on this task along with similar multivariable transfer follow-up tasks. The researchers argue that the early adolescent cognitive skills can be aided by both direct instruction and practice.

Understanding the full range of adolescent literacies and the role of literacy in adolescent development is important. This suggests the early adolescents need opportunities in school to explore both multiple texts and multiple literacies and to receive instruction and opportunities for practice and support from peers and adults. However, Cuban (1992) argues the contextual conditions afforded for the middle school students have not dramatically changed since the reconfiguration of the junior high school.

School Structure

O'Brien, Stewart, & Moje (1995) contend the infusion of content literacy into the middle school curriculum and school organization has changed very little over the last one hundred years. The institutional organization is formed around an approved formal curriculum divided by disciplines and is controlled by time and space through the context of school. Talbert and Bascia (1990) claim that this organization is framed around: (a) Six or seven class periods, about 50 minutes each, (b) Approved knowledge base divided into subject areas, and (c) Three or four elective classes that are mandated to meet core curriculum requirements. The success of the curriculum is gauged by content coverage, and the amount of seat time a student accrues (O'Brien et al., 1995).

Furthermore, given the unique individual differences among early adolescent literacy learners, curriculum delivery is often a one-size-fits-all practice (Alvermann, 2001; Ivey, 1999; Moore, 2000). Therefore, the integration of content literacy to meet the diverse needs for this population is challenged through the contextual structure and curriculum delivery. Regardless of the complications of the school structure, diverse needs of the early adolescent learner, and the pedagogical lens of the middle school

teacher, one of the significant complexities for this population is the political pressure from the reform movement with mandates and high- stakes testing.

Mandates and High- Stakes Testing

Historically, political and societal influences have impacted the educational process for the early adolescent learner for the last one hundred years (Cuban, 1992). The current effect of mandates and accountability through high-stakes testing as a result of the 2001 the *No Child Left Behind Act* (NCLB) creates the current academic dilemma for this population of literacy learners (RAND, 2005). Initially, prior to the advent of NCLB in 1997, Congress directed the Director of the National Institute of Child Health and Human Development (NICHD), in consultation with the Secretary of Education, to convene a national panel to assess the effectiveness of different approaches used to teach children to read. The National Reading Panel (NRP) issued a report in 2000 that responded to a congressional mandate to help parents, teachers, and policymakers identify key skills and instructional methods central to reading. Using these findings as a foundation for literacy instruction and implementation, NCLB established the Reading First initiative program under Title I, Part B, Subpart I of the Elementary and Secondary Education Act (ESEA). The goal of this initiative is to ensure that all children in America are reading at or above grade level by the end of the third grade (United States Department of Education, 2001).

The Reading First initiative focus was directed to reading improvement in instruction for grades kindergarten through third grade. As a result of this initiative, less focus was directed to the literacy learning needs of middle school students (Alvermann, 2001; Biancarosa & Snow, 2006; Guthrie & Wigfield, 2000; Kamil, 2002). McCombs, Kirby, Barney, Darilek, and Magee (2005) contend in their RAND report to the Carnegie

Corporation despite the reading progress made by primary grade students, this is not the situation for the early adolescent. deLeon, (2002) reports “ many children are not moving beyond basic decoding skills—deciphering and/or sounding out—to fluency and comprehension, even as they advance to the fourth grade and classes in history, mathematics, and science” (p. 1). McCombs et al., (2005) claim there is a need for continual instruction in reading beyond the third grade. However, teaching reading in the secondary schools to adolescents is an “orphaned responsibility” (deLeon, 2002).

In their study Amrein and Berliner (2002) suggest the Heisenberg Uncertainty Principle: “The more important that any quantitative social indicator becomes in social decision making, the more likely it will distort and corrupt the social processes it is intended to monitor”(p.5), applies to high-stake testing currently occurring in the schools. This principle, the researchers suggest, warns us that attaching serious consequences (e.g., high school graduation, retention, class remediation) to a high-stake testing environment may have serious personal and educational consequences.

The purpose of Amrein and Berliner’s study was to investigate whether the high-stakes testing program promotes the intended transfer of learning. A sample of eighteen states that had the most severe consequences because of testing results was used in this study. The effects of high-stakes tests on learning (general domain knowledge) as compared to training (narrow focus) were measured by examining indicators of student achievement with other standardized tests. The four different measures were:

- the ACT, administered by the American College Testing Program;
- the SAT, the Scholastic Achievement Test, administered by the College Board;

- the NAEP, the National Assessment of Educational Progress, under the direction of the National Center for Education Statistics; and
- the AP exams, the Advanced Placement examination scores, administered by the College Board. (p.20)

This study, according to Amrein and Berliner was to clarify the relationship between the scores obtained on a high-stakes test and the domain knowledge the test scores represents. The researchers used an archival time-series research design, to examine the state-by-state and year-to-year data on each transfer measure. The independent variables were before and after scores of high-stake testing for high school graduation. The dependent variables were and scores from year to year, (ACT, SAT, NAEP, and AP) before and after the implementation of the high-stake test. National trend lines were used as nonequivalent comparisons group along side the state trend lines. Also, correlations looked at participation rates in each state after high-school graduation tests.

Amrein and Berliner (2002) found the ACT data indicated 67% of the states that used high-school graduation exams posted decreases in ACT performance. These decreases were unrelated to participation rates, and on average, achievement on the ACT decreased. The SAT data indicated that 56% of the states using the high-school graduation exam posted decreases in SAT performance after the exams were implemented, however, these decreases were related to SAT participation. Nationally SAT participation showed a decrease of 61% in the states that used high school exit exams. Therefore, the researchers argue if these participation rates serve as indicators of

testing, the belief that high-stakes testing policies will prepare more students or motivate them to attend college, is not supported.

The NAEP data had limitations. Interpretation of data for the high school exams and the relationship with math and reading data for the fourth and eighth grade students is weak. The AP data however, showed high school graduation exams did not improve achievement for students as presented by the number of students passing the various exams. When participation rates were controlled the percent of students who passed the AP examinations decreased. Amrein and Berliner (2002) overall contend “ there is no compelling evidence from a set of states with high-stakes testing polices that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators” (p.54). In addition, the RAND Corporation was authorized by the Carnegie Corporation to investigate the data results of state instituted high-stakes testing and scores from the NAEP, regarding the state of literacy achievement for adolescents.

McCombs, Kirby, Barney, Darilek, and Magee (2005) were commissioned by the Carnegie Corporation to investigate the current state of adolescent literacy learning. Using data from the 2003 NAEP report, the researchers examined the results in reading achievement at the national level (NAEP), as compared to individual state reported achievement, for students who had reached proficiency in national literacy standards. McCombs et al., suggest that we need to be cautious, because there are differences in rigor of the state level tests and the testing at the national level. Specifically, when defining what it means to be proficient in reading. McCombs et al., suggest:

One important caveat to keep in mind is that, although we present data on the similarities and differences in the results of state assessments and the state NAEP, data from these two assessments are not directly comparable, because of the differences in the tests themselves and in the definitions of *proficiency levels* in the NAEP and state performance standards. While one could argue that state and national literacy goals should be reasonably similar, in reality there is debate about whether NAEP achievement standards are too challenging. Indeed, Linn (2003) points out that the proficiency standard on the NAEP is an ambitious one, intended to encourage greater effort. The National Assessment Governing Board (NAGB), which sets the standards for the NAEP, notes that the proficiency level on the NAEP indicates that students reaching this level have demonstrated competency over challenging subject matter. (p.4)

McCombs et al., findings suggest there are several concerns to meet the NCLB (2013) goal for proficiency. They are as follows:

1. Fewer than half the students meet state proficiency standards, less than half of the student's meet NAEP national proficiency literacy standard.
2. Overall, the pass rates on the middle school states assessments ranged from 21% to 88 %. However, between only 10% and 43 % of 8th graders scored at the proficient level of the NAEP Reading Assessment. The average pass rate of the 8th graders on the NAEP assessment was 32 %.
3. There is a wide disparity in reading achievement for the subgroups of students (disaggregated by race/ethnicity and poverty status). At the 8th grade level we see a difference of 26–28 percentage points between the proficiency rates of

white and African American students; 22–26 percentage points between white and Hispanic students; and 22–24 percentage points between economically advantaged and economically disadvantaged students.

4. At both grade levels (4th and 8th) students with limited English proficiency and students with disabilities trailed well behind their peers.

In conclusion McComb et al. (2005) contend that these findings are very disturbing, for our adolescent literacy learning, as they prepare for meeting the high demands of literacy needs for the new millennium. The researchers recommend that:

It is clear that simply mandating standards and assessments is not going to guarantee success. Unless we, as a nation, are prepared to focus attention and resources on the issue of adolescent literacy, our schools are likely to continue producing students who lack skills and who are ill-prepared to deal with the demands of post-secondary education and the workplace (p. 85).

Summary

Currently, although there appears to be emerging themes and important information being investigated about the developmental needs, contextual conditions, and instructional practices, the knowledge base for early adolescent literacy learners is still very much under-studied (Alvermann, 2002; Bean, 2000; Kamil, 2002; Moore, 1996). The research reviewed states some of the current complexities for this population of literacy learners. Historically, the early adolescent's literacy needs shifted from a superficial understanding of their development (physical, emotional) to a complex appreciation of the multidimensional nature of this literacy learner. However, the

pedagogical lens of the middle school teacher has not, it appears, addressed the unique needs for this population of learners.

Along with the instructional practices, the contextual conditions afforded to these students contribute to the complexities to meet their literacy needs. Curriculum delivery and contextual organization have continued to mimic historically the junior high. The structure of the divided departmental domains and the use of a single text for instruction is continued practice for this population.

The dominance of the accountability measures mandated by NCLB (2001) and evaluated through high-stakes testing has also added to the current complexities. It appears these tests have provided contributing factors as to how content literacy is provided to the early adolescent within the content specific classrooms. This is a result of sanctions applied to the teacher, school, district, and the state as: (a) denied diploma, (b) retention of students, (c) remediation mandates from the scores students attained, and (d) rewards and punishments for all stakeholders involved in this tests. This influences the instructional delivery for the early adolescent in literacy at many levels of their schooling. There are however, suggestions and recommendations for effective strategic practices practitioners, researchers, and all community members need to think about in order to prepare the early adolescent literacy learner for the future.

Recommended Effective Strategic Practices for the Early Adolescent Learner

Sociocultural Influences

A shift in the field of adolescent literacy in the last 10 years has occurred (Phelps, 2005). The research on the political and social climate afforded to this population of literacy learners has shifted historically from the “wasted grades” of the early 1900’s, to the developmental fit match of the 1980’s, to an appreciation of the sociocultural influences on literacy practices at the current times.

Cook- Gumprez, (1986), and Scriber and Cole (1981) suggest the sociocultural theories of literacy occur as literacy is used in specific contexts for specific purposes, and is socially constructed and constituted. The act of literacy is embedded in a network of social relations. Moje (1996) suggests that in the secondary content classroom the social context that shapes literacy practices is uniquely complex. Teachers and students in secondary classrooms move from class to class, teacher to teacher, and with a subgroup of peers. Teachers and students construct meaning about literacy and learning events based on values, beliefs, knowledge, depending on the contextual situation. Additionally, teachers and students bring meaning to these interactions through their past beliefs, values, and knowledge during social interactions (Moje, 1996). Studies that are guided by broad theories as a social construction have focused on how social interactions influence literacy learning (e.g., Myers, 1992).

Moje’s (1996) two year ethnographic study focused on how and why a content area chemistry teacher and her students engaged in literacy activities. Moje contends literacy in this classroom was practiced as a tool for organizing thinking and learning in the context of the classroom built on relationships with the teacher and students. Also, the

researcher explains that within this study the literacy strategies used were domain specific and did not transfer to other domains. Moje speculates the use of literacy strategies in content area classes should be domain specific and socially supported by the teachers and students in the classroom. Furthermore, strategies should be shown as how they could be used in other content areas. Moje also suggests that more research should investigate classroom interactions and how they play a part in shaping literacy practices.

Englert and Palinscar (1991) define their sociocultural approach to literacy instruction as the interdependence of social and individual processes in the construction of knowledge. When viewing literacy development from a sociocultural approach, literacy arises from the child's participation in social activities in which there are real reasons to use written language. Ryan's (2000) work investigates the research on peer groups' interactions, as a context for adolescent achievement, motivation, engagement, and socialization.

Peer Interactions

In her analysis on the research of peer group socialization for the early adolescent Ryan (2000) theorizes peers generally interact three ways with one another. During early adolescence, the peer group becomes a prominent context for development (Brown, 1990). The school and classroom provide opportunities for peers to interact throughout the day. Ryan (2000) reports "peer interactions consume significantly more time in adolescence compared to childhood" (p. 107). These interactions with peers can concern both academic (e.g., achievement) and nonacademic matters (e.g., engagement, motivation, self-efficacy, and interest). Ryan (2000) suggests three ways that early

adolescents generally experience peer interactions within the context of middle school: through information exchange, modeling, and peer pressure.

Information exchange occurs when adolescents have a discussion with their peers (Berndt, 1999). In an experimental study with eighth-grade students, Berndt, Laychak, and Park (1990) found that when adolescents had to make an academic decision, such as go to a rock concert or study for a test, they initially responded differently from one another. However, after discussing this dilemma with their peers, their answers were similar to their peers. This form of interaction could influence the early adolescent's choice to partake in the literacy task presented by the teacher if it was used effectively.

Modeling is another form of adolescent peer interaction. This interaction refers to individual changes in cognition, beliefs, or affect, which are a result of adolescents observing their peers (Ryan, 2000). Observing a specific behavior a peer performs or listening to a peer voice a certain belief can induce an adolescent to change their stance or adopt their peers' behaviors or beliefs. Schunk and Zimmerman (1996) reported peer modeling influences self-efficacy beliefs. In their study, they found that early adolescents who verbalized that they had difficulty with a task and then observed their peers have success with the same task then believed they could complete the task. The early adolescent, when faced with a literacy task, may have success by observing their peers. Peer pressure is the third way that the early adolescent interacts with their peers.

Peer pressure takes on the role of social reinforcement (Ryan, 2000). Brown, Lohr, and Eicher (1986) found that beliefs and behaviors that are discouraged by the groups are not likely to be displayed whereas beliefs and behaviors that are positively received by the group are more likely to surface. Therefore, participation in the literacy

tasks that the peer group positively received through this interaction could have a positive effect on the group's beliefs and decisions to participate by the group members.

Peer pressure may also play a role in how the peer group influences motivation. Brown, Lohr, and McClenahan (1986) report that peer pressure regarding school involvement, is significantly correlated with self-reported behaviors and attitudes regarding school. Ryan (2000) recommends further research on peer interactions within a domain specific classroom may fill in the gaps in the literature. The recommendations from the research of Moje (1996) and Ryan (2000) are used to frame this study's qualitative component. Ryan's theory on the three general categories of peer interactions will frame the interpretive case study, along with Moje's recommendations that research on interactions within the setting of the content classroom should be studied to inform practice as to how literacy learning could be shaped.

Effective Instructional Strategies

In 2004, to help address the issue of adolescent literacy learners, a panel of five nationally known educational researchers met with representatives of the Carnegie Corporation of New York and the Alliance for Excellent Education. The focus was to draw up a set of recommendations on how to meet the needs of adolescent literacy learners while propelling the field forward (Biancarosa & Snow, 2006). A list of 15 elements were reported and then divided into two sections: instructional improvements and infrastructure improvements.

The instructional elements consisted of: a) direct, explicit comprehension instruction, b) effective literacy instruction embedded in content, c) motivation and self-directed learning, d) text based collaborative learning, e) strategic tutoring, f) diverse

texts, g) intensive writing, h) technology, and i) formative assessment (Biancarosa & Snow, 2006). Researchers were urged to re-conceptualize how they perform research with early adolescent literacy learners. Investigations should combine different elements so important information about the early adolescent can be determined. The current study utilizes five of these elements. Biancarosa and Snow (2006) urge that we must meet these challenges because:

Literacy demands have increased and changed as the technological capabilities of our society have expanded and been made widely available; concomitantly, the need for flexible, self-regulated individuals who can respond to rapidly changing contexts have also increased. The goal in improving adolescent literacy should not simply be to graduate more students from slightly improved schools, but rather to envision what improvements will be necessary to prepare tomorrow's youth for the challenges they will face twenty and thirty years from now. America's schools need to produce literate citizens who are prepared to compete in the global economy and who have skills to pursue their own learning well beyond high school. (p. 9)

Direct Explicit Comprehension Instruction

There is an enormous amount of research on reading comprehension. Specially, Dunkin's (1978-1979) work is pivotal for understanding the need to address reading comprehension for middle school students. Durkin's monumental work in reading comprehension was in search of how teachers in the field assist children in developing a more critical and deeper understanding about what they read. A request for proposals from the National Institute of Education (NIE) for studies in reading comprehension led

Durkin to undertake this study. The NIE assumed reading comprehension could be taught, was being taught, and yet instruction in comprehension was not as effective as it should be.

Durkin, a veteran observer of the classroom was struck by the second assumption. In her frequent visits to the classroom she had witnessed almost no comprehension instruction being taught. This may be because studies in comprehension instruction were never the focus of previous research, and observations were centered at primary grades. To address this Durkin went in search of the literature to define comprehension and placed her focus on observations in middle and upper elementary grades looking not only at the reading block but also in the content area of social studies.

Durkin conducted observational studies for four years in an elementary fourth grade reading classroom, and in grades three to six during a social studies class period. She reported that comprehension instruction consisted primarily of answering questions, completing workbook pages, or taking tests. Researchers however, questioned Durkin's criteria for determining what constituted instruction (e.g., Hodges; Heap, 1982). Pearson and Fielding (1991) contend however, this work that motivated other researchers to pursue the meaning of comprehension instruction. The researchers suggested the first and most important issue was to recognize the complex process of reading comprehension is not a passive process, but an active one.

Pearson's (1985) work on explicit instruction for comprehension was an example of research motivated by Durkin's definition. He and his colleagues provided a model that teachers could use to support their students and demonstrate how strategies would build comprehension. The Gradual Release of Responsibility Model of Instruction

(Pearson & Gallagher, 1983) suggests teachers model an instruction strategy and have students practice that strategy with guidance followed by independent practice. The role of the teacher is to model, guide and release responsibility to their students. This is accomplished through teacher modeling of their cognitive processes, then assisting and scaffolding students to share their cognitive processes, and finally releasing responsibility to the individual learner. Using this model the teacher facilitates, models, and coaches the learner not to provide individualized instruction but to monitor progress individually. This form is aligned with Vygotskian (1978) principle of moving students when they are directed from an adult, to the point where they can take control of their own learning. Therefore, instruction is scaffolded, through support of the teacher to help students carry out the literacy task (Langer, 1984).

Strategy Instruction Embedded in the Content

Where this instruction takes place and how it assists the students to understand the material in the content area is important. In order to address these concerns and to meet the literacy needs of early adolescent, it is important to investigate how literacy is embedded into the content areas (Snow & Biancarosa, 2006). Literacy embedded in the content addresses two directions for instructional implementation (Snow & Biancarosa, 2006). First, within the Language Arts classroom these principles are not discrete skills or techniques, instead the emphasis should be to teach the strategy or skill using other content- area materials. Second, content area teachers should encourage literacy skills and strategies that emphasize the reading and writing practices that are specific to their subject area (Alfassi, 2004).

Alfassi's (2004) research investigated literacy that was embedded in content.

In her study Alfassi conducted two sequential experimental studies, over the course of a school year. The studies were interrelated, examining the efficacy of two models of reading strategy instruction (Reciprocal Reading Model and Direct Explanation Model). The studies were conducted in a Midwestern high school with proficient readers. The first study was in an intact heterogeneous freshman English language arts classroom, with 49 students. The experimental group consisted of 29 students, whereas, the control group included 20 students. Teachers of the treatment group were involved in a six-hour strategy training session.

Eight expository passages from the student's textbooks were used. Fry readability was conducted on all passages (Fry, 1977). In addition, 10 comprehension questions, created by the researcher, requiring short answers following the reading, were completed without the use of the text. Questions were both explicit and implicit (Pearson & Johnson, 1978). Two independent raters (reading specialists) read the questions and classified them. The internal consistency of the questions as measured by Cronbach's alpha ranged from .71 to .85. At the end of treatment the teacher gave a Gates-MacGinite Reading Comprehension Test (2000). This standardized test was used to investigate transfer effects from strategy instruction, to reading comprehension application.

Alfassi contends Study 1, demonstrated that using authentic texts and strategy instruction within the language arts class resulted in significantly better results, $F(2,44)=4.08, p < .05$ than their counterparts who were just exposed to literacy strategies, without the benefits of explicit instruction.

In Study 2 the sample participants were 277 sophomore students in four different content classes (science, arts, social studies, and math). Each of the four classes

combined specific strategy instruction with content specific instruction. The researcher investigated differential effects of combined strategy instruction to answer different types of questions (explicit, and implicit). Text-driven questions (explicit) related to information in the test, and knowledge-driven questions (implicit) information gleaned from the test. The results showed after the intervention there was a significant improvement on implicit questions, $F(1,276) = 12.84, p < .001$. The findings suggested students improved comprehension especially with implicit questions and with explicit strategy instruction. Overall, Alfassi claims that in order for readers to construct meaning from text explicit instruction embedded in the content area can support all readers.

Diverse Texts

Along with strategies to comprehend text, it is important to have texts the early adolescent is able to read. Too often texts in the content classroom are too difficult for students to understand. Diversity in text selection for the content classes addresses two issues: (a) interest, and (b) readability for the students to understand and access the materials taught.

In their studies Worthy, Moorman, and Turner (1999) and Ivey and Broaddus (2001) investigate middle school students' interests, engagement and motivation for reading. Worthy et al. (1999) conducted a two-part survey study of reading preferences with 12 sixth-grade language arts teachers and 426 of their students, from an economically and ethnically diverse district in Texas. They found a gap between students' preferred reading materials and what they were given in school. In addition, when students were interviewed, they were readily able to give the names of their

favorite books, or authors. Worthy et al., considered this evidence that students' attitudes toward reading are not as negative as assumed.

Ivey and Broaddus (2001) surveyed 1,765 sixth-grade students, in 23 diverse classrooms, located in the northeastern and Mid-Atlantic States. The purpose of the study was to describe the early adolescents' motivation to read. The researchers found that time to read books, and teacher read alouds are what appeared to motivate these students. While other researchers (e.g., Allington, 1977) have studied the benefits of time for reading to improve reading, this study found that students felt independent reading was a time to make sense out of what they read.

Along with affective issues related to reading, the need to read texts at the student's instructional reading level is important. Biancarosa and Snow (2006) suggest that too often students become frustrated when the book is too hard for them to read. Given the wide range of reading abilities at the middle school level (e.g., Ivey, 1999), texts must be accessible for this diverse population, and meet the various interest levels of the students. Therefore, middle school content classrooms should have diverse texts, especially high-interest texts of varying reading levels. The current study uses alternative genre of songs as diverse texts. The 200 songs have been analyzed for an instructional reading level.

Technology

Most middle school content area instruction in reading is textbook centered, which presents a formidable task for early adolescents in their reading. Alvermann (2003) suggests that this may be because the students are not able to gain the necessary background knowledge and specialized vocabulary because they are infrequently able to

read their textbooks. The early adolescent who may have difficulty with the linear textbook, is often more adept in media text, which also motivates and engages them while connecting them to real- world interactions (Alvermann, 2003).

The computer offers students more control in terms of support, pace and active processing of text (Kamil, 2002). The use of technology as an alternative text, links real world experiences and interests and provides a sound base for its use with early adolescent readers. The National Reading Panel (NRP) (2000) reports that there is little empirical research on the topic of the relationship of hypermedia that supports literacy learning and instruction for middle school readers. However, there is promising evidence from the synthesized work by (Leu, 2000) on the effectiveness of literacy instruction for this audience. Leu (2000) reports on the positive effects for middle school readers when print and visual texts (e.g., hypermedia, the internet, and interactive CD-ROMS) are utilized.

A meta-analysis of the effects of technology and reading for middle school learners was conducted by Pearson, Ferdig, Blomeyer, and Moran (2005). They were commissioned by the North Central Regional Educational Laboratory (NCREL) Center for Technology to investigate experimental and quasi-experimental studies over the last decade in literacy and technology. The purpose of the study was to investigate technology tools used with middle school students addressing the reading areas of: (a) strategy use, (b) metacognition, (c) reading motivation, (d) reading engagement, and (e) reading comprehension. However, Pearson et al. found little experimental research for reading and technology use in the middle grades.

The research that does exist according to the researchers focuses on comprehension with a slight emphasis on metacognition. The researchers acknowledge that even though the empirical knowledge is weak, there are many excellent theoretical arguments grounded in best practice. Many offer compelling cases that support the use of technology to enhance literacy learning. Although this analysis yielded no strong claims for practice, it did have several recommendations for further research. The researchers recommend that future studies investigating the use of literacy learning through technology for middle school students consider:

1. more experimental and quasi-experimental studies using some sort of correlated design (pretests used as covariates for posttest or repeated measures).
2. balance issues of focus on control and precision for five weeks or more, longer studies might have maturation effects or other confounding variables.
3. smaller sample sizes more manageable than larger samples. There might be a trade-off between statistical power and experimental precision, however, it may be easier for researchers to maintain a high degree of fidelity to treatment in smaller studies because of the greater manageability prospects.
4. follow the Complementarity Principle: (a) start with a small descriptive study, then (b) a formative experiment that narrow the range of relevant variable, followed by (c) carefully controlled randomized experiments, and finally (d) conduct a full scale experimental study.
5. more studies that explore the relationship between commercial products developed to address the literacy needs for the middle school. Little research has investigated commercial technology products used for improving literacy acquisition at the

middle school level. (pp. 19-23)

The current concurrent mixed methods study uses a quasi- experimental design for the quantitative phase investigating a commercial interactive sing-to-read program *Tune Into Reading* (TIR) (Electronic Learning Products, 2006), with middle school students in a music classroom. The alternative text format is an individual computer program, originally developed to improve singing, which uses a vocal range analyzer that tracks the singer's pitch and rhythm, comparing it to the correct pitch of the song. Each student uses a headset with a microphone, linked to the computer to sing along repeatedly and to record their singing. As suggested by Pearson et al (2005) there is a need to understand this alternative text format and its relationship to literacy learning for the middle school student.

Motivation and Engagement

Motivation in reading can be defined as the cluster of personal goals, values, and beliefs that an individual possesses and applies in a literacy situation (Guthrie & Wigfield, 2000). Central to most theories on motivation is a student's sense of self-efficacy, a belief in how competently he/she will perform a specific task (Bandura, 1997). Providing early adolescents with clear goals for a comprehension task and giving them feedback on their progress can lead to increased self-efficacy and greater use of comprehension strategies (Schunk & Rice, 1993).

In a longitudinal study of sixth and eighth grade students Wenzel (1996) investigated the social and academic constructs of motivation and how that affected academic achievement. A sample of 506 students in grades 6 and 8, participated in this study. All participation was voluntary and 92% of the population was white.

Multiple instruments were used to collect the data. They were *Motivational Strategies for Learning*, *Mastery of Goal Orientation*, and end of the year grades for English class. Wentzel found both the sixth grade and eighth grade students' social goal pursuit correlated significantly and positively with academic motivation in reading related values, reading self-efficacy, and generalized goal orientation. Pursuit of social goals also related to academic outcomes for both sixth and eighth grade students. In addition, social motivation was interrelated to academic motivation as well as performance. Academic motivation was not a predictor of students' efforts however. Wentzel concluded that if students see themselves as successful, dependable, wanting to learn new things, and get things done, they are in fact more successful.

Self-regulated behavior according to Zimmerman (2000) refers to students who are metacognitive, motivational, and behaviorally active in their learning. Learners, in other words, who have self-regulated strategies, believe they can perform efficaciously and set various and numerous goals for themselves within a social cognitive view of self-regulation. In their theory of self-determination, Deci and Ryan (1985) investigated the basic need for competence, claiming that intrinsic motivation is maintained when students feel competent in what they are doing.

In the theories of motivation through engagement, the focus has been on intrinsic and extrinsic motivation. Intrinsic motivation broadly means that students engage in an activity such as reading, out of curiosity, pursuit of interest, expressing a preference for challenging text, and demonstrating a disposition to read. Extrinsic motivation relates to engagement for students in an activity such as reading, towards the physical outcome of a reward or grades. The most highly internalized level of motivational development is

intrinsic motivation (Guthrie & Davis, 2003). At this point, the reader will engage in literacy activities for their own enjoyment, regardless of the reward or a grade.

This suggests an early adolescent reader who is engaged in their reading would be more motivated to read. In an extensive review of how instruction influences students' engagement, Guthrie and Wigfield (2000) concluded that the level of student engagement in reading influences student outcomes. Basically, to provide support for reading engagement for middle school readers, Guthrie and Wigfield (2000) suggest the use of their instructional model of engagement.

In this model, Guthrie and Wigfield (2000) suggest six characteristics of classroom instruction that influence reading engagement and motivation: (a) identify a knowledge goal of the lesson and announce it to the students; (b) provide real-world experiences related to the goal; (c) provide autonomy support to attain knowledge and learning of these goals; (d) use interesting texts for instruction that is relevant to the learning and knowledge goals being studied; (e) provide instruction of cognitive strategies that empowers students to succeed in reading these texts; and (f) provide opportunities for social collaboration of the students during teaching.

The current study provided real world experience with the use of the computer program. Autonomy with choice of songs was provided to the students in a diverse and interesting textual format.

Fluency

Biancarosa and Snow (2006) claim part of what makes teaching effective literacy strategies so difficult is the wide range of needs and experiences that present challenges for the early adolescent learner. Some readers at this level still have difficulty with

fluently reading the words accurately and with automaticity this could hamper their understanding of the various texts in the content area. Whereas, other early adolescent readers read accurately and quickly, enough for comprehension to take place however, they lack the ability to recall strategies to help them comprehend what they read. Still, others have learned the strategies but have not practiced them sufficiently. This is because they have only used them a limited amount of time, with a limited amount of different texts.

Fluency has been identified by The National Reading Panel (2000) as one of the five critical components of reading (Pikulski & Chard, 2005). As part of the NRP's review process two salient areas of fluency reading studies emerged, guided oral reading and silent reading. Guided oral reading studies included such approaches as repeated, impress, paired, shared, and assisted reading. Silent reading studies provided the student participants with time to read by him or herself.

Chall's (1996) model of reading development suggests readers go through stages in their reading, and each stage emphasizes a particular aspect of the reading process. According to this theoretical model the reader moves from: (a) early and emergent development with words, (b) through formal instruction, (c) building fluency for words, (d) then developing automaticity of word reading, and (e) finally placing emphasis on using reading to learn instead of learning to read to interpret and synthesize meaning.

This model can be interpreted as having the reader move from familiarity with the sound symbol relationship to automaticity with words to evaluate and synthesize text. However, as previously noted by Alexander (1998) in her Model of Domain Learning, she contends the early adolescent shifts in fluency in reading depending on the literacy

task. Alexander cautions that the stages of fluency are not grade or age specific, and that a reader may be competent or an expert fluent reader in one kind of literacy task however, they may drop back to acclimation (needed support) during another literacy task. This specifically occurs when there is a lack of prior knowledge, interests, of strategic processing. Topping (2006) concurs with Alexander when he contends:

Fluency is not an entity, a benchmarkable competence, or a static condition. Fluency is adaptive, context-dependent process that can operate at a number of layers or levels (this is also true of comprehension). Even expert readers will show dysfluency when confronted with a text on an unfamiliar topic that provides challenges beyond their independent reading level. Fluency is of little value in itself-it value lies in what it enables. (p. 106)

Topping suggests that there are a number of factors that interact with each other in the area of reading fluency. To demonstrate this interaction he created a model of fluency entitled, *The Deep Processing Fluency (DPF) Model* (Topping, 2006, pp. 106-129). Topping (2006) claims the relevant factors of reading fluency are arranged into four sequential sectors: (1) predisposing factors (entry skills and conditions that facilitate fluency, e.g., text difficulty, engagement, vocabulary, memory, motivation, and self-efficacy), (2) surface fluency (speed of accurate and automatic word recognition), (3) strategic fluency (control of speed of reading to yield comprehension and expression at the optimal level required for specific purpose), and (4) deep fluency (control of speed of reading to maximize comprehension, expression and deep reflection for specific purposes, enhancing explicit awareness an self-regulation of these processes) (p.107).

Topping also suggests an effective method to promote reading fluency is through repeated reading. He cautions however, that all methods are relative to text difficulty for the individual students because most students are “surface fluent” or word callers at readability level that are too difficult. Some teachers according to Topping advocate having students read and reread texts below their independent reading level, or just assess reading fluency for speed and word recall (surface fluency). These practices contribute to construing reading fluency in rather a “linear way” (Topping, 2006, p. 117) however repeated reading is seen a multidimensional event.

Repeated Reading: Accuracy, Automaticity, and Prosody

Repeated reading was often seen as a way to improve word recognition, accuracy and speed for beginning readers or older struggling readers (La Berge & Samuels, 1974). In 1979, Samuels tested the theory of automatic information processing in reading. Theoretically it was assumed that if a child could read a passage with accuracy and automatic reading recall (speed) they could then concentrate on comprehending what they read in text. To test this theory, Samuels conducted a study with a group of mentally challenged beginning readers by having them read and reread short passages (150 words) a number of times until they were able to read the passage with a rate of 85 words per minute (wpm).

Initially the children would have a copy of text at their reading level and listen while the passage was read aloud and modeled with the correct pacing, pitch, tone, emphasis, and volume. Then the children would go back to their seats and practice. When they felt they were ready they came to Samuel’s and read aloud the passage. Samuel’s would time the children’s reading and chart their progress. When they were able to read

the passage with a rate of 85 wpm they could move to the next passage. Samuels found however this was very time consuming process, then in 1985 O'Shea, Sindelar, & O'Shea found in their study with third grade readers that students only needed to reread text four times to get the benefits of fluent reading. Along with the rereading O'Shea et al., also tested for reading comprehension in their study by having students retell what they remembered after they read the passage. This study helped connect reading fluency (decoding) to reading comprehension.

Fluency connection to improve comprehension for readers of all ages and abilities has been established (Dowhower, 1987; Koskien & Blum; Schreiber, 1980). The explanation according to Schriber (1980) is the lack of prosodic information in printed text specifically; the pitch, stress, volume, and tone that help listeners obtain meaning from spoken language. Schriber suggests this could be compensated through repeated reading, which imitates speech. The prosody components of reading fluency address the use of phrasing and expression (Dowhower, 1987, 1991; Schreiber, 1980, 1987, 1991; Schreiber & Read, 1980). When readers adjust appropriate volume, tone, emphasis, phrasing, and other elements when reading aloud, they are providing evidence of comprehending text. In this sense fluency, can be seen as a multifaceted event with reading comprehension as the goal.

Taylor, Wade, and Yekovich's (1985) study with 45 struggling readers and 45 of their more proficient counterparts were indistinguishable in passage recall after their rereading intervention. Two recall scores were obtained, free recall and cumulative recall that included probes and direct questions. They found that practice through rereading texts was most effective to increase recall.

O'Shea, Sindelar, & O'Shea (1985) found average third grade readers who either needed word accuracy or speed and others who needed support in comprehension each met their goals through repeated reading. Whereas, Dowhower's (1987) research found that accurate but slow readers improved both within and between passages in their comprehension when rereading, especially when rereading several different passages at their instructional level.

For years, teachers thought if students could learn to decode words accurately, they would be successful in reading printed text (Rasinski, 2004). While it is true that accuracy in a students' ability to decode words is important for fluency, as Samuels believed in the 1970's, decoding needs to be automatic. However, this is still not sufficient. Rasinski (2004) points out the need to connect accuracy and automaticity to reading prosody.

Stayter and Allington (1991) suggest that "we have failed to consider some of the broader ramifications of an emphasis on fluency, especially with older and more developed readers" (pp.143-144). In their case study with a class of seventh grade students Stayter and Allington (1991) report that fluency instruction enriched the meaning of text. This study investigated a class of 25 heterogeneous seventh graders over five days as they reread and rehearsed short dramas. Interviews were conducted after the students performed for their class. The participants all came away with a different understanding of themselves as readers. As noted by the researcher one student said:

The first time I read to know what the words are. Then I read to know what the words say and later as I read I thought about how to say the words...As I got to know the character better, I put more

feeling in my voice. (p. 145)

Texts performed orally are ideal for repeated and prosodic reading (Rasinski, 2004). McGuire (2004) contends that the “rhythms and meter of spoken language are much like the lyrical rhythms and melodies of music” (p. 1). In her autobiographical narrative about her personal struggle to overcome her reading disability she uses music as the central metaphor to format the study. Rasinski, Homan, and Biggs (in press) report that “Singing lyrics to songs is a form of reading that is nearly ideal for fluency instruction. Songs are meant to be sung (read) orally and they are meant to be sung (read) repeatedly” (p.14). This form of repeated exposure through singing as a vehicle for reading, as in the case of the current study, can build reading fluency and comprehension and can be naturally embedded within the music content classroom.

Singing

Butzlaff (2000) contends there are similar characteristics with singing and reading: (a) music text and written text involve formal written notations that are read left to right, (b) the sensitivity to phonological distinctions and word recognition require a sensitivity to pitch and tonal distinctions in both reading and singing, (c) when students learn the lyrics to songs they are engaging in reading, and (d) learning song lyrics are often repetitive, so that rereading of text occurs through singing.

Music Learning Theory is an explanation on when and how music is learned. This theory’s primary objective is the development of students’ tonal and rhythm audiation (Gordon, 1979). The term “audiation” coined by Gordon, is the process when we hear and comprehend music for which the sound is no longer there. Gordon (1979) contends that the cognitive process is the “musical equivalent to thinking in language” (pp. 5-6).

When we listen to someone speak we must retain in memory their vocal sounds long enough to recognize and give meaning to the words the sounds represent. Music is similar, when listening to music we are audiating sounds that were recently heard. In addition, based on our schema of the tonal and rhythmic conventions a person can predict what comes next (Gordon. 1979).

However, singing in the music classroom is usually performed as a whole group with one song regardless of the variety of instructional reading levels of the student body. Hall, Boone, Grashel, and Watkins (1997) suggest students should sing independently, on pitch, and with rhythm. The *Tune Into Reading* study provided opportunities for students to sing independently, supported by background music, rhythm and pitch heard through their individual headsets.

Goetze, Cooper, and Brown (1990) conducted analysis of classroom singing studies over the last 25 years and concluded methods that included individual singing opportunities and immediate visual and verbal knowledge of results were warranted to increase accuracy in singing. While most singing in the music classroom is done in groups, minimal time is spent with students singing individually, making it difficult to assist each student to develop these specific faculties. Levinowitz (1989) found that students sang songs more accurately with text than without. In the *Tune Into Reading* study students have individual texts on their computer screens and scoring mechanism is displayed to record real-time pitch accuracy.

In the meta-analysis of over 150 articles, *Computer-Based Technology and Music Teaching and Learning*, Webster (2002) investigated various studies with computers in music education, including the categories of listening, performing, and composition.

Classrooms are more dominated by technology than ever and students' skill and understanding of computers often extend beyond those of their teachers. Webster reported that use of computers in the classroom, in partnership with teachers' orchestrating the learning environment, does seem to assist in actively engaging the student, increasing motivation and intellectual stimulation. Individualized instruction facilitates aural instruction can augment the efforts of music classroom teachers and increase learning in children in a number of different areas.

In a pilot study involving 48 struggling readers in the seventh and eighth grades in a rural central west Florida middle school, Biggs, Homan, Dedrick, Minick, & Rasinski (in press) used an interactive singing software program with real time pitch tracking that teaches users to sing in tune and in rhythm was used with middle school struggling readers. The computer program, *Carry-a Tune*, was originally developed to improve singing however, it was used in this study to determine its effect on comprehension and instructional reading levels with middle school struggling readers. The 9-week intervention was conducted with 24 struggling middle school readers. All participants had failed the state reading test, Florida Comprehensive Assessment Test (FCAT). Students utilized the software program for 30 minutes, 3 times a week. Treatment students were matched with a control group of students by FCAT level, gender, grade level, reading/language arts teachers and free and reduced lunch. Leveled texts from the Qualitative Reading Inventory (Qualitative Reading Inventory, 2004) developed as Cloze passages were administered to all 48 participants and served as pretest, posttest, and follow up measures of assessing comprehension and instructional reading levels.

A two-tailed *t*-test comparing pretest and posttest scores was used to determine the statistical significance at the end of nine weeks. A 2 (Group) x 3 (Time) repeated measures ANOVA of the group grade level averages was utilized for the follow-up testing at the end of the school year. No significant differences were found between treatment and control groups' pre-test scores however, the posttest results were highly significant for the treatment group. Mean scores of the Treatment students approached a 2-year gain in their instructional reading levels.

Current Study and Effective Practice

Tune into Reading (TIR) (Electronic Learning Products, 2006) is an interactive sing-to -read software program that can be used in the music classroom. This technological format provides diverse and interesting texts. Over two hundred songs are included on the TIR program. All songs were analyzed for readability level. The songs range from first to tenth grade level from traditional folk songs (e.g., *Amazing Grace*) to more recent pop songs (e.g., *Ain't No Mountain High Enough*).

Direct explicit comprehension instruction through repeated reading is modeled through singing. The music teacher modeled steps of effective singing by initially showing students how to get their individual vocal range (e.g., alto, soprano). Then the students proceeded with recording their individual vocal range. Once this is accomplished, all of the songs that the student sang matched their individual vocal range. Each student has an individual soundproof microphone headset for listening, singing, and recording while at their computer. The computer program has two different text formats. The first format, linear sheet music, allows the students to read the lyrics silently three times, while listening to the background music and tempo. In this way, repeated reading

is embedded into the singing program. This aligns with the recommended number of repetitions suggested by Samuels (1979). This is followed by a graphic textual view. This alternative text format provides a visual display of words broken into syllables without the accompanying musical staff and places each syllable accented at the appropriate pitch within each students' personal vocal range.

Along with the visual tracking of the words, a guideline is provided for accurate pitch and tone that provides a real time track line of the student's voice while they are singing and recording a song. After singing each time, a score is provided to the student. These scores, ranging from 0-100 represent accuracy of pitch and tone. The teacher uses these scores to determine when to change the level of songs. The students in this study sang and recorded the songs using the visual graphic format three times aloud, and saved the recorded version of their highest score. Strategy instruction with diverse texts through a technological format embedded in the content area of music led to engagement and motivation for the learners.

Summary

This review of the research clarified why gaps exist in the literature pertaining to the early adolescent and their literacy learning needs. Historically as noted by researchers (Beane, 2001; Brough, 1995; Cuban, 1992; Spring, 1986; Van Til, Vars, & Lounsbury, 1961) these learners have been caught in the tensions of whether the middle school should be more like the elementary school or like the high school. These tensions have also carried over to understanding the uniqueness of this population of learners, the ambiguity of the role of the middle school teacher, and the delivery of instruction, specifically reading in the content area.

Not unlike their historical predecessors there are current complexities for adolescents that are politically, socially, and academically influenced. The current dilemma of accountability and evaluation through high-stakes testing has compromised what has been learned to date about the complexity of the early adolescent literacy learner. It has also detoured effective practice of literacy embedded in the content classroom, by not addressing the unique needs for this population, especially when more literacy needs are needed to meet the challenges for the new millennium. How do we prepare students to be fluent active independent readers and comprehenders? Biancossa and Snow (2006) suggest the early adolescent literacy learner needs explicit direct literacy instruction, which is embedded in the content classroom to build comprehension. This can be achieved through the use of diverse and interesting texts that are accessible at the reading level of the student. Delivery of these texts could be through a technological format, which can be motivating and engaging for the adolescent. This study took place in the music content classroom, where singing instruction is taught using explicit instructions in rereading text to build comprehension. The current study will add to the body of knowledge on the early adolescent strategic processes and the need to provide literacy instruction in the content areas to these students of varying reading abilities.

CHAPTER THREE: METHOD

Chapter Three presented the methods used to conduct this study. This chapter contained five sections. The first section revealed the purpose of the study and outlines the research questions. The second section described the design of the study, the research context, and the participants. The third section presented the study's ethical considerations, instruments, measures taken to ensure reliability of the data, researchers' pre-study involvement, and the procedures. The fourth section provided specific details concerning data collection. The final section explained the manner in which data were analyzed and interpreted.

The Purpose of the Study and Research Questions

The purpose of this concurrent mixed methods study was to investigate the use of an interactive sing-to-read program *Tune Into Reading* (Electronic Learning Products, 2006) as an alternative text, embedded within a heterogeneous music classroom. Measured by the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006), the fluency, word recognition, comprehension, and instructional reading level of the treatment students were compared to their counterparts who sang as part of the regular music program. This investigation also provided a description of the peers' interactions during the literacy task assigned by the music teacher. The intent of this study was to address the following research questions:

Quantitative Research Questions

1. To what extent is the reading performance of word recognition, fluency, comprehension, and instructional reading level, as measured by the QRI-4, of students using the *Tune Into Reading* program, different from their regular music curriculum counterparts?
2. To what extent does the *Tune Into Reading* program differently impact the reading scores of students who are “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) reading scores?

Qualitative Reading Question

1. How do middle school readers interact with their peers, within the context of their music classroom?

The first quantitative research question addressed the readers’ use of the interactive sing-to-read program *Tune Into Reading* as an alternative text, and then was compared to their counterparts who are singing as part of the regular music program.. Prior to the treatment, I administered a pretest using the QRI-4. Scores from the pretest ensured that the students in the regular music class and the class using *Tune Into Reading* were not different in their performance in fluency (measured by words per minute), word recognition (measured by oral reading accuracy), comprehension (measured by implicit and explicit questions after the reading), and instructional reading level (measured by combining scores from word recognition and comprehension questions) before implementation. After the implementation of the interactive sing- to- read program, *Tune Into Reading*, I administered a posttest using the QRI-4 and compared the posttest scores with the pretest scores to determine if students in the experimental group gained

significantly over their counterparts in the control group. The students were initially assessed at posttest with a reading passage on the same instructional level attained during the pretest. The students were next assessed at posttest at the highest instructional reading level they attained.

The second quantitative research question investigated whether an interaction effect of the repeated reading methods occurred on the reading performance of the students “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) 2006 in reading, while using the sing-to-read program, *Tune Into Reading*, as an alternative text. The results in reading achievement level scores (achievement levels 1-5), according to the state of Florida Department of Education, are reported as: (a) students who scored a Level 1 or 2 are considered below proficiency in meeting grade level benchmarks, (b) students who scored a Level 3 are considered at grade level, and (c) students who scored at a Level 4 or 5 are considered above grade level (*FCAT Briefing Book*, 2005).

Concurrently, the qualitative observations were used to probe for significant themes by describing aspects of peer interactions (peer talk, peer modeling, and peer social reinforcement) among students who sang using the interactive program *Tune Into Reading*, versus the peer interactions among students who were singing in the traditional music class.

Design of the Study

In order to address the research questions, I used a mixed methods approach. The purpose of this approach was to collect, analyze, and mix or integrate both quantitative and qualitative data during the research process within a single study (Creswell, 2003;

Tashakkori & Teddlie, 2003). Both types of data were used because neither quantitative nor qualitative methods in isolation sufficiently capture the trends and details of situations, such as the complex issues of how the use of an alternative text supports literacy learning of the early adolescent and how these adolescents interact with their peers during the literacy task. When used in combination, quantitative and qualitative methods complement each other and provide a more complete picture of the research problem (Green, Caracelli, & Graham, 1989; Johnson & Turner, 2003; Tashakkori & Teddlie, 2003).

This study used a concurrent mixed methods design consisting of two distinct phases (Creswell, Plano Clark, Guttman, & Hanson, 2003; Tashakkori & Teddlie, 2003). The quantitative numeric data and qualitative text data were collected and analyzed concurrently. Integration of the data occurred during the interpretation of the study's findings. This interpretation can either note the convergence of the findings as a way to strengthen the knowledge claims of the study or explain any lack of convergence that may result (Creswell, 2003).

Quantitative Phase

The first two questions were answered utilizing a quasi-experimental design. The statistical technique that was used to answer the first question was analysis of variance (ANOVA) with repeated measures to assess differences in mean trend lines over time between the experimental and control group. Multivariate repeated measures ANOVA was conducted to assess the collective differences on the dependent variables overtime and by group (Stevens, 2002). The multivariate repeated measures ANOVA assessed if the combination of noncommensurate dependent variables differed over time and by

group. Simultaneous differences from pretests to posttest by group were further analyzed by conducting *t*-tests and determining effect sizes.

The independent variable for the first question was the literacy approach consisting of two levels: the early adolescents who use the alternative text *Tune Into Reading* during the literacy tasks and those who are part of the regular music program (treatment and control). The dependent variables were the scores from the QRI-4 on: (a) fluency- timed and measured by words per minute, (b) word recognition- measured the percentage of accuracy during the oral reading of the passages, (c) comprehension- measured by the percentage of correct response to questions asked, and (d) instructional reading level assigned a grade level (e.g., 6th) measured by the combination of scores on word recognition and comprehension at two points in time (pretest and posttest). Initially, the students were assessed at posttest with a reading passage on the same instructional level attained during the pretest. The students were next assessed at posttest on highest instructional reading level they attained.

The second question also addressed the students using the interactive sing-to-read *Tune Into Reading* program and those singing in their regular music class. The purpose was to investigate whether the repeated reading method with the sing-to-read alternative text program had a different effect on the performance of students who scored below, at or above in their reading level, as determined by their FCAT level scores.

Repeated measures ANOVAs were used to answer this question. It assessed differences in mean trend lines over time for the experimental and control groups classified according to below, at, or above grade level in FCAT reading scores. The dependent variables remained the same (pretest and posttest scores from the QRI-4).

However, the independent variables were the students in the two literacy approaches grouped by their 2006 FCAT level scores in reading.

Qualitative Phase

The qualitative phase in this study used an interpretive case study approach, with the data collection occurring through participant observation. Inductive analyses were conducted to identify conceptual themes or patterns in the data, and create categories needed (Bogdan & Biklen, 2003; Cohen, Manion, & Morrison, 2000; Merriam, 2001). These themes and categories were analyzed to identify subcategories, which helped to describe peer interactions (e.g., talk, peer modeling, peer reinforcement) during the literacy task (rereading through singing) assigned by their music teacher.

This was also considered a bounded case study because it had a defined time, a distinct social interaction focus, and a physical boundary (Stake, 1998). The case study was bounded in the context of one literacy task rereading through singing, for participants who used the *Tune Into Reading program* and those who were in the regular singing class, during the fourth quarter of the school year at the west central Florida middle school (March 26, 2007- May 25, 2007). In addition, the physical boundaries included two cases (one treatment group using the alternative text and one control group as part of the regular singing program) who were singing during their regularly schedule music class period.

This interpretive case study approach was used to describe peer interactions during the assigned literacy task. Thus, the quantitative data and results were used to provide a general picture of the research problem: whether the use of an alternative text *Tune Into Reading* supported literacy learning of early adolescents and improved their

word recognition, fluency, comprehension and instructional reading level. The qualitative data and analysis were used to describe the peer interactions during the literacy task assigned by their teacher.

Mixing the Methods

Priority was given to the quantitative approach because it looked at the statistical relationship between rereading through singing of the participants who used the sing-to-read program *Tune Into Reading* and their counterparts in the regular music class.

However, concurrently qualitative case study methods were used to better understand and describe the peer interactions occurring during the literacy task assigned by their teacher.

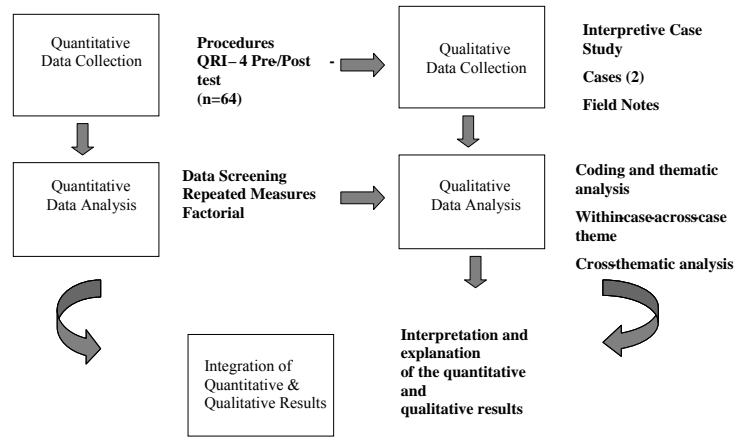
The integration of the two types of data might occur at several stages in the research process: the data collection, the data analysis, or the interpretation (Creswell, 2003). In

this concurrent mixed method study, the mixing of the data occurred during the qualitative findings section of the research project. The quantitative results and

qualitative descriptions were mixed applying a triangulation strategy in order to provide a clearer picture and answer the research questions. Figure 1 presents a diagram of the mixed methods concurrent design procedures in this study.

Figure 1

Diagram of the Mixed Methods Concurrent Design and Procedures



Research Context

The School Site

This study was conducted in a rural west central Florida middle school. This public middle school had 1079 students enrolled, and served grades sixth through eighth, (*School Improvement Plan, 2006*). The school term starts in August and extends through the end of May. The terms are divided into four quarterly reporting periods. I chose the school site because I had established a rapport with the principal and teachers prior to the study. In addition, the music teacher and I worked on previous research projects that have investigated literacy that is embedded in her music classroom.

The staff included 89 full-time teachers and 3 administrators. Ethnically, 95% of the staff were Caucasian, 1% were African American, and 4% were Hispanic. In addition, 72% were female and 28% were male. The school had one reading coach and 15 reading and/language arts teachers and support staff. All reading and language arts teachers were reading and ESOL (English as a Second Language) endorsed, through University of

South Florida (USF) partnership professional development courses, district professional development courses, or department of education state programs. In addition, all content teachers had professional development related to reading in the content areas. The ethnicity of the students is reported in Table 2.

Table 2

Percentage Enrollment By Ethnicity At The School Site

	Total	White	African American	Hispanic	Asian	Multiracial
Enrollment	1079	77%	7%	11%	1%	4%

This middle school is the cluster site for the district’s Exceptional Student Education program (ESE). It serves 240 ESE students (22%) with significant cognitive, behavioral and /or physical disabilities from around the district. Specifically, this population of students all have an active Individualized Educational Plan (IEP), and the students require direct and extensive instruction to acquire, maintain, generalize and transfer skills. In addition, students with significant cognitive disabilities are students whose cognitive abilities are 2.0 standard deviations or more below the mean of their grade level peers (Florida Department of Education, 2006). Additionally, less than 1% of the population is designated as qualifying for the ESOL program. The free and reduced lunch program benefits 51% of the student population at this school site. This qualifies the school as a Title 1 school, which receives funding from the state and national level to assist in providing remediation for struggling students based on the percentage of free and reduced lunch.

The achievement levels from the 2006 FCAT results in reading for student's in grades 6 through 8 at this middle school were reported as: (a) 47% were below grade level, (b) 35% were at grade level, and (c) 18% were above grade level. The middle school has not made Annual Yearly Progress (AYP) in reading for four consecutive years, however, they made a grade of A in Florida's A++ program. A needs assessment was completed to address the issues and review reading for the students. Based on the results of the needs assessment the school has taken several steps to improve reading. They have: (a) increased the reading remediation staff, (b) provided after-school tutoring, (c) continued with the Accelerated Reader Program through the purchase of more books at more levels, (d) increased student access to FCAT Explorer and supplemental technology tools, and (e) worked with the reading coach and professional development partnered from USF to support teachers in reading and reading across the content areas.

School instruction is provided through interdisciplinary team teaching by grade level. There are eight teams (two at each grade level) made-up of teachers in the core content areas (math, social studies, language arts, and science) plus one remedial math and one remedial reading teacher. The teams of students stay together as cohort groups for three years. Elective classes (Art, Music, Computers, and Consumer Education) are assigned to the students at the beginning of the school year, mixing grade levels across teams. Students are assigned an elective class per quarter (Wheel Class), so they have an opportunity for each of the four elective classes per year.

All students have a heterogeneously grouped language arts period daily for 90 minutes. However, students who require remediation (globally defined as those students who failed the reading portion of the FCAT 1 and 2) receive 45 minutes of reading

support during this 90 minute period. The first type of reading class includes students who performed at the lowest level on the FCAT: Level 1 (“Intensive” reading course). The second type of reading class includes students who performed at the second-to-lowest level on the FCAT: Level 2 (“Corrective” reading course). Students at Level 3, 4, and 5 on the FCAT have classes in reading, and they use the FCAT Explorer in reading and Accelerated Reader with leveled texts.

The FCAT Explorer is a free online educational program for Florida students, which provides FCAT sample questions in reading and math, related to the Sunshine State Standards (*FCAT Briefing Book*, 2005). The Accelerated Reader program (Renaissance Learning, Inc., 2006) is an individual computer program using multiple grade level texts of different readability levels. The students read the books at their instructional reading level and take a computer test, and then the teacher receives the print-out of the results. In addition to the language arts teachers’ literacy instruction, all content area teachers at this school site are required to incorporate literacy strategies in their lessons daily and must provide literacy objectives in their lesson plans (Improvement Plan, 2006).

The Music Classroom

In order to facilitate the visualization of the enactment of the literacy task (rereading through singing) a description of the physical configuration of the music class is provided. However, one must first enter the base of a single story rectangular shaped middle school to find the music classroom. The base of the rectangle houses the decision making and policy enforcement center of the school, containing the front desk and the different layers of administrative offices. Once allowed in the school, in an attempt to

reach the music classroom and see the inner working of the school and its physical layout, you must exit the administration building through a door parallel to the front desk. Outside the administration building the sound of music fills the air as different musical genres echo throughout the outdoor gardens. Picnic tables and benches, as well as a bird aviary, an alligator pen, and the school's Holocaust Memorial, are scattered around the center of the rectangle. All buildings at this school have outside access and to enter any of the buildings in the school you must follow a covered pathway that outlines the parameter of this rectangle. The right side of the rectangle contains the interdisciplinary grade level classrooms sequentially organized from 6th through 8th grade. The gym is located at the top of the rectangle and the service buildings, housing, the guidance center, the media center, the cafeteria, and the music wing are at the left side of the rectangle.

The music wing is located in the cafeteria building and runs parallel lengthwise to the cafeteria. The long hallway wall in this musical wing displays hand painted music notes, messages to the students, and different characters singing and playing instruments. The music wing is comprised of two large sound-proof classrooms: (a) the band room is first, and then (b) the chorus room, where this study took place, is second. When you open the door to enter the chorus classroom, the rhythms and sounds escape temporarily through the soundproof door.

The walls of this large room are print rich, covered with songs, musical notes, and schedules, extending all the way up the walls to the 100 foot ceiling. The left wall of the classroom is a hand painted musical scale displaying symbolic music notation and corresponding words. The right wall has an overhead screen projecting song lyrics from the projector. The white board at the front of the classroom has notes to the students and

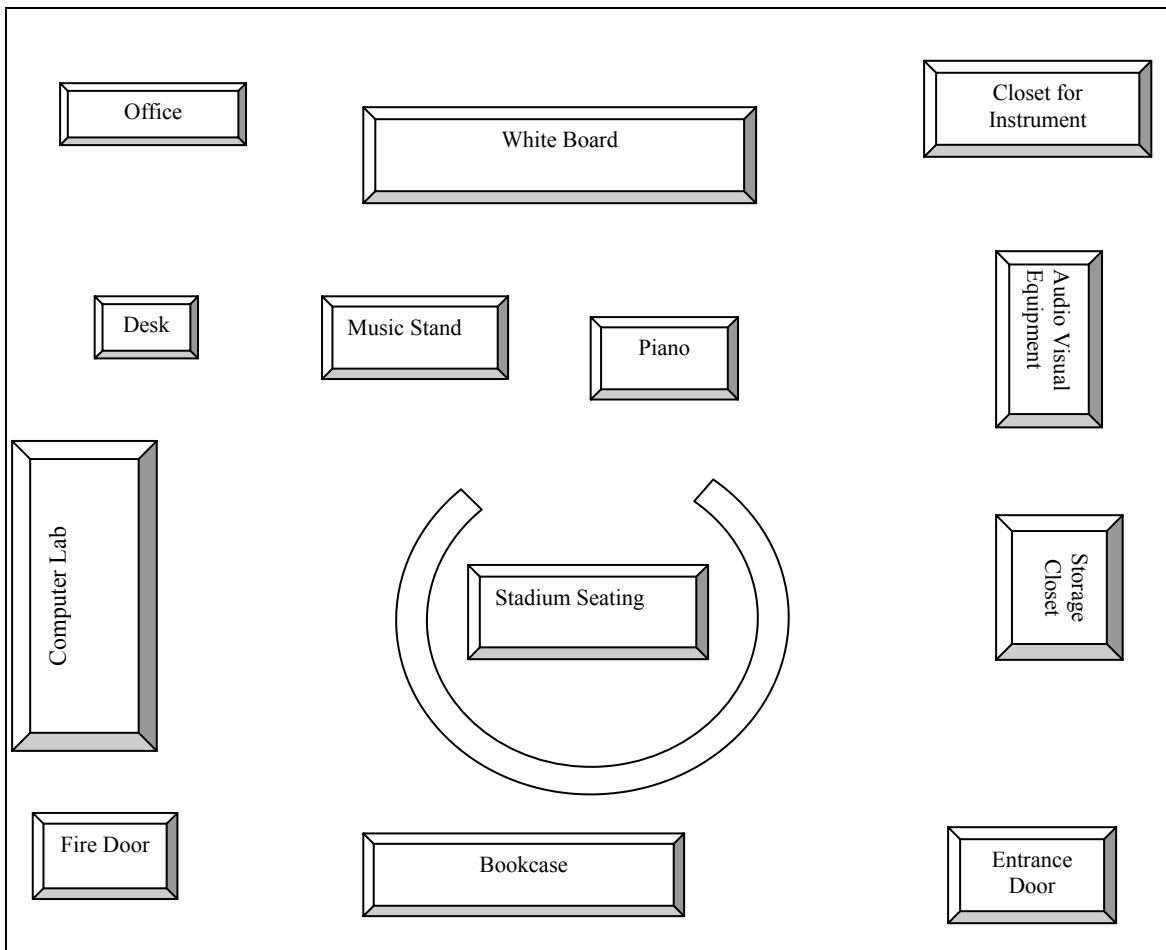
outlines the daily agenda, or assignments to be completed. The top of the back wall is the daily schedule and times for each class.

A painted bookshelf landscapes the back wall of the classroom and is autographed with handprints and names of the artists who created it. The classroom seats students in 3 semi-circle stadium steps that descend to the central stage of the classroom. To enter the stage, one could use the stairs or the ramp, wide enough for a wheelchair and hand rails to support balance. The performance areas' focal point is the piano surrounded by a garden of musical instruments: drums of all sizes, (both handmade and store bought), guitar, auto-harp, and a variety of different rhythm sticks. A music stand and a large African drum, which begs to be hit, are the standing position for the actors that enter the stage.

Before the walk to the stage, a soundproof computer lab housing 15 computers can be seen as you peer through the two-way glass window. Audio visual equipment hides in different corners of the classroom. A large television set, with a VCR and CD disk player, rests near the white board at the front of the classroom. A table outside the teacher's office holds the tape recorder and CD player, while the overhead projector gets pulled out daily and then neatly tucked next to a bookcase with song books and clipboards. The bell rings and the students rush into the classroom, select their instrument of choice, and sit down. The teacher enters the center stage and brings the group to attention with the beat of the drum: the children echo this beat and class begins. Figure 2 provides a floor plan of the music classroom.

Figure 2

Floor Plan of the Music Classroom



Participants

Sample Design

A total of 64 students from one rural west central Florida middle school music classroom participated in this study. A music classroom was chosen because it was appropriate to investigate singing as a method of rereading to build fluency embedded within the natural context of a chorus classroom. Biancarosa and Snow (2006) suggested when “instructional principles of literacy are embedded in content subject-areas, teachers

provide or reinforce instruction in the skills and strategies that are particularly effective in their subject area” (p. 24).

The sampling choice for this study was that of convenience. All of the study participants volunteered, were from the same school site, attended the same music class, and had the same music teacher. Although convenience sampling choice limits the generalizability of the findings to a larger population, this decision: (a) was consistent with the purpose of this study, and (b) is supported through the literature on technology and reading for middle school students.

As previously noted, the purpose of this study was to provide a description of the phenomena, rereading through singing using *Tune Into Reading* (Electronic Learning Products, 2006) program as an alternative text. This concurrent mixed methods study investigated the use of an interactive sing-to-read program embedded within a heterogeneous music classroom. Quantitatively, as measured by the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006), the fluency, word recognition, comprehension, and instructional reading level scores of the treatment students were compared to those of their counterparts who sang as part of the regular music program. Individual assessment of this sample provided opportunities to assess each participant, completing the full battery of the instrument while also noting and describing individual reading behaviors. Concurrently, qualitative observations were used to describe aspects of peer interactions (peer talk, peer modeling, and peer social reinforcement) among students who sang using the interactive program *Tune Into Reading*, versus the peer interactions among students who sang in the traditional music class.

In addition, for the purposes of the study, the literature supported the sampling choice and the current sample size. A meta-analysis of the effects of technology and reading for middle school learners was conducted by Pearson, Ferdig, Blomeyer, and Moran (2005). The researchers were commissioned by the North Central Regional Educational Laboratory (NCREL) Center for Technology to investigate experimental and quasi-experimental studies over the last decade in literacy and technology. Pearson et al. found little experimental research for reading and technology use in the middle grades. The researchers made the following recommendations for future studies to investigating the use of literacy learning through technology for middle school students:

1. More experimental and quasi-experimental studies using some sort of correlated design (pretests used as covariates for posttest or repeated measures).
2. Balance issues of focus on control and precision for about five weeks, longer studies might have maturation effects or other confounding variables.
3. Smaller sample sizes are more manageable than larger samples. There might be a trade-off between statistical power and experimental precision, however, it may be easier for researchers to maintain a high degree of fidelity to treatment in smaller studies because of the greater manageability prospects.
4. Follow the Complementarity Principle: (a) start with a small descriptive study, then (b) conduct a formative experiment that narrow the range of relevant variable, followed by (c) carefully controlled randomized experiments, and finally (d) conduct a full scale experimental study.
5. More studies that explore the relationship between commercial products developed to address the literacy needs for middle school. Little research has investigated

commercial technology products used for improving literacy acquisition at the middle school level. (pp. 19-23)

This mixed methods study used a quasi- experimental design for the quantitative phase, and an interpretive case study design for the qualitative phase, to investigate a commercial interactive sing-to-read program, *Tune Into Reading*, with middle school students in a music classroom. This seven-week descriptive study used a smaller sample size to maintain a high degree of fidelity to treatment, and to include measurements (pretest and posttest) for both groups (treatment and control).

To qualify for inclusion in the study, students were in grades six through eight and were a part of the elective Wheel Music Class during the fourth quarter of the 2006-2007 school year (March 12, 2007- May 31, 2007). The Wheel Music Class is an assigned elective class of new cohorts (a mix of sixth through eighth grade students) each quarter of the school year. The school year is divided into four quarters starting at the end of August and running until the end of May.

There were four intact Wheel Music Classes during the fourth quarter of this school year for this music classroom. Randomly assigning each individual student in intact curriculum classes to a treatment and control was not an option in this study (e.g., teacher lesson formats, scheduling, various grade levels). Therefore, participants were randomly assigned to treatment or control conditions by classes.

Prior to assigning each of the classes to treatment or control conditions, the numbers one through four were written on a piece of paper and placed in a bowl. A non-participant teacher from the school made four quick picks, alternating assignment for treatment then control. This way each class had an equal chance of being assigned to

either treatment or control. Classes one and three were assigned to receive the experimental treatment and classes two and four were assigned to the control. The classes were then combined. Classes one and three became the treatment group, and classes two and four became the control group.

Although the treatment and control groups were randomly assigned as classes, this told us little about the characteristics of the individuals within each group. In order to answer the research questions and compare the two groups, it was necessary to match as many of the sample characteristics of the subsets as possible prior to the experimental treatment.

Sample Characteristics

Many variables contribute to understanding how and why students perform during the complex process of reading. The control of all variables that contribute to understanding the outcomes in reading performance for these two groups is not within the scope of this study: therefore, it was necessary to provide information that matches characteristics of the two groups so that they could be compared prior to the experimental treatment.

A total of 64 students ages 12 to 14 participated in this study. The treatment and control groups had 32 students each. Initially, the treatment group had 33 students, whereas, the control group had 35 students. Two students, one treatment and one control moved. In addition, one student in the control group chose not to participate in the study. The changes in the total number of participants occurred prior to the pretest or any data collection. Therefore, these three cases were dropped from the study. In addition, attendance was taken for each session (14 sessions, 2 times a week, for 7 weeks) in both

groups. The music teacher provided a make-up session time available for students in both the treatment and control group each week. A total of six students (four treatment participants and two control participants) missed one session during this seven-week period, and all six students voluntarily made up the time during a make-up session.

It was originally assumed that each of the Wheel Music Classes would have a cohort of 6th through 8th grade students in each class, because of the inter-grade level structure of the elective classes at this school. However, after randomly assigning the students to a treatment or control groups, there were no 6th graders in either subset. In the treatment group 34% of the students were 7th graders and 66% were 8th graders, whereas, the control group had 33% 7th graders and 67% 8th graders.

Gender is a crucial variable for early adolescent literacy learners. Males and females bring different discourse styles and ways of understanding literacy to the middle school classroom (e.g., Moje, 2000). In this study, the treatment group had 37% females and 63% males, whereas, the control group had 41% female students and 59% males. Along with gender, the other classification variables (ethnicity, language, exceptional learning needs, and social economic status) influence adolescents' literacy development and their understanding of what they read and how they approach reading in school (e.g., Phelps, 2005).

The ethnic background of the students was predominately White (81% in the treatment group, and 78% in the control group). African American students accounted for 6% of the treatment population and 9% for the control, and the percentage of Hispanic students was 13% for both groups. Students identified as receiving services to support their learning, Exceptional Student Education (ESE) or language needs English Language

Learners (ELL) were: (a) 6% for ESE students in both groups, and (b) 3% for ELL students in both groups. Students' of poverty, low socioeconomic status (SES), is a critical issue for reading achievement. Some researchers contend that the academic achievement gap in reading is influenced by social, familiar, and economic factors. Allington (2002) asserts we hear more about the Black/White achievement gap or the urban issues in America schools and yet the rich/poor gaps in achievement are larger. In this study 72% of the treatment group students were considered to be of low SES (determined by free or reduced lunch programs), and 28% were considered to be of high SES. Whereas, 75% of the control group were low SES, and 25 % were considered high SES. Table 3 presents the percentages of classification variables for the students in the treatment and control groups.

Table 3

Students' Classification Variables Percentages by Treatment and Control

Group	Gender		Grade Level		Ethnicity			ESE	ELL	SES	
	Male	Female	7	8	White-	Black-	Hispanic			Low	High
Treatment (n=32)	63%	37%	34%	66%	81%	6%	13%	6%	3%	72%	28%
Control (n=32)	59%	41%	32%	68%	78%	9%	13%	6%	3%	75%	25%

A cursory examination of Table 3 of the percentages comparing the classification characteristics of both treatment and control groups, appear to suggest that that the groups are predominantly White low SES students. Male 8th graders represent a larger proportion for treatment and control groups than their female counterparts, or 7th grader peers. In addition only a small percent of the adolescents receive support services for learning or language needs. However, it is important to assess if there are any significant differences

between the sample characteristics of the two groups. Therefore, Chi-square tests at an alpha level of .05 were used to analyzed differences in gender (male and female), grade level (7th and 8th grades), ethnicity (White and Black), and SES (low and high) for the treatment and control groups. The results indicated that the proportions of classification characteristics do not differ significantly across groups, reported as: (a) gender, $\chi^2(1) = 0.0656, p = .7978$, (b) grade level, $\chi^2(1) = 0.0709, p = .7901$, (c) ethnicity, $\chi^2(1) = 0.2196, p = .6393$, and (d) SES, $\chi^2(1) = 0.0801, p = .7772$. These results verify that the treatment and control groups displayed homogeneity in the proportions of the classification variables, of gender, grade level, ethnicity, and socioeconomic status.

The matched characteristics of the sample groups provide useful information however; it did not address the research questions or provide needed information, about comparing the characteristics of reading performance for each of the groups. Prior to conducting pretests for both the treatment and control groups, each group was stratified by FCAT level reading scores.

The primary purpose of the FCAT in reading is to assess student achievement of higher order thinking skills (Florida Department of Education, 2005). FCAT level reading scores range from highest score (level 5) to lowest score (level 1). The scores for the treatment and control groups were stratified according to their FCAT level as: (a) Level 4 and 5 above grade level, (b) Level 3 at grade level, and (c) Level 1 and 2 below grade level. When this was accomplished a percentage was noted for each level by treatment and control groups. Table 4 displays the percentages by groups stratified by FCAT level reading scores as, above level, at level, and below level.

Table 4

Group Percentages of Students FCAT Level Scores

Group	Above Grade Level	At Grade Level	Below Grade Level
Treatment (n=28)	29%	42%	29%
Control (n=28)	29%	42%	29%

*Note each group was missing FCAT scores for some members: Treatment Group (-4) and Control Group (-4)

The percentages showed an equal distribution of FCAT level reading scores between the two groups, however, a concern was the missing reading scores for some of the participants. In the treatment group four students did not have FCAT level reading scores, whereas, in the control group four students did not have reading scores, also. A goal of this study was to understand and compare students of varying reading ability during the literacy task of rereading through singing. Consequently, it can not be assumed that they are compatible groups based on missing data, which could highly influence their scores in reading. In addition, reading is a very complex process. Using FCAT reading level scores alone does not provide sufficient information about the reader. As noted in the literature review, Amrein and Berliner (2002) overall contend that “ there is no compelling evidence from a set of states with high-stakes testing polices that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators” (p.54). Therefore, the use of a high-stake test scores alone can not account for the many variables associated with understanding the reading process and relating that to the characteristics of this group of early adolescent literacy learners and their fluent reading behaviors. Accordingly, it was necessary to conduct an analysis using reading pretest scores for both the treatment and the control groups.

Four Analyses of Variance (ANOVAs) at a .05 alpha level, were conducted to compare scores from the QRI-4 pretest for the treatment and control groups in fluency (wpm), word recognition, reading comprehension, and instructional reading level. The results of the analysis revealed no statistical significance difference in pretest reading scores for the treatment or the control groups in fluency (WPM) $p = .196$, word recognition (WR) $p = .180$, reading comprehension (COMP) $p = 1.00$, or instructional reading level (RL) $p = .720$. Table 5 provides a summary the descriptive statistics for the treatment and the control groups' QRI-4 pretest.

Table 5

Summary of the Descriptive Statistics of the Qualitative Reading Inventory Pretest Scores

	Treatment (n=32)				Control (n=32)			
	Mean	SD	Skewness	Kurtosis	Mean	SD	Skewness	Kurtosis
Fluency	125	32.9	0.068	-.706	136	36.09	-0.604	-0.152
Word Recognition	0.98	0.01	-.384	-1.55	0.98	0.02	-1.84	1.83
Comprehension	0.77	0.04	2.24	3.36	0.76	0.03	1.78	6.07
Reading Level	5.45	1.17	-0.10	1.65	5.58	1.22	0.35	-0.47

In conclusion, prior to experimental treatment, the treatment and control groups displayed homogeneity in proportions of the classification variables, of gender, grade level, ethnicity, and SES. In addition, the groups were no statistically different on FCAT reading level scores. Furthermore, there were not significant differences in the pretest scores of the QRI-4 in fluency (wpm), word recognition, comprehension, and instructional reading level prior to experimental treatment.

Ethical Considerations

I considered several ethical considerations before collecting the data, during the data collection, after the data were collected, and on completion of the research project.

Prior to Data Collection:

- Permission from the school and teacher where the study occurred was obtained.
- The study was reviewed and authorized by the Institutional Review Board from the University.
- Informed consent forms were used to obtain assent from the child and consent from the parents.
- Parents of the participants were sent a letter explaining the study and the role their child would play as a participant. No names were used that identify the children or their school. I provided my telephone number if any participant had questions.
- Along with the letter and the informed consent form, all participants were informed during a meeting that the study was voluntary. The participants would let me know if they did not wish to continue or in the case of the child, the parent or teacher would advise me if the child no longer wished to be a part of the study.

During data collection

- Data were backed up regularly using coded disks.
- Security codes were in place to control access to the data.
- All data were stored in a locked file cabinet at the university.

Completion of the Project

- All the field notes and data were kept secure.
- No identifiers (names, schools) of the participants were used in any written report.

- All research material will be kept for three years.
- When the data are no longer needed it will be shredded, electronic data will be destroyed.

Instruments

Qualitative Reading Inventory-4

The Qualitative Reading Inventory-4 (QRI-4) was used at two points in time to investigate the impact of using the alternative text program *Tune Into Reading* compared to the regular music curriculum. The following is a summary of the reliability and validity of the QRI-4 scores taken from the technical development report (Leslie & Caldwell, 2006). In addition the scoring procedures used are described.

Validity and Reliability of the QRI-4 Scores

The QRI-4 is intended to determine instructional reading levels for students and for diagnostic purposes (strengths and weaknesses in their reading) to determine fluency, word recognition, and comprehension. Therefore, the crucial test properties to determine reliability and validity are consistency, construct representation, and penetration (Cross & Paris, 1987). Consistency relates to the reliability of the QRI-4, and construct representation and penetration relate to the validity of the test. The QRI-4 measures consistency of scores in three ways: inter-scorer reliability, internal consistency reliability, and alternate-form reliability.

Leslie and Caldwell (2006) wanted to investigate whether the QRI-4 was consistent across examiners, to ensure that differences in judgment did not affect the consistency of the examiners' ratings. They used three expert scorers with master's degrees in reading and scorers who did not have extensive training in the subject. The

judges scored 304 student passages for reading levels and agreed on 299 of them, for an inter-scoring reliability of .98, indicating a high degree of consistency.

The internal consistency reliability or how well the score is representative of a student's true reading comprehension was also assessed. Cronbach's (1951) alpha reliability indicated a high degree of consistency (.98) for comprehension. The standard error of measurement (SEM) should be between .00 and $.25/\sqrt{n(i)-1}$, with lower numbers being more desirable (Crocker & Algina, 1986). The SEM was found to be between .10 and .23 for each passage and grade level. The reliability increases and the SEM decreases when students complete two passages of the same type (e.g., two narrative passages), as the number of similar questions the student must answer rises.

Alternative-form reliability methods were used to determine the consistency of test results over time or conditions, in order to ensure students were placed in appropriate instructional levels. This was accomplished by having students read two similar passages (e.g., two narrative passages). The reliabilities of the instructional-level decisions were all above .80, and 75% were above or equal to .90. These methods also found that 71% to 84% of the time the same instructional level would be found on both passages, according to the comprehension scores for each passage.

Leslie and Caldwell wanted consistency in the QRI-4 ability to successfully illustrate the student's strengths and weaknesses. The QRI-4 would be considered reliable in this regard if a student performed similarly when orally reading a passage and on a word list of a comparable level of readability. Two examiners independently scored 108 students to determine their level of word recognition and comprehension. The scorers agreed on the diagnostic category for the abilities of the students 87% of the time. When

the judges did not agree, it was generally when the student's patterns of strengths and weaknesses were unclear.

Another concern was that the QRI-4 should be sensitive, or responsive, to both immediate and long-term changes in students' abilities. Leslie and Caldwell examined the changes in students' reading abilities by assessing them over a four-month period. They found the QRI-4 could successfully measure change in word recognition and comprehension over this short time period. Longitudinal studies were also completed, over both the course of one school year, and over several school years. Researchers found that the QRI-4 was also sensitive to changes in abilities over a longer period of time.

Content-validity evidence speaks to the extent to which the sample of items on a test is representative of some defined domain or content (Ary, Jacobs, & Razavieh, 1996). Researchers evaluating the QRI-4 wanted to represent the field of reading in a systematic manner that reflected research findings as well as classroom practice. To accomplish this approach, researchers included both narrative and expository passages for a wide range of levels, from pre-primer to high school. The passages at the beginning levels include pictures so they represent age-appropriate materials children generally encounter.

However, reading research shows the importance of prior knowledge when reading and the significance of miscues in oral reading that alter the meaning of the passage as compared with miscues that do not (Snow, 2004). To provide for these findings, researchers included a measure of prior knowledge in the QRI-4 and two ways to score oral reading accuracy.

Researchers also provided three ways of measuring comprehension, which include the use of explicit questions, the use of implicit questions, and retelling. Word lists contained words that could be figured out using the rules of the English language and words that could not be figured out because the spellings were irregular. The QRI-4 also provides a way to evaluate a student's oral reading fluency by measuring the student's correct words per minute when reading aloud. Researchers included all of these factors in order to create a valid test that fully covers the domain of reading.

Criterion-related validity was measured by comparing students' instructional level based on the QRI-4 with students' equivalent scores on standardized reading tests, including the California Achievement Test, the Iowa Test of Basic Skills, and Terra Nova tests. The researchers examined the correlation (within grade) between the instructional level obtained from the QRI and the student's national curve equivalent (NCE) or standard score on a group administered standardized reading test.

The standardized test data for grades 1 through 3 were obtained from the California Achievement Test or the Iowa Test of Basic Skills. The standardized test data from grades 4 through 9 were the Terra Nova Test. Statistically significant correlations were found between the instructional level in narrative texts and standardized tests scores for all grade levels. Table 6 displays the correlation between the instructional level obtained from the QRI and the students' scores on the various standardized tests.

Table 6

Correlation of Instructional Level Qualitative Reading Inventory Scores and Standardized Tests Scores by Grade Level

Sample	Total	Grade Level	Correlations	Standardized Test
n=50	206	1	.85	California Achievement/
n=32		2	.65	Iowa Test of Basic Skills
n=39		3	.55	Grades 1-3
n=31		4	.66	
n=35		5	.44	
n=21		6	.27	Terra Nova
n=17		7	.43	Grades 4-9
n=22		8	.47	
n=19		9	.52	

Leslie and Caldwell assessed construct validity by determining whether the QRI-4 successfully measured word-recognition ability and comprehension. Expectations were that word identification, oral reading accuracy and reading rate would be strongly related to comprehension when dealing with beginning readers, while prior knowledge of concepts in the passage would be connected with comprehension with more advanced readers with a higher level of word recognition. Researchers found word identification from word lists; oral reading accuracy, semantically acceptable accuracy rate, rate of reading, and corrected rate were positively correlated and statistically significant through the 3rd grade from .34 to .59. Statistically significant correlations between prior knowledge and comprehension existed from the primer level and above, but correlations were much stronger above the 3rd grade. The correlations from the primer level to the 2nd grade level ranged from .18 to .30, while the correlations from 3rd grade to middle school ranged from .35 to .86.

The researchers also found students' comprehension at beginning reading levels was best predicted by the percentage of miscues that do not change the meaning of the passage and whether they read narrative or expository passages. At more advanced reading levels, researchers found comprehension was most successfully predicted by the reader's background knowledge of the concept being presented.

Qualitative Reading Inventory Adminstrating and Scoring Procedures

The QRI-4 is an informal reading inventory that provides grade level word lists, and narrative (literature) and expository (science, social studies, historical) passages for pre-primary through high-school reading levels. The choice of using narrative passages for the participants at pretest and posttest in this study came as a result of reviewing the technical report which provided support for the validity and reliability of only the narrative genre.

All passages in the QRI-4 are assigned ordinal numbers corresponding to readability levels (e.g., 1st grade reading level). However, that was not the case for upper middle school (7th and 8th grades) and high school (9th and 10th grades). They are labeled as upper middle school and high school with no corresponding readability levels. Instructional reading level is a dependent variable in this study and therefore it was important to determine the readability levels for the middle and high school narrative passages.

A Fry (1979) readability analysis co-scored with another literacy expert and approved by a university literacy professor was calculated resulting in a readability level of 7.5 (7th grade 5th month) for the upper middle school passages, and 9.5 (9th grade 5th month) for the high school passages. Therefore, when calculating the instructional

reading levels for all participants in the study, all reading levels were extrapolated (e.g., 6.0 sixth grade zero months) so that all the scores could be commensurable.

The purpose of this instrument according to Leslie and Caldwell (2006) is to determine: (a) timed automaticity of words in context (fluency), (b) accuracy of oral reading (word recognition), (c) the level of understanding in reading by answering explicit and implicit question (comprehension), and then (d) a reading level by combining word recognition and comprehension level scores (instructional reading levels). Leslie and Caldwell contend that unlike other reading inventories this instrument has extensive piloting with approximately 1,000 students at multiple grade levels.

The administration of this assessment begins by determining the appropriate grade level passage for the individual students. The authors recommend that using either the graded word list provided in this assessment or any extant data, which approximates their reading level. In this study the FCAT reading level scores were used to approximate the appropriate the beginning reading levels for assessment for two reasons: (a) it addresses the second research question of this study concerning the comparison of the relationship with reading performance and FCAT levels, and (b) the primary purpose of the FCAT in reading is to assess student achievement of the higher- order thinking skills (Florida Department of Education, 2005). Therefore it was assumed that a student who attained a higher FCAT level score in reading (e.g., level 4) would be above grade-level peers in reading. FCAT level reading scores (level 1-5) ranging from highest score (level 5) to lowest score (level 1) were used to determine the grade-level passage to start with the students. For that reason a student in grade seven who scored at a level 2 in his or her FCAT level reading score first passage would start with a 6th grade reading level.

Students however, were given as many passages as necessary until they reached frustration in order to determine their highest instructional reading level (described below).

Scores from the reading instrument are calculated for the separate components of reading as (a) fluency measured by the rate the student reads the words per minute, (b) word recognition measured by oral reading accuracy, (c) comprehension measured by implicit and explicit questions after the reading, and (d) instructional reading level determined by combining level scores from word recognition and comprehension questions. The following describes the scoring procedure for each of the components.

The administrator goes over the procedures for the assessment with the student. The student and administrator both have a copy of the passage however, only the administrator has a copy of the comprehension questions. Reading rate is calculated to determine automaticity in fluency. The administrator uses a timer with a second hand noting the student's start and end times on the assessment. To obtain the reading rate in words per minute the following formula is used: $WPM = (\text{number of words in the passage} \times 60) / \text{divided by the number of seconds it took the students to read the passage}$.

Word recognition is measured by the number of miscues in the student's oral reading. Miscues are mistakes the student makes by substituting, omitting, or inserting words, or if the administrator tells the student a word because he or she does not know it. The administrator circles mistakes on his or her copy of the passage while the student orally reads their passage. When a student self-corrects or repeats words or phrases this is not considered an error. However, the administrator of the assessment should note the self-correction because it provides evidence of comprehending or in some cases offers

evidence of struggling with the passage. In addition, an omission of an entire line by a student is counted as one miscue because it is considered as a loss of place. At the end of each passage the administrator counts the number of miscues, and the results determine whether the performance reflects an independent, instructional, and /or a frustration level in a student's word accuracy in reading. The total accuracy for reading level performance in word accuracy is (a) independent level- reads text with 98% accuracy, (b) instructional level-reads text with 90% to 97% accuracy, and (c) frustration level – reads text below 90% accurately. A chart after each passage provides the number of miscues designated for each reading level.

To determine percentages for word recognition in reading, the administrator subtracts the number of miscues from the number of words in the passage (total words are listed at the bottom of each passage). This yields the number of words read correctly. Then the administrator divides the number of words read correctly by the number of total passage words, rounding up to find the percentage of total accuracy.

Comprehension is measured by the students' responses to either eight or ten implicit and explicit questions asked after the reading. Only the administrator has a copy of the questions. The questions are scored as either right or wrong, and under each question the correct responses are provided to the administrator. At the end of each passage the administrator counts the number of correct responses and the results determine whether the performance reflects an independent, instructional, and /or a frustration level in the student's comprehension in reading. The total correct responses for reading level performance in comprehension are (a) independent level- answers questions correctly 90% or above, (b) instructional level- answers questions correctly

67% - 89% and (c) frustration level – answers questions below 67%. A chart after each passage provides the number of correctly answered questions needed for each reading level. To determine percentages for comprehension in reading, the administrator divides the correct responses by the total number of questions.

Instructional reading level is determined by the combination of word recognition level plus comprehension reading level, on a particular grade level passage. Therefore, a student who reads a 6th grade passage and scores at the independent level for word recognition, and the instructional level in comprehension, would represent a 6th grade instructional reading level. Table 7 displays how the combinations of levels determine the students reading level.

Table 7

Determining Instructional Reading Levels from the Qualitative Reading Inventory

Word Recognition	+	Comprehension	=	Total Passage Level
Independent	+	Independent	=	Independent
Independent	+	Instructional	=	Instructional
Independent	+	Frustration	=	Frustration
Instructional	+	Independent	=	Instructional
Instructional	+	Instructional	=	Instructional
Instructional	+	Frustration	=	Frustration
Frustration	+	Independent	=	Instructional
Frustration	+	Instructional	=	Frustration

Leslie and Caldwell (2006) recommend that if the assessment is being used as a pretest/ posttest measure, that the posttest passage should be at the same instructional level attained during the pretest. Then the administrator continues to test the students

until they reach frustration. One level above frustration is their new instructional reading level.

Florida Comprehensive Assessment Test

The primary purpose of the FCAT is to assess student achievement of higher-order thinking skills for reading, writing, math, and science. Students take the FCAT in grades 3 through 11. In grades 4, 8, and 10 students take the writing portion of the test. In grades 5, 8, and 11 students take the science portion of the FCAT, and students in grades 3 through 10 take the reading and mathematics portions. Florida Comprehensive Assessment Test (FCAT) level 2006 scores in reading were used in this study to divide students before treatment in three groups. The students in this study were divided into groups for the purpose of data analysis based on their levels as “below, at, or above” in reading. The following is a summary of the reliability and validity of the FCAT level scores as reported by the Florida Department of Education (FCAT Briefing Book, 2005).

Reliability and Validity of the FCAT Scores

Criterion-referenced tests are designed to identify an individual’s status with respect to an established standard of performance. For the FCAT, these established standards are the Sunshine State Standards. The FCAT’s secondary purpose is to compare the performance of Florida students with students across the nation, which is accomplished by using a norm-referenced test (NRT) for reading and math. The current NRT is the Stanford Achievement Test 10 (*SAT 10*), published by *Harcourt Assessment, Incorporated*, 2005. A research based norm-reference achievement test provides information on student performance based on its nationwide standardization program conducted in the spring and fall of 2002 on the K-12 population.

The 2003 Florida legislature enacted HB 915 that required the Department of Education to determine the score relationships of the SAT, ACT, PSAT, and PLAN to the FCAT. They conducted concordance studies, a technical procedure for converting scores from one standardized test to another. The study was based on students who had taken the FCAT in spring of 2000 or 2001 and had taken one of the other four tests. They found positive correlations between FCAT scores and the scores on the other four tests, all within the range of correlations between those of the SAT and ACT. The state of Florida had by far the strongest correlations, with a .96 correlation between high and low stakes test score levels and a .71 correlation between the year-to-year gains on high and low stakes tests (Florida Department of Education, 2006).

The degree of difficulty of FCAT items is categorized in two ways – by item difficulty and cognitive complexity. Item difficulty consists of two meanings. Before testing, it is the prediction of the percentage of students who will choose the correct answer. After testing, it is the percentage of students who actually chose the correct answer. When 70% of the students chose the correct answer items are categorized as easy. When 40-70% of the students answered correctly items are considered average, and challenging questions are answered correctly by fewer than 40% of the students.

The cognitive complexity refers to the cognitive demand associated with each item. This is currently determined using a system based on Webb's (2002) work related to the Depth of Knowledge Levels. Webb developed four levels of cognitive complexity as an alignment method to examine the consistency between the cognitive demands of the standards and the cognitive demands of the assessment. Bloom's taxonomy was previously used to determine the cognitive complexity, but it was found to depend too

much on the abilities and prior knowledge of the students as opposed to the expectations of the items. Therefore, the cognitive complexity classification no longer relies on the student's approach to the question but on the actual test item itself.

After a student takes the FCAT in reading and mathematics, the student receives a developmental scale score that ranges from 0 to 3000. These scores provide additional information to help interpret scores from the FCAT Sunshine State Standards (SSS) test. Developmental scores are used because simply looking at the scale scores that the FCAT reports, which range from 100 to 500, do not reflect students' progress within a level. Students should receive higher developmental scores as they move from grade to grade according to increased achievement. Since reading and mathematics are tested every year, this score is used to help parents and schools understand students' year-to-year progress. Based on the developmental scale score, the student is then assigned one of five Achievement Level Classifications ranging from 1 to 5.

A level 5 score indicates the student has had success with the most challenging content of the SSS and has answered most of the test questions correctly, including the most challenging questions. Students who earn a level 4 score have had success with challenging content of the SSS, and have answered most of the test questions correctly, but may have had only some success with questions concerning the most challenging content. Level 5 and 4 are considered above grade level in reading. A level 3 score means that the student had partial success with the challenging content of the SSS, but their performance is inconsistent. They may have answered many of the test questions correctly, but they are generally less successful with the most challenging questions. Level 3 denotes meeting the basics for the grade level or at grade level. Students at this

level are considered on grade level in reading and mathematics. A student who earns a level 2 score has had limited success with the challenging content of the SSS. A level 1 score indicates little success with the challenging content of the SSS. Both Level 2 and level 1 are considered below grade level and not meeting grade level expectations.

Reliability of the Data

The following section reports how I ensured the information in the concurrent mixed method study was reliable. The quantitative phase addressed the measure taken to address reliability through measures of internal consistency and interrater reliability.

Internal Consistency Reliability

The reliability of a measuring instrument is the degree of consistency with which it measures whatever it is supposed to measure (Nunnally, 1978). One way to measure reliability involves assessing a test's internal consistency, the extent to which all test items are measuring the same thing. Cronbach's alpha is the most common estimate of internal consistency of items in a scale. Alpha measures the extent to which items responses obtained at the same time correlate highly with each other. However, when items are dichotomously scored, as in this study, as right or wrong (0 and 1) Kuder – Richardson 20 (KR20) is used to assess a test's internal consistency. Kuder and Richardson devised a procedure for estimating the reliability of a test in 1937. It has become the standard for estimating reliability for single administration of a single form. Kuder-Richardson measures inter-item consistency. It is tantamount to doing a split-half reliability on all combinations of items resulting from different splitting of the test (Sapp, 2006).

In this study, comprehension reading scores for 10 students (5 treatment and 5 control) consisting of 10 questions were labeled as right (1) or wrong (0). The alpha was computed for internal consistency on the 10 students followed by internal consistency measures for the 5 treatment and the 5 control groups separately. The raw coefficients for each of these variables were .75, .72, and .70 respectively. Nunnally (1978) suggests .70 as an acceptable reliability coefficient; smaller reliability coefficients are seen as inadequate. These numbers are considered satisfactory following Nunnally's guidelines and indicate that for these variables, the test scores in reading comprehension had an acceptable level of internal consistency.

Interrater Reliability Training and Scoring

In addition to internal consistency, another reliability issue is the consistency of scoring of test items. To measure the extent to which I accurately and reliably applied the scoring criteria from the *Qualitative Reading Inventory-4* (QRI-4) for fluency, word recognition, comprehension, and reading level, a stratified random sample of 20 students (10 treatment and 10 control), at pretest and posttest were double-scored. Prior to any work completed by the second scorer, I conducted two training sessions. The second scorer was a literacy education professional with extensive experience in reading content and pedagogy. In addition, she has for the last three years used the QRI-4 in the field with me on various research projects.

The first session explained the procedure for co-scoring with a student. Since the co-scorer was familiar with the instrument, the first session developed the procedures we followed in the field. The second session was a practice session with the procedures for co-scoring with a student which included looking at rate, measuring words per minute for

fluency, word recognition, miscues (number of mistakes made by the students), comprehension questions answered correctly, and assessment of instructional reading level based on the scores from word recognition and comprehension. After the scoring session was complete, both the scorer and I calculated the assessment independently and then discussed any differences in scores.

The scorer then went out into the field on two occasions, during pretest and posttest. Ten students were selected using a stratified random sample from the treatment and control groups. The same students, selected at pretest were co-scored during the posttests. Two Pearson correlation coefficients were calculated on two of the dependent variables of this study, fluency and word recognition to investigate the relationship of the scores between the co-scorer and researcher. The fluency scores and the word recognition scores were both highly correlated $r=.999$. The correlation results for word recognition also showed a strong relationship that was significant $r = .943$.

Procedures

The following section describes the procedures used for the treatment and control groups during the literacy task of rereading through singing. However, before a description of the procedures for both of the groups, a discussion of my pre-study involvement with the interactive sing-to-read program *Tune Into Reading* is necessary because the protocol, developed from previous research, was used in the current study.

Pre-Study Involvement Developing the Protocol for the Current Study

Over the past three years I have been involved in several quasi-experimental studies investigating the impact of the interactive sing-to-read program, *Tune Into Reading* (previously referred to and adapted from the *Carry-A-Tune* program).

Interestingly, the program was designed to improve singing; however, the developer of the interactive sing-to-read program, *Tune Into Reading* received a call from a parent of a middle school student who struggled in her reading suggesting to the developer that the use of the program improved the student's reading. The developer brought the program to The University of South Florida and asked a literacy professor if a study could be conducted on this assumption. I was assigned as a research assistant to conduct a pilot study. The purpose of the initial study and the following replication studies were to investigate the impact on reading performance measured by the *Qualitative Reading Inventory* of the students who used this singing program compared to their counterparts who did not.

A total of four hundred west central Florida struggling readers (the struggle was determined by FCAT scores levels 1 and 2) from three school districts in grades four through twelve were participants over the last three years. The initial study ($n=48$) was conducted in a middle school music classroom for 9 weeks, 3 times a week, for 30 minutes each session. During this study I developed a protocol for use with the sing-to-read program that was used in the current study. This protocol was adapted from Samuels's (1979) theoretical recommendations for building reading fluency.

The program *Tune Into Reading* uses a vocal-range analyzer that tracks the singer's pitch and rhythm, comparing it to the correct pitch of the song. Each student uses a microphoned headset linked to the computer to sing along repeatedly and to record his or her singing. Following Samuels's (1979) theoretical recommendation for building reading fluency with struggling readers, I developed a protocol for treatment using this interactive sing-to-read program.

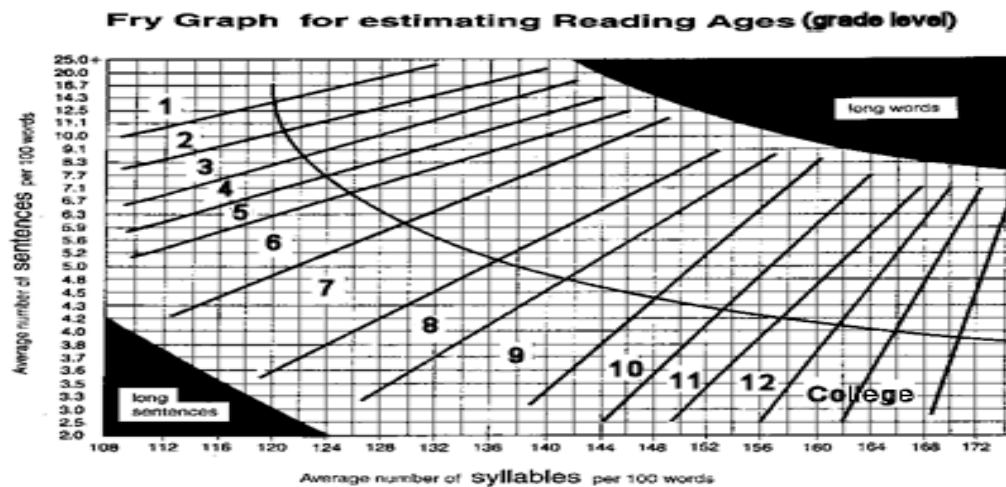
Samuels (1979) recommends: (a) students be provided a model of fluent reading, (b) reading material should be at their instructional reading level (reading with 90-94% accuracy), (c) practice rereading the material at least three times independently, and then (d) orally read the passage for assessment and feedback. Following these recommendations I adapted the reading fluency protocol to meet the needs for this study.

The first recommendation was to provide a model of fluent reading. In this study the students had background music with words (broken into syllables) emphasizing pace, pitch, volume, rhythm, and tone. This provided a model of reading fluency specifically, relating to prosody of text. Then, as recommended by Samuels (1979), the reading material used to build fluency should be at the students' instructional reading level (students can read passage with 90-95% accuracy). There were 24 songs on this program, and to determine readability for the songs a literacy professor and I co-scored all the songs. We both had individual copies of the song lyrics and independently scored the songs for readability levels using the Fry (1979) readability formula. When this was accomplished we compared each song and if there was any disagreement we discussed it and made the appropriate adjustments. In the end each song on the program had a readability grade level so that the students could sing songs at their instructional reading level.

The Fry readability formula is calculated by the averaging the number of sentences and syllables per hundred words. These averages are plotted onto a specific graph, and the intersection of the average number of sentence and average number of syllables determines the reading level. Figure 3 is a copy of the graph used for the Fry (1979) readability formula.

Figure 3

Fry Readability Formula Graph (1979)- Copyright Free



Instructional reading levels were determined from leveled passages using the *Qualitative Reading Inventory* administered to all 48 participants and served as pretest, posttest, and follow up measures of assessing comprehension and reading levels. Therefore, when the students sang their songs they used material on their instructional reading level. Once the students' instructional reading level was determined, and the songs (reading material) were at the students' correct level, Samuels recommended rereading the passage at least three times.

Prior to using the software each student enters the signs-in component of the computer. After typing in their name all data collected for the student became permanently stored into his or her personal portfolio on the computer specifically, all the students singing scores, recordings, and their individual vocal range. In order for the students' to get his or her vocal range they record themselves singing at their highest vocal level followed by their lowest vocal level, holding a single note or vowel sound (e.g., do or ah). Then the program calculates the vocal range by combining the high and

low levels. As a result all of the songs that the students sang were at the individuals' appropriate reading level.

The sing-to-read program has two different textual formats for rereading. The first text format, linear sheet music, allows the student to read the lyrics silently three times, while listening to the background music and tempo. This aligns with the recommended number of repetitions suggested by Samuels (1979). The linear sheet-music view is followed by a graphic textual view, where students record their singing. This alternative text format provides a visual display of words broken into syllables without the accompanying musical staff and places each syllable accented at the appropriate pitch within each student's personal vocal range. The graphic view of the song that is used to guide students' pitch matching while they sing selected songs.

Along with the visual tracking of the words, a guideline for accurate pitch and tone provides a real time track line of the student's voice while he or she is singing and recording a song. After singing each time, a score is provided to the student. These scores range from 0-100 on their representation of pitch accuracy and tone for the song. The students in this study sang and recorded the songs using the visual graphic format three times aloud. Then the program saved all their recorded versions of their highest score for each song. Therefore, I could review their singing and assess their progress.

How the Students Used the Tune Into Reading Program

The teacher was a veteran music teacher of 20 years. She had used the program for three years and was the same music teacher with whom the protocol was developed. This was important so that teaching can be undisturbed by trying to learn the program. Prior to starting the experimental treatment with the students the teacher reported to me

there were no participants in this study who has used the program. In addition, the students were aware of the selection process and they appeared to be comfortable with how it was handled. Following the protocol from previous studies the music teacher introduced the students to the interactive sing-to-read program. *Tune Into Reading*. However, unlike previous studies the current study was a seven-week treatment and sessions were twice a week for forty-five minutes per session.

Using an overhead projector the music teacher presented the *Tune Into Reading* program to the whole group of students. She went over all the components of the program, showing the students: (a) how to sign-in, (b) how to determine their vocal range, (c) how to use the two different textual formats, (d) how many times to listen to the song and reread silently, (e) how many times to record their singing, (f) how to interpret their scores and how this represents the accuracy of matching the pitch of the song while singing and recording, and finally (g) how to access their individual folder that contained the songs they would work with for each week. Then the student went to their individual computers and the teacher had them sign-in and record their vocal range.

The teacher walked around and made sure that the students had this in place. All of the students practiced the fluency protocol using the same song *Hot Cross Buns*. This particular song was used because it has a 2nd grade readability level. Therefore, all of the students were able to read the words of the song while they were learning how to use the program. When the students returned for the next session, they had individual songs in a folder under their name at their individual instructional reading levels. Instructional reading levels for the students were determined through their pretest scores from the QRI-4.

Control Group

The same music teacher worked with the control students for seven weeks, two sessions a week, for 45 minutes per session, not unlike their counterparts using the sing-to-read software, *Tune Into Reading*. The students during this seven-week study learned three songs with multiple stanzas, while learning and individually playing simple drum rhythms to accompany their singing. The music teacher suggested, “drum circles are a way to build a sense of community in the classroom. They keep the students motivated and engaged in the singing process.... And drumming provides a rhythmic background that supports the student while learning a song” (March 26, 2007).

Initially, the music teacher presented the simple drum patterns to the students. All of the students had individual drums, as did the teacher. She taught the rhythmic pattern and the students echoed the same pattern during the first two sessions. This was followed by teaching a song. The procedure for teaching a song went as follows:

1. The song was presented to the entire group using an overhead projector.
2. The meaning of the song was discussed along with some pertinent vocabulary words within the song.
3. The music teacher sang the song first, and then the students followed along reading the text on the projector screen.
4. The song was broken down by stanzas the teacher sang first, and then the students echoed her singing for each stanza.
5. Each stanza was sung repeatedly three times.

6. When the song in its entirety was repeatedly sung, the students played the simple drum rhythm while they sang.
7. The music teacher spent two weeks on each of the three songs.
8. The final week was a performance of the students' singing and playing the drums for a school assembly.

Data Collection

Quasi-Experimental Design Data Collection

Quantitative data collection consisted of administering the QRI-4 assessment to participants in both the alternative text *Tune Into Reading* program and the regular music curriculum program at two points in time (pretest and posttest). Prior to the experimental treatment and upon approval of the informed consent forms, groups by class were assigned randomly to the control and experimental conditions. One treatment group of 32 students used the alternative text program *Tune Into Reading*, and one control group of 32 students sang as part of their regular music program. Scores from the pretest were used to ensure that the students in the experimental treatment and control groups were not different in their performance in word recognition, fluency, comprehension, and instructional reading level before the experimental treatment.

The students were individually tested during their Wheel Music Class periods. Each Wheel Music Class period ran for 50 minutes each day, and each student took approximately 25 minutes to test during these periods. As previously discussed this study included four Wheel Music Classes that were randomly assigned to a treatment or a control condition. The four classes had different class periods each day and there were different numbers of participants in each class. A total of eight students, four from the

treatment and four from the control condition, were tested daily. The total testing time was accomplished in 10 days for all participants. All pretests for the 64 participants were completed within two weeks (March 19th- March 30th, 2007) prior to the 7- week experimental treatment (April 2nd –May 15th, 2007). Table 8 presents the schedule of pretests by class period for the treatment and control participants.

Table 8

Schedule of Pretests for Treatment and Control Participants

Class Period	Time	Treatment/ Control	Number of Participants	Duration
Period 1	7:30-8:20	Treatment Condition	12 Participants	6 days
Period 2	8:30-9:20	Control Condition	18 Participants	9 days
Period 3	9:30-10:20	Treatment Condition	20 Participants	10 Days
Period 6	1:45-2:30	Control Condition	14 Participants	7 Days

After the implementation of the interactive sing-to-read program, I administered a posttest using the QRI-4 and compared the posttest scores with the pretest scores to determine if students in the experimental group had gained significantly over their counterparts in the control group. All posttests for the treatment and the control conditions were completed after the 7-week experimental treatment following the same procedures as the pretests (May 17th- May 31st, 2007).

Interpretive Case Study Data Collection

I developed a schedule of observations for the two cases in this study based on the middle school calendar. A total of 14 classroom visits were made in the music classroom during the fourth quarter of the 2006-2007 school year (April 2 - May 15, 2007) over the seven week experimental treatment period. As previously noted, four classes were

randomly assigned by class to the treatment and control conditions. Two classes were combined and became the treatment group and two classes were combined and became the control group. Observations occurred twice a week for both the treatment and control groups in all four classes on the same day. Figure 4 depicts a schedule of qualitative observations for both the treatment and control groups.

Figure 4

Qualitative Observations Schedule

	Monday	Tuesday	Wednesday	Thursday
April 2		Observations All Classes		Observations All Classes
April 9	Observations All Classes		Observations All Classes	
April 16	Observations All Classes		Observations All Classes	
April 23			Observations All Classes	Observations All Classes
April 30		Observations All Classes		Observation All Classes
May 7		Observations All Classes		Observations All Classes
May 14	Observations All Classes	Observations All Classes		

Observational field notes were taken during each class session twice a week during the 50-minute class periods for each of the four classes assigned to the treatment or control condition. Field notes were taken on a pad of paper during the Wheel Music Class periods noting time, place, attendance, and all the peer interactions during the observation. These observations focused on describing the relationship, if any, between the literacy task the music teacher assigned (rereading through singing). Focusing on the interactions (peer talk, peer modeling, and peer social reinforcement) among students

who were singing using the interactive program *Tune Into Reading*, versus the peer interactions among students who sang in the traditional music class.

Strauss (1993) recommended that to assist with this difficult process for beginners, researchers should develop a coding paradigm. The paradigm, which applies to this study, consisted of: (1) the literacy task (rereading through singing) assigned by the music teacher and (2) interactions among the peer groups during the literacy task assigned by the music teacher for the two cases (students using the interactive sing-to-read program and students in the regular music class).

Following a theory suggested by Ryan (2000), there are generally three ways that early adolescents experience peer interactions within the context of middle school: (a) through information exchange (discussion), (b) modeling (peer observation and imitations), and (c) peer pressure (social reinforcement). I used these three categories as preliminary coding categories and as a framework to focus my observations.

Information exchange refers to discussions and talk amongst the peers, capturing direct quotes from the various conversations that the peers exchanged during the literacy task: Peer 1 “ How did you get the song to slow down” Peer 2 “ Click on this button” (Observational notes April, 7, 2007). Peer modeling on the other hand refers to the act of peers observing one another that result in changes in behaviors or understanding within the student(s). This is achieved by describing the interactions during the literacy task that documents these changes: [He looked around the classroom for two minutes then he smiled and went back to playing the drums] (Observational notes, April 7, 2007). Finally, peer pressure occurs through social reinforcement, both negative and positive. Descriptions of peers’ accepting or rejecting behaviors exhibited by their counterparts

through body language, facial expressions, smiling, or laughing during the literacy task: [T hit the drum wrong... M laughed...and then the class laughed...T turned red and put his head down] (Observational notes, April 7, 2007). Ryans' (2000) three categories became preliminary coding categories. They were then put into a matrix that was used for data analysis.

Field notes were reviewed daily after all the observations were completed. Initially, I would read through the notes three times to get a holistic sense of the data collected. Then the notes were bracketed and coded as one of the three peer interaction categories. Units of data were, conversations amongst the peers, or paragraphs that described peers observing or applying pressure to other peers, were bracketed and labeled as one of the three peer interaction categories. This was followed by transferring the bracketed notes to a matrix (Appendix B) with the three categories. The matrix was used to ensure that the observations did not stray from the focus of the study. Once the data were transferred the difficult job of data analysis began. Figure 5 provides an example of the matrix used in this study

Figure 5

An Example of the Observational Notes Transferred to the Categorical Matrix for the Peer Interactions

Information Exchange Peer discussion/talk direct quotes from conversations during the literacy task	Peer 1“ How did you get the song to slow down” Peer 2 “ Click on this button”
Modeling Peer Observation/ through descriptions of interactions during the literacy task	He looked around the classroom started to smile and went back to playing the drums
Peer Pressure Social reinforcement/ descriptions through looks / comments/ laughs during the literacy task	T hit the drum wrong, M laughed and then the class laughed T turned red and put his head down,

Data Analysis

Data analyses were concerned with the research questions and the integration of the data to meet the study’s design. The quantitative methods used for data analysis are explained first. This explanation is followed by the qualitative methods used for data analysis. The final section explains how the data were integrated.

Quasi-Experimental Design Data Analysis

The research question concerned with this phase was:

1. To what extent is the reading performance of word recognition, fluency, comprehension, and instructional reading level, as measured by the QRI-4, of students using the *Tune Into Reading* program, different from their regular music curriculum counterparts?

2. To what extent does the *Tune Into Reading* program differently impact the reading scores of students who are “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) reading scores?

Data for the quantitative phase came from the participants’ performance on the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006), reading assessments during the pretest and posttest administrations. Leslie and Caldwell recommend that if the assessment is being used as a pretest/ posttest measure, that the posttest passage should be at the same instructional level attained during the pretest. Then the administrator continues testing the students until they reach frustration, so that the new instructional reading level can be determined. Therefore, analysis for the first two questions was completed utilizing the same instructional reading level scores attained on the pretests, and then another analysis was completed at the students’ higher instructional reading level if appropriate.

All quantitative data analyses were conducted using SPSS software (Statistical Package for Social Sciences) Version 15.0 (Stevens, 2002). The analyses included computation of differences in mean performances between the experimental and control group on the QRI-4.

Question 1. The first quantitative research question addressed the readers’ literacy performance after using the interactive sing-to-read program *Tune Into Reading* as an alternative text and how this compared to the performance of their counterparts who were singing in the regular music class. Prior to the treatment, I administered a pretest using the QRI-4. Scores from the pretest were examined to ensure that the students in the regular music class and the students using *Tune Into Reading* were not different in their

performance in fluency (measured by words per minute), word recognition (measured by oral reading accuracy), comprehension (measured by implicit and explicit questions after the reading), and instructional reading level (measured by combining scores from word recognition and comprehension questions) before implementation. After the implementation of the interactive sing- to- read program, *Tune Into Reading*, I administered a posttest using the QRI-4 and compared the posttest scores with the pretest scores using their reading level scores from their pretest initially to determine if students in the experimental group gained significantly over their counterparts in the control group. This was followed by a comparison of pretest scores and posttest scores at the higher instructional reading level. Then I analyzed the scores at their higher reading level at posttest if appropriate.

Doubly multivariate repeated measures ANOVA at an alpha level of .05 was used to examine the simultaneous differences in the dependent variables fluency (WPM), word recognition (WR), comprehension (Comp), and instructional reading level (RL) on the same instructional reading level attained at the pretest initially at two points in time (pretest to posttest). The multivariate repeated measures ANOVA assessed if the combination of noncommensurate (differing measurement scales) dependent variables differ over time and by group. Before analyses were initiated, preliminary inspections of all variables were completed to check distributions (observations outside the normal distribution). Means, standard deviations, skewness, and kurtosis, were calculated for all continuous variables, and percentages for all categorical variables, were derived in order to describe the sample and be able to compare results with data from other published studies.

Simultaneous differences reported by the *F* test statistics from pretests to posttest by group were analyzed first by checking for significant interactions. If the interactions were significant, then comparisons were conducted using *t*-tests on each of the dependent variables and determining effect sizes. Initially, the scores were analyzed at the same instructional pretest reading level and then this was followed by a between-groups analysis of variance (ANOVAs) for each of the four dependent variables at the increased reading level posttest scores.

Question 2. The second quantitative research question investigated whether an interaction effect of the repeated reading methods occurred on the reading performance of the students “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) 2006 in reading, while using the sing-to-read program, *Tune Into Reading*, as an alternative text. The results in reading achievement level scores (achievement levels 1 through 5), according to the state of Florida Department of Education, are reported as (a) students who scored a Level 1 or 2 are considered below proficiency in meeting grade level benchmarks, (b) students who scored a Level 3 are considered at grade level, and (c) students who scored at a Level 4 or 5 are considered above grade level (*FCAT Briefing Book*, 2005). The students were grouped by FCAT reading level scores and then analyses were conducted on the four dependent variables for the three levels.

Repeated measures ANOVAs at an alpha level of .05 were used to examine the differences for each of the dependent variables fluency (WPM), word recognition (WR), comprehension (COMP), and instructional reading level (RL) at the same instructional pretest level for each of the three FCAT levels. The repeated measures ANOVA assessed

if each of the dependent variables differed over time, instructional group, and by FCAT Reading Levels. Before analyses were initiated, means, standard deviations, skewness, and kurtosis, were calculated for all the continuous variables.

Differences reported by the F test statistics from pretests to posttest by groups were analyzed by first checking for significant group level interaction. If the interactions were significant, then comparisons were conducted using t -tests and determining effect sizes for each of the dependent variables for the three levels at the same instructional pretest reading level. This was followed by a between-groups analysis of variance (ANOVAs) for each of the four dependent variables at the increased reading level posttest scores.

Interpretive Case Study Data Analysis

The analysis of the data required qualitative analysis procedures. Patton's (2002) guidelines for content analysis recommended reading through the data at a specific time and making notes in the margins pertaining to specific notions about meanings. Moerman's (1988) suggestions for conversation analysis guided the analysis of peer interactions through conversations. In addition, Miles and Huberman's pattern analysis (1994) was used to code data and look for emerging patterns.

The data analysis for this case study involved a careful review of data gathered from the observations of peer interactions within the treatment and control groups during the literacy task of rereading through singing. This study consisted of two cases. The experimental treatment group using the interactive sing-to-read program *Tune Into Reading* and the control group singing as part of their regular music class. Therefore, the constant comparative method was used to analyze the data (Glaser & Strauss, 1967;

Patton, 2002). Using constant comparative form of analysis, I began the process of analyzing text after each observation (Glaser & Strauss, 1967). It involved “continually comparing one unit of data with another in order to derive conceptual elements of theory” (Merriam, 2002, p.8). The comparison initially took place within each case but eventually moved across cases.

My first task involved typing the field notes from the observations of the Wheel Music Classes. The notes were typed-up daily after all the classroom observations so the information could remain fresh. Once this task was accomplished, I began the difficult task of reading and analyzing the data. First, I read the field notes from the classes through three times to gain a holistic sense of the data. Then I returned to the data and bracketed the categories of peer interactions and labeled them as information exchange, modeling, and peer pressure so that it could be transferred to the peer interaction matrix (Appendix B). I then read each line of the data in the matrix and highlighted units of meaning, patterns where repeated phrases and or words occurred (Patton, 2002).

Construct names emerged from these data. The construct names came directly from the data. One example that illustrates how this was done was from a phrase that described peer modeling, “In the four corners of the computer lab small groups of females look at one another and start to laugh softly, as they secretly glanced around the room”. This sentence was highlighted and was bracketed with the construct name, *Peer Observation*.

Once in the matrix the data were further analyzed to determine the elements of peer interactions during the literacy task. After the elements were identified and assigned construct names, they were added to the Construct Key (Appendix C). I used the Construct Key to be consistent with construct names from the emerging data, but also

added any new emerging constructs from the consecutive observations to the construct key.

The elements were then grouped according to the construct names. The elements with the construct names assigned were then cut-up and placed in a folder. The frequency of each construct was tallied to determine whether or not an element was emphasized during the peer interactions. The frequency calculations were followed by organizing the constructs into categories. Each category of constructs was placed on a bulletin board and further analysis determined the themes that emerged from these data. These themes were presented first as individual cases, then a cross case analysis.

I repeated this process for 28 observations (14 observations for the treatment case and 14 observations for the control case), then I analyzed these data again with the finalized Construct Key. To ensure that the qualitative phase of this study is credible, qualitative researchers with background in literacy were utilized as a second observer and conducted an analysis check of the data.

Credibility of the Data

The qualitative phase was devoted to addressing the issues of credibility in this study. Credibility ensures the accuracy of the data. The researcher is responsible to ensure the truthfulness of the findings and to report the findings with care. Therefore, to address the issue of credibility a second observer was used and analysis checks were done with two qualified literacy researchers. In addition, a triangulation strategy for this concurrent mixed methods study is described and also addresses supporting the credibility of this study.

Second Observer

A second observer ensured the analysis was systematic and verifiable, strategy suggested by experts in qualitative research. This enhances the accuracy of data recording (Miles & Huberman, 1994). In this study a second observer was utilized. The second observer's was a literacy education professional with extensive experience in reading content and pedagogy. In addition, she has a strong qualitative research background. Prior to any observations, I conducted a training session with the second observer.

During the training session I discussed the paradigm that applies to this study: (1) Literacy task (rereading through singing) assigned by the music teacher, and (2) Interactions among the peer groups during the literacy tasks assigned by the music teacher. Following a theory suggested by Ryan (2000) that early adolescents experience peer interactions within the context of middle school generally in three ways: (a) through information exchange (discussion), (b) modeling (peer observation and imitation), and (c) peer pressure (social reinforcement). These general categories were used as a framework to focus our observations in the field.

Once in the field we each took observational notes with both the treatment and control groups. Immediately after the observations a discussion occurred. This helped to ensure I was capturing and accurately recording the peer interactions during the literacy task.

Analysis Checks

Two qualitative researchers with backgrounds in literacy education and extensive experience in reading content and pedagogy read several transcripts. The qualitative researchers checked for credibility at two points during the qualitative phase of the study.

The researchers were given the Construct Key (Appendix C) I developed for coding purposes. The Construct Key included the constructs with descriptions. They were given several transcripts of field notes. One literacy expert was given transcripts from the group using the interactive software program, and the other literacy expert was given transcripts from the group in the regular singing class. Their coded transcripts were compared to the same transcripts I coded to determine the clarity of the constructs and definitions. We discussed any areas of disagreement and reworded descriptions presented in the construct key that were unclear for a better understanding.

Concurrent Triangulation Strategy

Triangulation involves both qualitative and quantitative formats to better measure concepts gauged individually (Creswell, 2003). This technique is an attempt to confirm, cross-validate, or corroborate findings within a single study (Morgan, 1998; Steckler, McLeroy, Goodman, Bird, & McCormick, 1992). In incorporating the two, a researcher can look for or measure data normally associated with quantitative methodologies such as outcomes as well as data commonly used in qualitative research such as perceptions (Tashakkori & Teddlie, 1996). In combination, this strategy can target a larger or more varied series of indicators or data sets usually limited within conventional research formatted studies (Creswell, 2003). In addition, “it can result in well-validated and substantiated findings” (Creswell, 2003, p. 217). It also limits the weaknesses inherent in both formats and enhances their strengths as the diversity establishes a greater reliability and reduces errors or threats. Triangulation of the data occurred in Chapter Four of this study.

Integration of the Data

Priority was given to the quantitative approach because it looked at the statistical relationship between the treatment group who used the sing-to-read program *Tune Into Reading* and the control group who were rereading through singing in their regular music class. The analysis for this approach was executed first to answer the first two questions of this study. However, concurrently qualitative case study methods were used to better understand and describe the peer interactions occurring during the literacy task assigned by their teacher. The integration of the two types of data occurred during the qualitative findings section of the research project. The quantitative results and qualitative description were mixed in order answer the research questions and to provide a clearer picture

Chapter Summary

Chapter 3 presents the methods that were used to conduct this study. It outlines the research questions, describes the design of the study, and describes the study population and participants. In addition, this chapter delineates ethical considerations, instruments, and reliability of the data. Finally, it outlines the procedures, data collection, data analysis, and credibility for the study.

CHAPTER FOUR: FINDINGS

The purpose of this study was to investigate the use of an interactive sing-to-read program, *Tune Into Reading*, as an alternative text embedded within a heterogeneous music classroom. As measured by the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006), fluency, word recognition, comprehension, and instructional reading level of the treatment students were compared to their counterparts who sang as part of the regular music program. This investigation also provided a description of the peers' interactions during the literacy tasks assigned by the music teacher. This chapter presents the results of this concurrent mixed methods study organized according to the research questions. The first two questions were concerned with the quantitative phase of the study. The descriptive and inferential statistical results, as well as interpretations, are provided. The third question is concerned with the qualitative phase of the study. Peer interactions during the literacy task of rereading through singing were examined and described. The statistical findings and the qualitative description were integrated within the qualitative findings in this study.

Question One: Quantitative Findings for Treatment and Control Groups

The findings in this section address the following research question: to what extent, as measured by the QRI-4, is the reading performance of word recognition, fluency, comprehension, and instructional reading level of students using the *Tune Into Reading* program different from their regular music curriculum counterparts?

The first quantitative research question addressed the readers' use of the interactive sing-to-read program, *Tune Into Reading*, as an alternative text compared to their counterparts in the regular music class. Prior to the treatment, I administered a pretest using the QRI-4. Scores from the pretest ensured the students in the regular music class and the students in the class using *Tune Into Reading* were not different in their reading outcomes, specifically in Fluency (WPM) measured by words per minute, Word Recognition (WR) measured by oral reading accuracy, Comprehension (COMP) measured by implicit and explicit questions after reading, and Instructional Reading Level (RL) measured by combining scores from word recognition and comprehension questions before implementation. After the implementation of the interactive sing-to-read program, *Tune Into Reading*, I administered a posttest using the QRI-4 and compared the posttest scores with the pretest scores to determine if students in the experimental group gained significantly over their counterparts in the control group. Initially, the students were assessed at posttest with a reading passage on the same instructional level attained during the pretest. This was followed with statistical analysis of the posttest on the highest instructional reading level attained by the students.

Results

A doubly multivariate repeated measure ANOVA at an alpha level of .05 was conducted on Fluency (WPM), Word Recognition (WR), Comprehension (COMP), and Instructional Reading Level (RL) from pretest to posttest by treatment group (Control vs. Treatment). Students were initially assessed using the same instructional reading level scores attained during their pretest. Means, standard deviations, and values for skewness

and kurtosis (Table 9) for WPM, WR, COMP, and RL from pretest to posttest by treatment group (n=32) and control group (n=32) are presented.

Table 9

Descriptive Statistics for Fluency (WPM), Word Recognition (WR), and Comprehension (COMP)

	Group	Pretest				Posttest			
		M	SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis
WPM	Control	136.56	36.09	-0.06	-.015	146.81	53.18	1.02	1.02
	Treatment	125.28	32.95	0.07	-.071	160.34	47.52	-0.06	0.40
WR	Control	0.98	0.02	-1.84	1.83	0.98	0.01	-1.38	0.42
	Treatment	0.98	0.01	-0.38	-1.56	0.99	0.01	-1.15	1.44
COMP	Control	0.76	0.03	1.78	6.07	0.75	0.03	2.50	11.79
	Treatment	0.77	0.04	2.24	3.36	0.85	0.07	-0.27	-0.15
RL	Control	5.58	1.22	0.35	-.047	5.58	1.22	0.35	-0.47
	Treatment	5.45	1.17	-0.10	1.65	5.45	1.17	-.010	1.65

**Note Instructional reading level is the same at pretest and posttest*

An examination of Table 9 suggested higher reading achievement scores were attained for students classified as treatment than by students classified as control at posttest in WPM, WR, and Comp on the same instructional reading level attained at pretest. The treatment group exhibited a means change in WPM from 125 at pretest to 160 at posttest, showing an increase of 35 in WPM scores; whereas, the control group went from 137 at pretest to 147 at posttest, a difference of 10 in the WPM scores. In addition to the WPM changes, the treatment group exhibited a means change in WR from .98 at pretest to .99 at posttest, illustrating an increase in word recognition scores. Whereas, the control group showed no increase in word recognition scores across the two points in time, .98 at pretest and posttest respectively on the same instructional reading level attained at pretest. Furthermore, the treatment group exhibited a means change in

COMP from .77 at pretest to .85 at posttest, demonstrating an increase in comprehension scores of .08; while, the control group's scores decreased across the two points in time in COMP, .76 at pretest and .75 at posttest on the same instructional level attained at pretest..

However, chance must be eliminated as a plausible explanation for the observed sample differences found in the population. A doubly multivariate repeated measures ANOVA was conducted with an alpha level of .05. Due to the kurtosis and skewness numbers found in the descriptive statistics, normality was assumed for two of the group distributions, WPM and WR for treatment and control at two points in time (pretest and posttest). For the third group distribution, COMP normality appeared questionable for the control group at two points in time due to leptokurtic kurtosis. Specifically, a distribution with positive kurtosis (6.07 at pretest and 11.79 at posttest) exhibits a superior acute "peak" around the mean (a higher probability than a normally distributed variable of values near the mean) and "fat tails" (a higher probability than a normally distributed variable of extreme values). Consequently, Stevens (1996) contended that "deviation from multivariate normality has only a small effect on Type 1 error" (p. 243). In addition, Tabachnick and Fidell (1989) report that "for grouped data if there is at least 20 degrees of freedom for error in the ANOVA, the reported *F* test is said to be robust to violations of normality" (p. 71).

Homogeneity of variances might be assumed, as the largest variance ratio was less than 2, which was not large enough to be considered problematic. Because the sample sizes were equal in each group, the analysis was expected to be relatively robust to

violations. Based on the analysis assumptions, it appeared reasonable to conduct the doubly multivariate repeated measures ANOVA.

There was a simultaneous difference on WPM, WR, and COMP at the same instructional reading level from pretest to posttest by treatment group, $F(4, 59) = 10.539$, $p < .001$, $\eta^2 = .417$. The Wilks' Lambda for within subjects (time) was $F(3, 60) = 14.623$, $p < .001$, $\eta^2 = .422$. The Wilks' Lambda for within subjects time by treatment interaction was $F(3, 60) = 12.039$, $p < .001$, $\eta^2 = .376$. Table 10 presents an ANOVA of WPM, WR, COMP, and RL.

Table 10

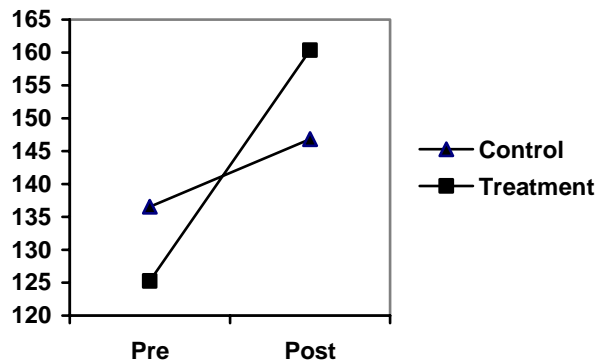
ANOVA Table on Fluency (WPM), Word Recognition (WR), Comprehension (COMP), And Instructional Reading Level (RL)

DV	F	Sig.	η^2
WPM			
Group (gp)	.014 (2870.637)	.906	.000
Time (T)	18.957	.000	.234
gp * T	5.684 (866.47)	.020	.084
WR			
Group (gp)	.278 (.000)	.600	.004
Time (T)	4.641	.035	.070
gp * T	.364 (.000)	.549	.006
COMP			
Group (gp)	43.447 .002	.000	.412
Time (T)	16.484	.000	.210
gp * T	27.356 (.002)	.000	.306
RL			
Group (gp)	.175 (2.862)	.677	.003
Time (T)	--	--	--
gp * T	--	--	--
	--		

Fluency (WPM)

There was a Group (treatment vs. control) and Time (pretest vs. posttest) interaction for WPM, $F(1, 62) = 5.684, p = .020, \eta^2 = .084$. This indicated that the observed differences between the pretest and posttest for students in the treatment condition were different from the observed differences for students in the control condition in WPM. The main effect for Group was not statistically significant, $F(1,62)=.014, p=.906$, which suggested the observed average scores between students in the treatment condition and in the control condition were not large enough to indicate a difference existed between the groups in WPM. However, the main effect for Time, $F(1,62)=18.96, p=0.00$, was found to be statistically significant, which suggested the overall mean score at Time 1 differed from the overall mean score at Time 2. To indicate relative positions of the sample means, an interaction graph is provided in Figure 6.

Figure 6
Group (Treatment vs. Control) and Time (Pretest vs. Posttest) Interaction for Fluency (WPM)



The interaction graph of Group and Time for WPM illustrates a disordinal interaction. Relative to Fluency (WPM), the data indicated a mean 13.53 points lower for control students than for treatment students at posttest. The size of the interaction effect

could also be expressed using eta squared (η^2) effect size (eta squared small .0-.3, medium .3-.5, and large above .7). The calculated value, $\eta^2 = .084$, indicated a fairly small effect size; however, it was of statistical significance. To further examine the interaction for Fluency (WPM), two *t*-tests at an alpha level of .05 each were conducted.

A dependent samples *t*-test was conducted for the control group, and no significant difference existed from the pretest ($M = 136.56$, $SD = 36.08$) to posttest ($M = 146.81$, $SD = 53.18$), $t(31) = -1.255$, $p = .219$, showing a small effect size of $d = .2$. A dependent samples *t*-test was conducted for the treatment group, and pretest scores ($M = 125.28$, $SD = 160.34$) were significantly lower than the posttest scores ($M = 160.34$, $SD = 47.52$), $t(31) = -5.434$, $p < .001$ with a large effect size of $d = .8$, indicating WPM treatment group's scores significantly increased from pretest to posttest.

In summary, it was found that pretest and posttest scores for WPM were significantly different between control and treatment groups. The treatment group showed a significant increase from pretest to posttest with a large effect size; whereas, within the control group there was no significant increase from pretest to posttest with a small effect size. It could therefore be interpreted that the treatment group made a significant increase from pretest to posttest in their fluency (WPM), as measured by words per minute on the same instructional reading level attained at pretest and compared to the control group.

Word Recognition (WR)

Word recognition (WR) data revealed no statistically significant interaction for Group By Time. In the control group scores reported from pretest ($M = .9819$, $SD = .02$) to posttest ($M = .9847$, $SD = .01$), there was a small effect size of $d = .2$. Whereas, the treatment

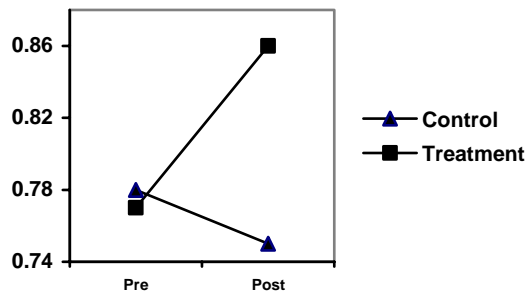
group scores demonstrated from pretest ($M=.9819$, $SD=.02$) to posttest ($M=.9869$, $SD=.01$) a medium effect size of $d=.6$. This suggested WR was more effective for the treatment group compared to the control group with a small effect from pretest to posttest at the same instructional level attained at pretest.

Comprehension (COMP)

There was a statistically significant Group By Time interaction for comprehension (COMP), $F(1, 62) = 27.356$, $p < .001$, $\eta^2 = .306$. This indicated the observed differences between the pretest and posttest for students in the treatment condition were different from the observed differences for students in the control condition in reading comprehension (COMP). The main effect for Group was statistically significant, $F(1,62)= 43.44$, $p=.000$, which suggested the observed average difference between students in the treatment condition and in the control condition was large enough to indicate a difference existed between the groups in COMP. In addition, the main effect for Time, $F(1,62)=16.48$, $p=0.00$, was found to be statistically significant, which suggested the overall mean score at Time 1 differed from the overall mean score at Time 2. To indicate relative positions of the sample means, an interaction is provided in Figure 7 to indicate relative positions of the sample means.

Figure 7

Group (Treatment vs. Control) and Time (Pretest vs. Posttest) Interaction for COMP



The interaction graph of Group and Time for COMP type illustrates a disordinal interaction. Relative to the reading comprehension (COMP) scores, the data indicate a mean that was .10 points lower for control students than for treatment students at posttest. The size of the interaction effect could also be expressed using eta squared (η^2) effective size. The calculated value, $\eta^2 = .306$, indicated a medium effect size that demonstrated statistical significance. To further examine the interaction for reading comprehension (COMP), two *t*-tests at an alpha level of .05 each were conducted.

When a dependent samples *t*-test was conducted for the control group, there was no significant difference from pretest ($M = .76, SD = .03$) to posttest ($M = .75, SD = .03$), $t(31) = 1.404, p = .170$ with a small effect size of $d = .3$. However, the treatment group posttest scores ($M = .85, SD = .07$) were significantly higher than their pretest scores ($M = .77, SD = .04$), $t(31) = -5.110, p < .001$, showing a very large effect size of $d = 1.17$. Therefore, indicating that for reading comprehension COMP, the treatment group's scores significantly increased from pretest to posttest.

In summary, it was found that the pretest and posttest scores for COMP were significantly different between control and treatment groups. The treatment group

illustrated a significant increase from pretest to posttest with a large effect size $d=1.17$; whereas, within the control group there was no significant increase from pretest to posttest with a small effect size $d=.3$. It could therefore be interpreted that the treatment group made a significant increase from pretest to posttest in their reading comprehension (COMP), as on the same instructional reading level attained at pretest when compared to the control group.

Highest Instructional Reading Level

Analysis was conducted on Fluency (WPM), Word Recognition (WR), and Comprehension (COMP) at the highest Reading Levels (RL) attained at posttest for the control and treatment groups. Four between-groups analyses of variance (ANOVAs) were conducted on highest reading level scores for WPM, WR, COMP, and RL. Type I error was controlled by using the Bonferroni adjustment of the significant level to .02.

The results revealed that between the Groups at the highest instructional reading level there were no statically significant differences for WPM, WR, and COMP. However, it was found that for RL (instructional reading level) by Group (treatment vs. control), the treatment group showed a significant increase compared to the control group at the highest instructional reading level. The between-groups analyses of variance (ANOVA) indicated the Treatment instructional reading level scores RL ($M = 6.58$, $SD = .1.59$), $F(1, 62) = 31.28$, $p < .001$, $\eta^2 = .335$ were significantly higher than the Control RL ($M = 5.77$, $SD = 1.44$) at the highest instructional reading level. This suggested that even though the treatment and control groups showed no significant difference in WPM, WR, or COMP, the treatment group increased significantly in their instructional reading

levels at the highest level attained at posttest. Table 11 displays the percentages by group of the instructional reading changes at posttest.

Table 11

Changes in Instructional Reading Levels for Treatment and Control Groups

	Treatment (n=32)	Control (n=32)
Attained a Higher Level	81%	12%
	n=26	n=4
Stayed at the Same Level	19%	88%
	n=6	n=28

Summary of Finding for Question 1

In conclusion, the treatment group, using the interactive singing software *Tune Into Reading*, demonstrated a significant increase with large effect sizes in Fluency (WPM) $d=.8$ and Reading Comprehension (COMP) $d=1.17$ as compared to the control group who were singing in the regular music class at the same instructional reading level attained during the pretest. In addition, although there were no observed differences noted in the interaction for Word Recognition (WR), the treatment group effect size was larger $d=.6$ than the control group effect size of $d=.3$. This suggested that from pretest to posttest the treatment group had a larger effect for WR than the control group. Furthermore, at the highest Instructional Reading Level (RL) the treatment group showed a significant increase in RL with a medium effect size $d=.7$ as compared to the control,

whose effect size reported was very small. Table 12 displays the interactions and effect sizes for the groups by variables.

Table 12
Interactions and Effect Sizes for WPM, WR, COMP, and RL by Groups

Variables	Time X Group	Treatment n=32	Control n=32
WPM	*	d=.8	d=.2
WR	NS	d=.6	d=.2
COMP	***	d=1.17	d=.3
RL	***	d=.7	d=.1

* Note * small significant effect, *** large significant effect, and NS no significant effect.

These findings suggests the treatment students of varying reading abilities that used the interactive singing program, *Tune Into Reading*, illustrated a significant increase in their Fluency (WPM), Reading Comprehension (COMP), and Instructional Reading level (RL) as compared to their counterparts who were singing in the regular music class. In addition, for the treatment students Word Recognition (WR) indicated a larger effect from pretest to posttest than the control group. Specifically, this suggests that rereading through singing, using the interactive singing program, *Tune Into Reading*, was more effective regardless of the reading levels for treatment students compared to control students. These results can be interpreted as rereading through singing in the music classroom alone, as was the case for the control students, does not improve WPM, WR, COMP, and RL for the students of varying reading abilities.

Furthermore, at the increased reading level reported at posttest, even though the treatment group had a significant increase in their instructional reading level (RL), there was no significant difference between the groups in WPM, WR, or COMP. This suggested that, even though the treatment students increased in their instructional reading level (pretest M=5.45 and posttest M=6.58), their reading scores at the higher instructional reading level in WPM, WR, and COMP were lower than their scores at posttest on the same instructional reading. Specifically, as the early adolescents in the treatment condition increased in text difficulty, their fluency (WPM), word recognition (WR), and comprehension (COMP) shifted from a fluent expert reader, on the similar level passage attained at pretest, to a surface fluent reader (e.g., Topping, 2006) at a higher level.

Question Two: Quantitative Findings for Group by FCAT Reading Levels

The findings in this section address the following research question: To what extent does the *Tune Into Reading* program impact reading scores of students who are “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) reading scores?

The second quantitative research question investigated whether an interaction effect of the repeated reading methods through singing occurred on the reading performance of students stratified as “*Below, At, or Above*” grade level in the treatment condition as compared to their counterparts in the control condition. The results from the QRI-4 pretest and posttest reading scores were used to determine reading outcomes (WPM, WR, COMP, and RL) for the treatment and control groups. Then, students achievement level scores (levels 1-5) in reading were used to stratify the groups as

“*Below, At, or Above*” grade level in the treatment condition (students who used the interactive sing to read program) compared to their counterparts in the control condition (students who were singing as part of the regular music program). The achievement levels as determined by the Florida Department of Education are: (a) Levels 1 or 2 that are considered *Below* proficiency in meeting grade level benchmarks, (b) Level 3 that is considered *At* grade level, and (c) Levels 4 or 5 that are considered *Above* grade level (*FCAT Briefing Book, 2005*).

Fluency (WPM) for FCAT Levels 1-5

A repeated measures ANOVA at an alpha level of .05 was conducted for FCAT Levels 1-5 in reading scores on WPM (words per minute). Students were initially assessed on the same instructional reading level attained during their pretest. Means, standard deviations, skewness, and kurtosis for WPM, from pretest to posttest by treatment group (Control n= 56 vs. Treatment n=56), are presented in Table 13.

Table 13

Descriptive Statistics for Fluency (WPM) for FCAT Levels 1-5

	Group	M	SD	Pretest		M	SD	Posttest	
				Skewness	Kurtosis			Skewness	Kurtosis
WPM 1 & 2	Control	115.00	44.32	-0.51	-1.65	124.75	30.09	-0.85	0.21
	Treatment	110.25	28.25	0.11	0.02	129.38	51.54	-0.02	-0.01
WPM 3	Control	135.75	28.16	-0.11	-1.32	138.50	34.06	-0.23	-1.37
	Treatment	123.08	34.66	0.33	-0.94	180.50	58.57	1.02	1.03
WPM 4 & 5	Control	164.13	30.70	-1.58	2.44	201.88	65.82	-0.14	-1.54
	Treatment	151.50	25.01	-0.35	1.33	186.50	34.72	1.69	3.33

An examination of Table 13, stratified groups as *Below* (FCAT Levels 1 & 2), *At* (FCAT Level 3), and *Above* (FCAT Levels 4 and 5) for Fluency (WPM), suggested students classified as treatment from FCAT Levels 1-3 attained higher reading achievement scores than students classified as control at posttest on the same instructional reading level attained at pretest. The groups stratified as FCAT Level *Below* (1 & 2) exhibited a mean change in WPM from 110 at pretest to 129 at posttest; whereas, the control group went up across the two points in time, 115 at pretest to 125 at the posttest. The groups stratified as FCAT Level *At* (3) exhibited a means change in WPM from 123 at pretest to 181 at posttest; whereas, the control group increased across the two points in time, 136 at pretest and 139 at the posttest on the same instructional reading level attained at pretest. However, for groups stratified as FCAT Level *Above* (4 and 5), the control group appeared to have higher reading achievement scores than the treatment group. The control group FCAT Level *Above* (4 & 5) exhibited a mean change in WPM from 164 at pretest to 202 at posttest; whereas, the treatment group mean change increased across the two points in time, 152 at pretest and 187 at posttest on the same instructional reading level attained at pretest.

However, to suggest that differences would be found in the population, chance must be ruled out as a plausible explanation for the observed sample differences. A repeated measures ANOVA was conducted with an alpha level of 0.05. The kurtosis and skewness numbers found in the descriptive statistics suggested normality could be assumed for all three of the group distributions. In addition, homogeneity of variances might be assumed, as the largest variance ratio was less than 2, which was not large enough to be considered problematic. Furthermore, since the sample sizes were equal for

each group, the analysis was expected to be relatively robust to violations of the homogeneity of variance assumption. Based on the analysis assumptions, it appeared reasonable to conduct the repeated measures ANOVA for the FCAT Level groups 1-5 on WPM.

A repeated measures ANOVA was conducted in order to determine if a difference exists in the reading scores between groups, across time, and for the different FCAT levels. Alpha level was set at .05. Table 14 illustrates the results of the repeated measures ANOVA.

Table 14

ANOVA Table FCAT Levels 1-5 for WPM

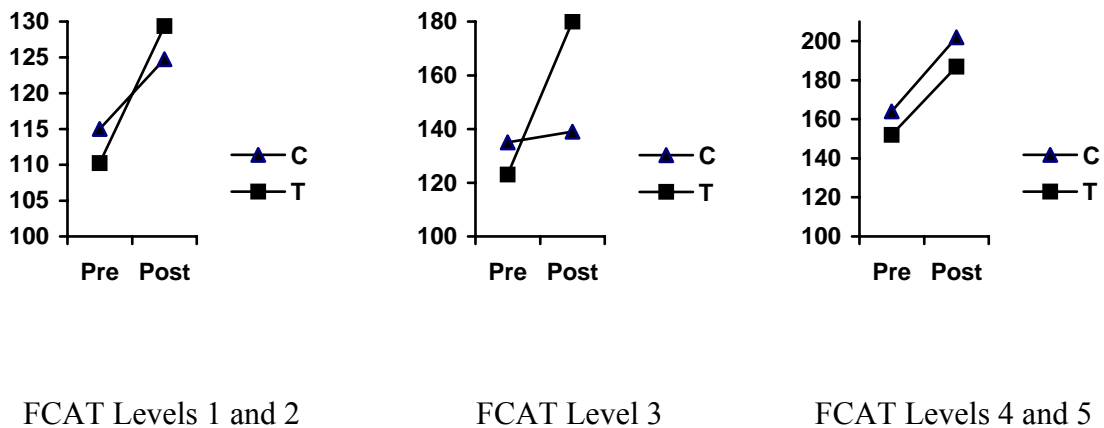
DV	F	Sig.
WPM		
Group (gp)	0.01 (14255.512)	.930
Level (L)	7.83	.001
gp * L	0.27 (6443.604)	.762
Time (T)	18.02	.000
gp * T	3.85 (907.8478)	.062
L * T	1.56 (1332.9236)	.221
gp*L*T	3.28 (854.65278)	.047

A review of the ANOVA table indicated for FCAT Levels 1-5 on WPM, there was a statistically significant interaction for Group By Time By Level $F(3, 55) = 3.28, p = .04, \eta^2 = .197$. This indicated that observed differences between pretest and posttest for

students in the treatment condition were different from the observed differences for students in the control condition within the three FCAT Levels on WPM. To indicate relative positions of the sample means, interaction graphs for the three Levels by Groups across Time are provided in Figure 8.

Figure 8

Interaction Graphs of FCAT Levels 1-5 on WPM



The interaction graph of Group By Time By Level for WPM reveals disordinal interactions for FCAT Levels 1-3. However, for FCAT Levels 4 and 5 the interaction is ordinal. Relative to fluency scores (WPM), the data indicate a mean for FCAT Levels 1 and 2 that were 4 points lower for control students than treatment students at posttest. The size of the interaction effect exhibits a calculated value of $\eta^2 = .017$, indicating a small effect size. In addition, the fluency scores (WPM) for FCAT Level 3 indicated control students were 41 points lower than treatment students at posttest. The size of the interaction effect could also be expressed using eta squared (η^2) effective size. The calculated value, $\eta^2 = .240$, indicated a small effect size. However, relative to fluency scores (WPM), the data indicate a mean for FCAT Levels 4 and 5 that was 15 points

lower for treatment students than for control students at posttest. The size of the interaction effect could also be expressed using eta squared (η^2) effective size. The calculated value, $\eta^2 = .001$, indicated a small effect size. To further examine the interaction for fluency (WPM), three *t*-tests at an alpha level of .05 each were conducted.

For FCAT Levels 1 and 2, there was no statistically significant difference between the groups. The control group scores reported from pretest ($M=115.00$, $SD=44.32$) to posttest ($M=124.75$ $SD=30.09$) illustrated a medium effect size of $d=.7$. Whereas, the treatment group scores demonstrated from pretest ($M=110.25$, $SD=28.25$) to posttest ($M=129.38$, $SD=51.54$) a large effect size of $d=1.1$. This suggested the treatment group WPM had a larger effect in their scores compared to the control group with a medium effect from pretest to posttest on the same instructional level attained at pretest.

For FCAT Level 3 on WPM scores, the treatment group ($M = 180.50$, $SD = 58.57$) was significantly greater than the control group ($M = 138.50$, $SD = 34.06$), $t(22) = -2.148$, $p = .043$. The control group scores reported from pretest ($M=135.75$, $SD =28.26$) to posttest ($M=138.50$ $SD=34.06$) illustrated a small effect size of $d=.1$. Whereas, the treatment group scores exhibited from pretest ($M=123.08$, $SD=34.66$) to posttest ($M=180.50$, $SD=58.57$) a large effect size of $d=1.4$. This suggested the treatment group in FCAT Level 3 outperformed the control group on WPM and had a larger effect in their scores for WPM compared to the control group with a medium effect from pretest to posttest on the same instructional level attained at pretest.

For FCAT Levels 4 and 5, there was no statistical significant difference between the groups. The control group scores reported from pretest ($M=164.13$, $SD =30.70$) to posttest ($M=201.88$, $SD=65.82$) with a small effect size of $d=.3$. Whereas, the treatment

group scores revealed from pretest ($M=151.50$, $SD=25.01$) to posttest ($M=186.50$, $SD=34.72$) a small effect size of $d=.4$. This suggested that for FCAT Levels 4 and 5 WPM had little effect on the scores for both groups. These findings can therefore be interpreted as when the students are grouped by FCAT Levels in reading, the variable of Fluency (WPM), measuring reading rate, is more effective for students in FCAT Levels 1 and 2 (*Below*) and FCAT Levels 3 (*At*) than FCAT Levels 4 and 5 (*Above*). This suggests that when thinking about WPM for the higher performing students, reading rate may not be an important variable.

Word Recognition (WR) for FCAT Levels 1-5

A repeated measures ANOVA at an alpha level of .05 was conducted for FCAT Levels in reading scores (FCAT levels 1 -5) on WR (word recognition). Students were initially assessed on the same instructional reading level attained during their pretest. Means, standard deviations, skewness, and kurtosis for WR, from pretest to posttest by treatment group (Control $n= 56$ vs. Treatment $n=56$) are presented in Table15.

Table 15

Descriptive Statistics on Word Recognition (WR) for FCAT Levels 1-5

	Group	Pretest				Posttest			
		M	SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis
WR 1 & 2	Control	0.97	0.03	-2.54	6.73	0.98	0.01	-0.64	-2.24
	Treatment	0.98	0.01	-0.64	-2.24	0.99	0.00	2.83	8.00
WR 3	Control	0.99	0.01	0.72	-0.79	0.99	0.01	-2.54	6.77
	Treatment	0.98	0.01	-0.00	-1.88	0.98	0.01	-1.15	-0.25
WR 4 & 5	Control	0.99	0.01	-0.64	-2.24	0.99	0.01	-2.83	8.00
	Treatment	0.99	0.01	0.00	-2.80	0.99	0.00	-2.83	8.00

An examination of Table 15, stratified groups as *Below* (FCAT Levels 1 & 2), *At* (FCAT Level 3) and *Above* (FCAT Levels 4 and 5) for Word Recognition (WR), suggested that there was a higher reading achievement scores attained for the students classified as treatment from FCAT Levels 1 and 2 than for the students classified as control at posttest on the same instructional reading level attained at pretest. The groups stratified as FCAT Level *Below* (1 & 2) in treatment exhibited a mean change in WR from .98 at pretest to .99 at posttest; whereas, the control group increased across the two points in time, .97 at pretest and .98 at the posttest. The groups stratified as FCAT Level *At* (3) exhibited no means change in WR for either the treatment or the control group. The treatment group scores for WR were .98 from pretest to posttest, and the control group had a slightly higher score of .99 from pretest to posttest. The groups stratified as FCAT Levels *Above* (4 and 5) WR scores showed no changes for either the treatment or control group at .99 on the same instructional reading level that was attained at pretest. This “ceiling effect” in word recognition suggests this test may be too easy for this group of students (Glass & Hopkins, 1996).

However, to suggest that differences would be found in the population, chance must be ruled out as a plausible explanation for the observed sample differences. A repeated measures ANOVA was conducted with an alpha level of 0.05. The kurtosis and skewness numbers found in the descriptive statistics suggested normality could not be assumed for the three of the group distributions. Specifically for FCAT Levels 1 and 2, there was a leptokurtic kurtosis distribution of 6.53 for the control group at pretest and a leptokurtic kurtosis distribution of 8.00 for the treatment group at posttest. In addition,

FCAT Level 3 control group at posttest for WR was negatively skewed -2.34 and had a leptokurtic kurtosis distribution of 6.77. Furthermore, FCAT Levels 4 and 5, the control and treatment groups at posttest for WR were negatively skewed at -2.84 and had a leptokurtic kurtosis distribution of 8.00. However, Stevens (1996) contended that “deviation from multivariate normality has only a small effect on Type 1 error” (p. 243). In addition, Tabachnick and Fidell (1989) report that for grouped data with an equal sample size, the reported *F* test was said to be robust to violations of normality.

Homogeneity of variances might be assumed, as the largest variance ratio was less than 2, which was not large enough to be considered problematic. Furthermore, since the sample sizes were equal for each group, the analysis was expected to be relatively robust to violations of the homogeneity of variance assumption. Based on the analysis assumptions, it appeared reasonable to conduct the repeated measures ANOVA for the FCAT Level groups 1-5 on WR.

A repeated measures ANOVA was conducted in order to determine if there was a difference in the reading scores across groups, between time and for the different FCAT Levels. Alpha level was set at .05. Table 16 shows the results of the repeated measures ANOVA for WR.

Table 16

ANOVA Table for FCAT Levels 1-5 on WR

DV	F	Sig.
WR		
Group (gp)	0.00 (0.6253501)	.992
Level (L)	10.92	.003
gp * L	0.00 (0.1574536)	.998
Time (T)	4.16	.053
gp * T	0.76 (0.0000736)	.391
L * T	1.16 (1113.786)	.323
gp*L*T	0.50 (0.0000880)	.619

A review of the ANOVA table indicated that the data revealed for FCAT Levels 1-5 on WR, there was no statistically significant interaction for Group By Time By Level $F(3, 53)=0.50, p=.619$. In addition, there was also no Group By Level, Time By Group, or Level By Time interactions. However, the main effects for Time $F(3, 53)=04.16, p=.053$ and Level $F(3, 53)=10.92, p=.003$ were statistically significant. This can be inferred as the means for the three Levels were different from Time 1 to the overall mean score at Time 2.

The group effect size for FCAT Levels 1 and 2 was $\eta^2 = .067$, a small effect size. However, within the group, the treatment students showed a large effect size of $d=1.0$ compared to the control group with a small effect size of $d=.3$.

In addition, for FCAT Level 3, the group's effect size was small $\eta^2 = .091$. However, within the group, the treatment students showed a medium effect size of $d = .5$ compared to the control group with a small effect size of $d = .3$. Furthermore, for FCAT 4 and 5, this level too had a small effect size for WR $\eta^2 = .036$. However, within the group, the treatment students showed a medium effect size of $d = .6$ compared to the control group with a small effect size of $d = .2$. These findings might be interpreted as the variable of WR was more effective for the treatment students than the control students on the same instructional reading attained at pretest.

Comprehension (COMP) for FCAT Levels 1-5

A repeated measures ANOVA at an alpha level of .05 was conducted for FCAT Levels in reading scores (FCAT levels 1 -5) on COMP (reading comprehension). Students were initially assessed on the same instructional reading level attained during their pretest. Means, standard deviations, skewness, and kurtosis for WR, from pretest to posttest by treatment group (Control $n = 56$ vs. Treatment $n = 56$) are presented in Table 17.

Table 17

Descriptive Statistics on Reading Comprehension (COMP) for FCAT Levels 1-5

	Group	M	Pretest			Posttest			
			SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis
COMP 1 & 2	Control	0.75	0.00	0.00	0.00	0.75	0.00	0.00	0.00
	Treatment	0.78	0.06	1.44	0.00	0.83	0.07	-0.64	-2.24
COMP 3	Control	0.76	0.04	3.46	12.00	0.75	0.00	0.00	0.00
	Treatment	0.77	0.04	2.82	8.06	0.87	0.10	-0.40	-0.58
COMP 4 & 5	Control	0.78	0.04	-1.32	0.88	0.75	0.07	1.62	2.47
	Treatment	0.77	0.05	2.83	8.00	0.86	0.05	-1.63	1.61

An examination of Table 17, stratified groups as *Below* (FCAT Levels 1 & 2), *At* (FCAT Level 3), and *Above* (FCAT Levels 4 and 5) for Comprehension (COMP), suggested higher reading achievement scores attained for the students classified as treatment from FCAT Levels 1-5 than for the students classified as control at posttest on the same instructional reading level attained at pretest. The groups stratified as FCAT Level *Below* (1 & 2) treatment group exhibited a mean change in COMP from .78 at pretest to .83 at posttest; whereas, the control group COMP scores showed no change across the two points in time .75 at pretest and posttest.

FCAT Levels *Above* (4 and 5) the treatment group exhibited means change in COMP from .77 at pretest to .87 at posttest; whereas, the control group COMP scores decreased slightly across the two points in time .76 at pretest and .75 posttest.

However, to suggest that differences would be found in the population, chance must be ruled out as a plausible explanation for the observed sample differences. A repeated measures ANOVA was conducted with an alpha level of 0.05. The kurtosis and skewness numbers found in the descriptive statistics suggested normality could not be assumed for the two of the group distributions FCAT Level 3 and FCAT Levels 4 and 5. FCAT Level 3 normality appeared to be questionable for the control and treatment groups. The control and treatment groups displayed a positive skewness and leptokurtic kurtosis distributions for COMP at pretest, specifically the control (sk= 3.46, ku=12.00) and the treatment group (sk=2.82, ku=8.06). For FCAT Levels 4 and 5, the treatment groups displayed a positive skewness of 2.83 and had a leptokurtic kurtosis distribution of 8.00 for COMP at pretest. However, Stevens (1996) contended that “deviation from multivariate normality has only a small effect on Type 1 error” (p. 243).

Homogeneity of variances might be assumed, as the largest variance ratio was less than 2, which was not large enough to be considered problematic. Furthermore, since the sample sizes were equal for each group, the analysis was expected to be relatively robust to violations of the homogeneity of variance assumption. Based on the analysis assumptions, it appeared reasonable to conduct the repeated measures ANOVA for the FCAT Level groups 1-5 on COMP.

A repeated measures ANOVA was conducted in order to determine if a difference existed in reading scores across groups, between time and for the different FCAT levels. Alpha level was set at .05. Table 18 shows the results of the repeated measures ANOVA on COMP.

Table 18

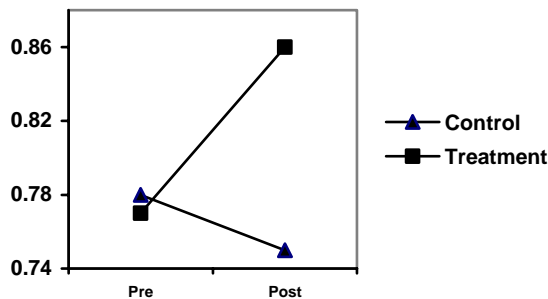
ANOVA Table for FCAT Levels 1-5 on COMP

DV	F	Sig.
WPM		
Group (gp)	0.18 (0.41365448)	.674
Level (L)	11.06	.000
gp * L	0.02 (0.0017548)	.897
Time (T)	11.80	.002
gp * T	20.16 (0.00272309)	.002
L * T	0.66 (0.00124236)	.523
gp*L*T	1.03 (0.00189312)	.366

A review of the ANOVA table indicated the data revealed for FCAT Levels 1-5 on COMP no statistical significant interaction for Group By Time By Level, $F(3, 53)=1.03, p=.366$. In addition, there were also no Group By Level or Level By Time interactions. However, there was a statistically significant Time By Group interaction $F(3, 53)=20.16, p=.002$. This suggests the observed differences between the pretest and posttest for students in the treatment condition were different from the observed differences for students in the control condition in COMP. To indicate relative positions of the sample means, an interaction graph combining the three FCAT Levels by Groups across Time is provided in Figure 9.

Figure 9

Interaction Graph of the Groups By Time on COMP



The interaction graph of Group By Time for COMP illustrates a disordinal interaction. Relative to reading comprehension (COMP) scores, the data indicate a mean that was 10 points lower for the control students than for the treatment students at posttest. The size of the interaction effect could also be expressed using eta squared (η^2) effective size. The calculated value, $\eta^2 = .349$, indicated a medium effect size. To further

examine the interaction for reading comprehension (COMP), three *t*-tests at an alpha level of .05 each were conducted on the three FCAT Levels.

For FCAT Levels 1 and 2, there was a statistically significant difference in comprehension (COMP), such that the Treatment group ($M = .83$, $SD = .07$) effect size $d = .3$ was significantly greater than the Control group ($M = .75$, $SD = .00$), $t(7) = -3.416$, $p = .011$ $d = 0$, on the same instructional reading level attained at pretest. In addition, FCAT Level 3 on COMP scores, demonstrated the treatment group COMP posttest scores ($M = .87$, $SD = .10$) effect size $d = 1.2$ were significantly higher than control COMP posttest scores ($M = .75$, $SD = .00$), $t(22) = 4.01$, $p < .001$, effect size $d = 0$. Furthermore, for FCAT Levels 4 and 5, the control group scores reported no difference on COMP scores, such that pretest scores ($M = .77$, $SD = .04$) were not significantly different from posttest scores ($M = .75$, $SD = .07$), $t(7) = 1.000$, $p = .351$, with effect size of $d = -.4$ on the same instructional reading level attained at pretest.. However, the treatment group showed there was a significant difference on the COMP scores between posttest ($M = .86$, $SD = .05$), and pretest ($M = .77$, $SD = .05$), $t(7) = -2.714$, $p = .030$, with a medium effect size of $d = .6$. This suggested treatment group had a significantly higher mean in their reading comprehension (COMP) scores at posttest compared to the control group, for students stratified by FCAT Levels 4 and 5. The findings reported suggested that for the treatment students in FCAT Levels 1-5, COMP was more effective compared to the control groups in FCAT Levels 1-5.

Highest Instructional Reading Level on WPM for FCAT Levels 1-5

Three between-groups analyses of variances (ANOVAs) were conducted. Type I error was controlled by using the Bonferroni adjustment of the significant level to .02 at

the highest instructional reading level attained at posttest on WPM by group (treatment vs. control) for each of the FCAT levels (*Below* level, *At* level, and *Above* level). Means, standard deviations, skewness, and kurtosis at the highest reading level attained at posttest for WPM are reported by group (treatment vs. control) for each FCAT level (*Below* level, *At* level, and *Above* level) and presented in Table 19.

Table 19

Means, Standard Deviations, Skewness, and Kurtosis at the Highest Reading Level for WPM Scores by Group (Treatment vs. Control) and FCAT level

FCAT Level	Group	M	SD	Skewness	Kurtosis
Below n=16	Treatment	119.88	26.08	0.91	-0.54
	Control	127.88	50.63	-1.11	0.73
At Level n=24	Treatment	138.50	34.06	0.69	0.36
	Control	181.25	61.85	-0.23	-1.37
Above n=16	Treatment	184.88	59.22	0.33	1.35
	Control	127.63	26.64	-0.16	-0.10

The students stratified as FCAT level 1 and 2 *Below* level showed no difference at the increased reading level at posttest for WPM by group (treatment vs. control), $F(1, 14) = .158, p = .697, \eta^2 = .011$. However, for the students stratified as FCAT level 3 *At* level, the treatment group was significantly higher than the control group for WPM, $F(1, 22) = 4.399, p = .048, \eta^2 = .167$, at the highest reading level. When students were stratified for FCAT levels 4 and 5 *Above* level in reading, the treatment group was significantly higher than the control group, $F(1, 14) = 6.217, p = .026, \eta^2 = .308$, at the highest instructional reading level for WPM attained at posttest.

Highest Instructional Reading Level on WR for FCAT Levels 1-5

Three between-groups analyses of variances (ANOVAs) were conducted. Type I error was controlled by using the Bonferroni adjustment of the significant level to .02, at the highest instructional reading level attained at posttest, on WR by group (treatment vs. control) for each of the FCAT levels (*Below* level, *At* level, and *Above* level). Means, standard deviations, skewness, and kurtosis at the highest reading level attained at posttest for WR by group (treatment vs. control) for each FCAT level (*Below* level, *At* level, and *Above* level) are presented in Table 20.

Table 20

Means, Standard Deviations, Skewness, and Kurtosis at the Highest Reading Level for WR scores by Group (Treatment vs. Control) and FCAT level

FCAT Level	Group	M	SD	Skewness	Kurtosis
Below n=16	Control	0.98	0.01	-0.64	-2.24
	Treatment	0.98	0.01	-0.64	-2.24
At Level n=24	Control	0.98	0.01	-1.71	2.23
	Treatment	0.98	0.01	-0.19	-2.25
Above n=16	Control	0.99	0.01	-2.83	8.00
	Treatment	0.99	0.01	-2.83	8.00

For the students stratified as FCAT levels 1 and 2 *Below* level, there was no difference at the highest reading level in WR by group $F(1, 14) = .000, p = 1.000, \eta^2 = 1.000$. In addition, for the students stratified as FCAT level 3 *At* level, there was no mean difference at the highest reading level at posttest in WR by group, $F(1, 22) = .672, p = .421, \eta^2 = .030$. Furthermore, when students were stratified for FCAT levels 4 and 5 *Above* level, there was no mean difference at the increased reading level attained at posttest in WR by group, $F(1, 14) = 1.000, p = .334, \eta^2 = .067$.

Highest Instructional Reading Level on COMP for FCAT Levels 1-5

Three between-groups analyses of variances (ANOVAs) were conducted Type I error was controlled by using the Bonferroni adjustment of the significant level to .02 at the highest instructional reading level attained at posttest on COMP by group (treatment vs. control) for each of the FCAT levels (*Below* level, *At* level, and *Above* level). Means, standard deviations, skewness, and kurtosis are reported at the highest reading level attained at posttest for COMP by group (treatment vs. control) for each FCAT level (*Below* level, *At* level, and *Above* level) and presented in Table 21.

Table 21

Means, Standard Deviations, Skewness, and Kurtosis at the Highest Reading Level for COMP scores by Group (Treatment vs. Control) and FCAT level

Below n=16	Control	0.75	0.00	0.00	0.00
	Treatment	0.74	0.02	-2.83	8.00
At Level n=24	Control	0.75	0.01	-3.46	12.00
	Treatment	0.76	0.06	0.89	0.68
Above n=16	Control	0.74	0.04	0.40	-0.23
	Treatment	0.75	0.03	0.00	3.50

For the students stratified as FCAT level 1 and 2 *Below* level, there was no difference on the highest reading level at posttest and no mean difference for COMP by group (treatment vs. control), $F(1, 14) = 1.000, p = .334, \eta^2 = .067$, at the highest reading level attained at posttest. In addition, for the students stratified as FCAT level 3 *At* level, there was no mean difference on the increased reading level at posttest COMP by group, $F(1, 22) = .428, p = .520, \eta^2 = .019$. Furthermore, when students were

stratified for FCAT levels 4 and 5 *Above* level, there was no mean difference on the highest reading level attained at posttest in for COMP by group, $F(1, 14) = .636$, $p = .438$, $\eta^2 = .043$.

Highest Instructional Reading Level on RL for FCAT Levels 1-5

Three between-groups analyses of variances (ANOVAs) were conducted. Type I error was controlled by using the Bonferroni adjustment of the significant level to .02 at the highest instructional reading level attained at posttest on RL by group (treatment vs. control) for each of the FCAT levels (*Below* level, *At* level, and *Above* level). Means, standard deviations, skewness, and kurtosis at the highest reading level attained at posttest by group (treatment vs. control) for each FCAT level (below level, at level, and above level) are presented in Table 22.

Table 22

Means, standard deviations, skewness, and kurtosis for RL at the Highest Reading Level for RL scores by Group (Treatment vs. Control) and FCAT level

FCAT Level	Group	M	SD	Skewness	Kurtosis
Below n=16	Control	4.50	1.27	0.00	-2.80
	Treatment	5.56	1.29	-0.84	2.14
At Level n=24	Control	5.25	0.45	1.33	-0.33
	Treatment	7.46	1.61	0.15	-1.40
Above n=16	Control	7.63	1.33	0.37	-0.66
	Treatment	6.63	1.27	2.11	4.17

For students stratified as FCAT level 1 and 2 *Below* level for RL, the treatment group was significantly higher than the control group, $F(1, 22) = 45.000$, $p < .001$, $\eta^2 = .763$. The effect size for the treatment group was $d = .7$; whereas the control group was

$d=.3$, indicating the variable RL for treatment group had a higher effect size compared to the control group. In addition, for the students stratified as FCAT level 3 *At* level at the higher instructional reading level, the treatment group was significantly higher than the control group, $F(1, 22) = 14.474, p = .001, \eta^2 = .397$. The effect size for the treatment group was $d=.9$; whereas the control group was $d=.0$, indicating the variable RL for treatment group had a higher effect size compared to the control group. However, students were stratified for FCAT levels 4 and 5 *Above* level at the higher instructional reading level attained at posttest, and no mean difference existed on FCAT RL by group (treatment vs. control), $F(1, 14) = 2.966, p = .107, \eta^2 = .175$. However, the effect size for the treatment group was $d=.5$; whereas the control group was $d=.0$, indicating the variable RL for treatment group had a higher effect size compared to the control group. The findings suggest that for the treatment students, RL was more effective compared to the control group.

Summary of Findings for Question 2

In conclusion when the students were stratified by FCAT Levels 1 and 2 *Below* and FCAT Level 3 *AT*, at the same instructional reading level attained at pretest on WPM, there was a statistically significant difference between groups across time and within levels. Further analysis suggested reading rate was more effective for treatment students than control students in FCAT Levels 1-3. However, for FCAT Levels 4 and 5 on WPM, there was no significant difference between the groups across time. In addition, although there were no observed differences noted in the interaction for Word Recognition (WR), the treatment group effect size for each level was larger than the control group effect size. This suggested, from pretest to posttest, the treatment group had

a larger effect for WR than the control group. Furthermore, reading comprehension COMP for the treatment group, using the interactive singing software *Tune Into Reading*, demonstrated a significant increase with large effect sizes. Finally, at the highest Instructional Reading Level (RL), the treatment groups showed a significant increase in RL with a larger effect size as compared to the control groups. Table 23 displays the effect sizes for the groups by FCAT Levels on the four variables.

Table 23

Effect Sizes for WPM, WR, COMP, and RL by FCAT Level Groups

Variables	FCAT 1 and 2		FCAT 3		FCAT 4 and 5	
	Treatment	Control	Treatment	Control	Treatment	Control
WPM	d=1.1	d=.7	d=1.4	d=.1	d=.4	d=.3
WR	d=1.0	d=.4	d=.5	d=.2	d=.6	d=.2
COMP	d=.3	d=0	d=1.2	d=0	d=.6	d=-.4
RL	d=.7	d=.1	d=.9	d=0	d=.5	d=0

These findings suggests the treatment students of varying reading abilities that used the interactive singing program, *Tune Into Reading*, illustrated a significant increase in their Fluency (WPM), Reading Comprehension (COMP), and Instructional Reading level (RL) as compared to their counterparts who were singing in the regular music class. In addition, on the variables by level, it appears FCAT Levels 1-3 had a larger effect than FCAT Level 4 and 5 on WPM, and FCAT Levels 1 and 2 showed a larger effect on WR than Levels 3-5. Further examination revealed that for reading comprehension COMP, it appeared to be more effective for FCAT Levels 3-5 than for Levels 1 and 2. This implies

the different levels, when using the interactive singing software *Tune Into Reading*, appeared to be more effective for each group differently. Particularly, those students needing fluency (WPM) increased in reading rate; whereas, those needing more opportunities for reading comprehension increased in their scores. This suggests the use of the interactive sing-to-read program provides for its user's differentiated instruction.

Question Three: Qualitative Findings for Peer Interactions

Due to the interpretive case study design of this phase of the study (Patton, 2002), this section is devoted to presenting an analysis of the data within individual cases, followed by a cross-case analysis. This was conducted to determine the major themes for each case as well as those themes across the cases. The constructs and themes that emerged from the data were useful in answering the research question that guided this phase of the study: How do middle school readers interact with their peers within the context of their music classroom?

Case studies included two groups of early adolescent peers who participated in this study. The two cases consisted of early adolescents in a music classroom who used the interactive singing program, *Tune Into Reading*, and their counterparts who sang as part of their regular music class. The focus was on the descriptions of peer interactions during the literacy task of rereading through singing. This section provides a focused understanding of how the peers interacted during the literacy task. The individual case studies of each group of peers are presented separately, first in an effort to demonstrate how peers interact within the different instructional formats provided by the music teacher. After analysis of the data for each of the cases, a cross case analysis is presented to address the similarities and differences across cases. Integrated within both individual

cases and across the cases were the relevant statistical findings drawn from the quantitative phase of the study.

Prior to the analysis of the individual cases, an overview of the study is provided. This included a description of the participants, my role as a researcher during this phase of the study, and the theoretical considerations that apply to this interpretive case study. In addition to the overview, an understanding is presented of who these literacy learners are through interactions during the assessment period. This provided a better understanding of the participants and how they see themselves as readers.

Overview

Participants

During the 2006-2007 School Year, a total of 64 middle school students, in 7th and 8th grade, voluntarily agreed to participate in this study. These students were members of the fourth quarter Wheel Music Class (March 12, 2007- May 31, 2007). The Wheel Music Class was an assigned an elective class of new cohorts (mix of 6th -8th grade students) each quarter of the school year. However, within this sample of students in this study, there were no sixth grade student participants.

During the fourth quarter, there were four intact classes of Wheel Music students randomly assigned by classes to a treatment or control condition for the study. When the classification characteristics were compared (as noted in Chapter 3, Table 3) for the treatment and control groups, it suggested the groups were predominantly White low SES students. Male eighth graders represented a larger proportion for the treatment and control groups than their female counterparts or seventh grader peers. In addition, only a small percent of the adolescents received support services for learning or language needs.

The Role of the Researcher

I was trained to take observational field notes when I collected data for the National Longitudinal Evaluation of Comprehensive School Reform (NLECSR) for the David C. Anchin Center at the University of South Florida (USF) and the American Institutes of Research (AIR) during the 2003-2004 school year. In addition, I have taken two Qualitative Research classes at the University of South Florida.

For the current Interpretive Case Study, my role as a researcher was participant observer. Initially, I planned to observe more than participate by sitting in the back of the music classroom while observing and taking field notes. The reason for this decision was I did not want to have an adverse impact on the peer interactions during the literacy task of rereading through singing. Once the study began, I realized the impracticality of this plan to sit in the back of the music classroom while taking field notes on peer interactions. In order to capture the peers' interactions, I needed to move amongst the students and simultaneously take field notes.

Since this research study commenced during a new quarter of the school year, the students were accustomed to the presence of a researcher having never had this class or teacher prior to the study. Therefore, the students recognized me, and I was expected in the classroom. My presence did not deter them from their daily routine, as they greeted me by name whenever I came to the classroom.

Theoretical Considerations

Sociocultural Theory

Many variables influence reading performance of early adolescents' literacy learning within the context of the middle school classroom. Ivey (1999) contends early

adolescent readers are complex and multidimensional in their literacy learning. Cook-Gumprez, (1986) and Scriber and Cole (1981) suggest sociocultural theories of literacy occur as literacy is used in specific contexts for specific purposes and is socially constructed and constituted. The act of literacy is embedded in a network of social relations. Moje (1996) suggests that in the secondary content classroom the social context that shapes literacy practices is uniquely complex. Teachers and students in secondary classrooms move from class to class, teacher to teacher, and with a subgroup of peers. Teachers and students construct meaning about literacy and learning events based on values, beliefs, and knowledge, depending on the contextual situation. Additionally, teachers and students bring meaning to these interactions through their past beliefs, values, and knowledge during social interactions (Moje, 1996). Studies that are guided by broad theories as a social construction have focused on how social interactions influence literacy learning (e.g., Myers, 1992).

Moje also contends more research should investigate classroom interactions and how they play a part in shaping literacy practices. Sociocultural theories informed this case study, especially using Ryan's (2000) theoretical work on peer interactions.

Peer Interactions

Ryan's (2000) work investigated the research on peer groups' interactions, as a context for adolescent achievement, motivation, engagement, and socialization. In her analysis on the research of peer group socialization for the early adolescent Ryan (2000) theorizes peers generally interact three ways with one another. During early adolescence, the peer group becomes a prominent context for development (Brown, 1990). The school and classroom provide opportunities for peers to interact throughout the day. Ryan (2000)

reports “peer interactions consume significantly more time in adolescence compared to any other time in childhood” (p. 107). These interactions with peers can concern both academic (e.g., achievement) and nonacademic matters (e.g., engagement, motivation, self-efficacy, and interest). Ryan (2000) suggested three ways early adolescents generally experience peer interactions within the context of middle school: through information exchange, modeling, and peer pressure.

Information exchange occurs when adolescents have a discussion with their peers (Berndt, 1999). In an experimental study with eighth-grade students, Berndt, Laychak, and Park (1990) found that when adolescents had to make an academic decision, such as attending a rock concert or study for a test, they initially responded differently from one another. However, after discussing this dilemma with their peers, their answers were similar to their peers. This form of interaction could influence the early adolescent’s choice to partake in the literacy task presented by the teacher if it was used effectively.

Modeling is another form of adolescent peer interaction. This interaction refers to individual changes in cognition, beliefs, or affect, which are a result of adolescents observing their peers (Ryan, 2000). Observing a specific behavior a peer performs or listening to a peer voice, a certain belief can induce an adolescent to change their stance or adopt their peers’ behaviors or beliefs. Schunk and Zimmerman (1996) reported peer modeling influenced self-efficacy beliefs. In their study, they found early adolescents who verbalized difficulty with a task and then observed their peers have success with the same task, then believed they could complete the task. The early adolescent, when faced with a literacy task, may have success by observing their peers. Peer pressure is the third way the early adolescent interacts with their peers.

Peer pressure takes on the role of social reinforcement (Ryan, 2000). Brown, Lohr, and Eicher (1986) found beliefs and behaviors that are discouraged by the groups are not likely to be displayed; whereas, beliefs and behaviors that are positively received by the group are more likely to surface. Therefore, participation in the literacy tasks that the peer group positively received through this interaction could have a positive effect on the group's beliefs and decisions to participate by the group members.

Peer pressure may also play a role in how the peer group influences motivation. Brown, Lohr, and McClenahan (1986) reported peer pressure regarding school involvement is significantly correlated with self-reported behaviors and attitudes regarding school. Ryan (2000) recommended further research on peer interactions within a domain specific classroom may fill in the gaps in the literature. The above named recommendations from the research of Moje (1996) and Ryan (2000) are used to frame this study's qualitative component. Ryan's theory on three general categories of peer interactions framed the interpretive case study, along with Moje's recommendations that research on interactions within the setting of the content classroom should be studied to inform practice as to how literacy learning could be shaped.

Assessments

Prior to and after the experimental treatment, all 64 students were individually assessed in their reading performance in Fluency (WPM), Word Recognition (WR), Reading Comprehension (COMP), and Instructional Reading Level (RL). The students were told this assessment would not be a part of their personal records, and any information obtained was confidential so no one in school would ever see any results

from the assessments with their names attached to the scores. However, it was reinforced to the students that it was important to try their very best when reading.

I continued with an explanation pertaining to the assessment process and asked their permission to proceed with the assessment. It was explained that they would have two passages (sometimes more) to orally read and while they were reading I would be taking notes. On completion of their oral reading of the passages, they would be asked comprehension questions about what they read. Interestingly, what I observed and what the students provided through unprompted self-reports confirmed Ivey's (1999) contentions about the complexity of this population of heterogeneous middle school students of varying reading abilities

FCAT Reading Level Scores

The scores of Florida FCAT levels 1- 5 in reading are indicators, according to the state, of reading ability and performance. Levels 1 or 2 are considered below grade level; whereas, Levels 3 through 5 are considered at or above grade level in reading. There were 56 students out of 64 (28 in treatment and 28 in control groups) who had FCAT level reading scores in this study. Eight of the students did not have 2006 FCAT level reading scores for various reasons (e.g., relocated from another state).

Of the 56 students 40 (71%) were considered meeting grade level proficiency or above grade level, according to their FCAT level reading score. Students at this level are considered proficient readers and are not necessarily provided any individual support in their reading skills or strategies. Specifically, in this study 40 students out of 56 were determined to be proficient (FCAT level reading scores of 3 through 5) in their reading, according to results from the FCAT high-stakes tests.

However, when given the QRI-4 reading assessments, their instructional reading levels illustrated only 19% of the students (11 out of 56) at pretest were on grade level or above in reading. A combined mean score of instructional grade level in reading was 7.5 for all participants at FCAT levels 3 through 5. At posttest only 27% of the students (15 out of 56) were on grade level or above in reading with a combined mean score of grade level reading at 8.1 for all participants in both treatment and control groups. This suggested only 15 students out of the 40 students, determined by their FCAT level reading scores, were in fact meeting grade level proficiency in their reading at posttest.

Student participants who scored at FCAT level reading score 1 or 2 are considered below grade level in their reading. In this study, 16 students (29%) out of 56 scored at a level 1 or 2. Specifically, these students were determined to be below grade level proficiency in their reading, according to the results from their FCAT reading level scores. However, when given the QRI-4 reading assessments, co-scored for reliability (see Chapter 3), their instructional reading levels showed that 80% (45 out of 56) of the students (combining all FCAT levels) at pretest were below grade level in reading with a combined mean score of grade level reading at 4.94. At posttest 73% (41 out of 56) of the students were below grade level with a combined mean score of grade level reading at 5.49. This suggested that for the 41 students who were reading below grade level only 16 of the students are receiving remediation in their reading. It is therefore suggested that for the 56 students in this middle school 27% are proficient readers; whereas, 73% are reading below grade level.

These descriptive findings concur with the statistical results for the students stratified by their FCAT levels 3 (At level) and 4 and 5 (Above level). The statistical

results showed that the groups had different pretest scores on their instructional reading levels (RL) than the control group. Specifically, for the treatment group on Level 3 (*At*) had significantly higher pretest scores $p = .019$ than the control; whereas for control group Level 4 and 5 (*Above*) had significantly higher pretest scores than the treatment student $p = .002$. However, the statistical findings reported for both groups (treatment vs. control) on their pretest instructional reading level scores found no significant difference on the pretest scores $p = .677$.

In addition, the descriptive findings suggest according to the students' FCAT levels 3, 4, and 5 (at or above level), 71% (40 out of 56) of students were determined as meeting or above grade level in their reading based on FCAT level reading scores. However, when assessed using the QRI-4, only 27% were performing at or above grade level in their reading. This suggests that using FCAT reading level scores as benchmarks to determine instructional reading level do not appear to correlate to scores from the QRI-4 assessment. Amrein and Berliner (2002), overall, contend that "there is no compelling evidence from a set of states with high-stakes testing policies that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators" (p.54). Therefore, the use of a high-stake test scores alone can not account for the many variables associated with understanding the reading process and relating that to the characteristics of this group of early adolescent literacy learners and their fluent reading behaviors.

Coincidentally, these results align with the report from The National Assessment of Educational Progress (NAEP) (2005) that contends 73 % of eighth grade students perform below or at a basic level in their reading achievement. In addition, consistent

with NAEP results, Biancarosa and Snow (2006) reported, to the Carnegie Corporation, over 70% of adolescents struggle with their reading in some manner and, therefore, require differentiated and strategic instruction.

Fluency: Absence of Prosodic Reading

Fluency is a necessary aspect of successful reading as it allows readers to read with speed, accuracy, and proper expression (National Reading Panel, 2000; Rasinski, 2004). Rasinski (2004) contends reading fluency is a “bridge between two major components of reading- wording, decoding, and comprehension. At one end of the bridge, fluency connects to accuracy and automaticity in decoding. At the other end, fluency connects to comprehension through prosody, or expressive interpretation” (p. 1).

The students in both groups read their reading passages orally with speed and a high level of accuracy in word recognition, and yet they struggled with comprehension during the pretest assessment. Their oral reading was absent of volume, tone, pitch or any expression. There was no pausing at punctuation, rereading for clarification, or self-corrections made in 53 out of 64 students. Interestingly, 75% of the students asked prior to reading the passage, “how fast do you want me to read,” or “I need to read this fast,” (Assessment Notes April 2, 2007). Indicating for this group of literacy learners, proficient fluent reading was related to speed. My response to all the students was I want you to read at a pace so you can understand what you are reading and be able to answer the questions when you finish. Regardless of this suggestion at pretest, as noted in the statistical findings, there was no significant difference between the groups in fluency (WPM), word recognition (WR), comprehension (COMP), or instructional reading levels (RL).

However, at posttest, the treatment group of students outperformed their counterparts significantly from pretest to posttest in reading comprehension COMP and instructional reading levels RL. The oral reading of the students in the treatment group, although fast, had expression, pitch, and volume, unlike their counterparts. Specifically, 81% of the treatment students, or 26 out of 32, read their passage making self-correction, pausing at punctuation, and rereading phrases or sentences. Whereas, the control group of students, only 28% or (9 out of 32) of these students incorporated these same prosodic elements in their reading. This was particularly noted in the statistical finding showing a significant difference between the groups in COMP reading comprehension, $p < .001$, and instructional reading levels, $p < .001$. Rasinski (2004) contends, when reading, prosody is incorporated in the rereading with accuracy and automaticity then the student's comprehension will improve.

Students' Self- Reports on Their Reading Disposition

Self-reports by the students during the assessment sessions provided an opportunity for me to hear the students' perceptions of how they see themselves as readers and their personal relationship with the reading process. There were two self-reports that had an overwhelming frequency of responses from the students regarding their relationship with and disposition for the reading process. The first self-report was a dislike for reading. An example of this was, "I hate to read- there are no books in this school I like" (Assessment Notes, April 1, 2007). Specifically, 39% of the students (25 out of 64) made a statement similar to this response before the reading assessment even began. Interestingly, this concurs with Ivey's (1999) findings in her case study on the

early adolescent's attitude towards reading and book choice in the middle school classrooms.

The second self-report with the highest frequencies related to how cognizant the students were of their difficulties in reading comprehension. An example of this was, "I can never remember what I read" (Assessment Notes, April 1, 2007). Specifically, 30% of the students or 19 out of 64 students made reference to not being able to remember what they read. This was evident in the scores on reading comprehension for both treatment and control groups at pretest and posttest at the higher instructional reading level. The means for both groups was the same in reading comprehension (COMP) at pretest, 77%. The reading comprehension (COMP) significantly increased at posttest to 85% on the same instructional reading for the treatment group. However, at posttest on the highest instructional reading level, COMP was again the same for both treatment and control groups at 78%, even though the treatment students had increased their reading level (RL) over a year compared to the control group.

Biancarosa and Snow (2006), in their report to the Carnegie Corporation, contend that reading comprehension is an area of concern for the early adolescent. Some students may have trouble with decoding words accurately and with automaticity; whereas, others may read words fluently, but they do not remember what they read. In addition, other students may know comprehension strategies but do not have sufficient practice or opportunities for use. Biancarosa and Snow suggest this may be a result of limited understanding, support, and practice for strategies used to develop comprehension in the various content areas. Topping (2006) suggests that "even fluent readers will show dysfluency with text beyond their independent reading levels" (p.106).

Treatment and Control Group Data Analysis

The analysis of the data required qualitative analysis procedures. Patton's (2002) guidelines for content analysis recommend reading through the data at a specific time and making notes in the margins pertaining to specific notions about meanings. Moerman's (1988) suggestions for conversation analysis guided the analysis of peer interactions through conversations. In addition, Miles and Huberman's pattern analysis (1994) was used to code data and look for emerging patterns.

Observational field notes were taken during each 50-minute class session, twice a week for each of the four classes assigned to the treatment or control condition. Field notes were taken on a pad of paper during the Wheel Music Class periods noting time, place, attendance, and all the peer interactions during the observation. These observations focused on describing the relationship, if any, between the literacy task and music teacher assigned (rereading through singing) and focused on interactions (peer talk, peer modeling, and peer social reinforcement) among students who were singing using the interactive program, *Tune Into Reading*, versus the peer interactions among students who sang in the traditional music class.

Strauss (1993) recommended that to assist with this difficult process for beginners, researchers should develop a coding paradigm. The paradigm, which applies to this study, consisted of: (1) the literacy task (rereading through singing) assigned by the music teacher and (2) interactions among the peer groups during the literacy task assigned by the music teacher for the two cases (students using the interactive sing-to-read program and students in the regular music class).

Following a theory suggested by Ryan (2000), there are generally three ways that early adolescents experience peer interactions within the context of middle school: (a) through information exchange (discussion), (b) modeling (peer observation and imitations), and (c) peer pressure (social reinforcement). I used these three categories as preliminary coding categories and as a framework to focus my observations.

Information exchange refers to discussions and talk amongst the peers, capturing direct quotes from the various conversations that the peers exchanged during the literacy task: Peer 1 “How did you get the song to slow down,” and Peer 2, “Click on this button” (Observational notes April, 7, 2007). Peer modeling, on the other hand, refers to the act of peers observing one another and results in changes in behaviors or understanding within the student(s). This is achieved by describing the interactions during the literacy task that documents these changes: [He looked around the classroom for two minutes then he smiled and went back to playing the drums] (Observational notes, April 7, 2007). Finally, peer pressure occurs through social reinforcement, both negative and positive. Descriptions of peers’ accepting or rejecting behaviors exhibited by their counterparts through body language, facial expressions, smiling, or laughing during the literacy task: [T hit the drum wrong... M laughed...and then the class laughed...T turned red and put his head down] (Observational notes, April 7, 2007). Ryan (2000) suggested three categories became preliminary coding categories. They were then put into a matrix used for data analysis.

Field notes were reviewed daily after all the observations were completed. Initially, I would read through the notes three times to get a holistic sense of the data collected. Then the notes were bracketed and coded as one of the three peer interaction

categories. Units of data, conversations among peers or paragraphs that described peers observing or applying pressure to other peers, were bracketed and labeled as one of the three peer interaction categories. This was followed by transferring the bracketed notes to a matrix (Appendix B) with the three categories. The matrix was used to ensure observations did not stray from the focus of the study. Once the data were transferred, the difficult job of data analysis began. Figure 5 provides an example of the matrix used in this study

My first task involved typing the field notes from the observations of the Wheel Music Classes. The notes were typed-up daily after all classroom observations so the information could remain fresh. Once this task was accomplished, I began the difficult task of reading and analyzing the data. First, I read the field notes from the classes three times to gain a holistic sense of the data. Then, I returned to the data, bracketed the categories of peer interactions, and labeled them as information exchange, modeling, and peer pressure, so it could be transferred to the peer interaction matrix (Appendix B). I then read each line of the data in the matrix and highlighted units of meaning, patterns where repeated phrases and/or words occurred (Patton, 2002). Construct names emerged from these data. The construct names came directly from the data. One example that illustrates how this was done was a phrase that described peer modeling, “In the four corners of the computer lab, small groups of females look at one another and start to laugh softly, as they secretly glanced around the room.” This sentence was highlighted and was bracketed with the construct name, Peer Observation.

Once in the matrix, the data were further analyzed to determine the elements of peer interactions during the literacy task. After the elements were identified and assigned

construct names, they were added to the Construct Key (Appendix C). I used the Construct Key to be consistent with construct names from the emerging data but also added any new emerging constructs from the consecutive observations to the construct key.

The elements were then grouped according to the construct names. The elements with the construct names assigned were then cut-up and placed in a folder. The frequency of each construct was tallied to determine whether or not an element was emphasized during the peer interactions. The frequency calculations were followed by organizing the constructs into categories. Each category of constructs was placed on a bulletin board and further analysis determined the themes that emerged from these data. These themes were presented first as individual cases, and then a cross case analysis was conducted.

I repeated this process for 28 observations (14 observations for the treatment case and 14 observations for the control case), and then I analyzed these data again with the finalized Construct Key. To ensure that the qualitative phase of this study is credible, qualitative researchers with background in literacy were utilized as a second observer and conducted an analysis check of the data.

After the constructs were identified, they were grouped accordingly under a construct heading. There were two construct headings that emerged from the data for both the treatment and control group. They were *Group Characteristics of Peer Interactions and Peer Interactions During the Instructional Procedures*. The frequency for each construct was then calculated to determine the themes for the cases. These themes are presented first as individual cases then as a cross case analysis.

Integration of the Data

Priority was given to the quantitative approach. It looked at the statistical relationship between students who used the sing-to-read program, Tune Into Reading, compared to students who sang as part of their regular music program. Therefore, the analysis for this approach was done first to answer the first two questions of this study. However, concurrently, qualitative case study methods were used to better understand and describe the peer interactions occurring during the literacy task assigned by their teacher. The integration of the two types of data occurred within the qualitative findings section of this study. The quantitative results and qualitative description were triangulated mixing the quantitative results with the qualitative descriptions in order to provide a clearer picture and more fully answer the research questions

Treatment Group

Description of Classroom Computer Lab

A single door opens into a small rectangular sound-proof computer lab, located at the back left hand corner of the music classroom. Three quarters of the parameter of the room housed 15 permanent computer docking stations. An empty table, located at the front of the lab and a small table in the middle of the lab, was also used during the intervention for the remaining students. Laptops were placed there from a laptop bunker that stayed during the 7-week intervention. At each computer station, an individual microphoned sound-proof headset for each of the students was attached to the computers. Students would walk into the lab, sign-in, and retrieve their personal folders to keep a record of songs that they sang and recorded during each session.

Data Analysis

The data collected from the treatment group came from 14 classroom observations. The classroom observations took place in the music classroom during the fourth quarter of the 2006-2007 school year (April 2nd - May 15th, 2007), over the 7-week experimental treatment period. As previously noted, four classes were randomly assigned by class to the treatment or control conditions. Two classes were combined and became the treatment group, and two classes were combined and became the control group. Observations occurred twice a week for the treatment group from the week of April 2, 2007- May 15, 2007.

There were 20 constructs that emerged from the treatment group observational data. The constructs were tabulated to determine the frequency of each construct in the data. Constructs with a frequency count of five or less were not included in the analysis. Table 24 specifies the 9 constructs that emerged from these data with the highest frequency. Refer to Appendix C for a description of the constructs.

Table 24

Constructs from the Treatment Group Observational Data

Construct	Frequency
Extrinsic Motivation	34
Peer Observations	29
Peer Hierarchy	18
Peer Support	13
Autonomy	12
Intrinsic Motivation	10
Students' Perspectives of Alternative Text	9
Safe Risk –Free Environment	8
Disequilibrium	7

There were three themes that emerged from the treatment group’s classroom observational data. They were developed after a thorough analysis of the data, which included reading through the data at least three times for a holistic sense of the data, analyzing the data for meaningful units, developing constructs from the emerging meaningful units, and tallying the constructs for frequency. The themes that emerged were *Group Dynamics*, *Motivation*, and *Singability vs. Readability*. These themes encompassed the essence of peer interactions during the treatment groups’ use of the interactive sing to read program, *Tune Into Reading*. Table 25 presents these themes and the frequency with which they occurred in the data collected from the treatment group.

Table 25

Themes from the Treatment Groups’ Observational Data

Theme	Frequency
Group Dynamics	68
Motivation	56
Singability vs. Readability	18

Group Dynamics

Safe Risk-Free Environment. The music teacher was aware that this non-conventional alternative middle school task of rereading through singing made the students nervous. To help alleviate some of their apprehensions and fears, she would constantly walk around during the sessions and give both verbal (telling them it would be alright) and non verbal (pat on the back or a warm smile) support. The students did in fact approach the task with some nervous laughs, self-reports of their lack of singing ability, and intense observation of one another. The peers constantly looked around at one

another to be assured this was an acceptable activity in which they should partake. The following is an example from the observational notes taken:

He looked around the room, focusing on each person for over 20 seconds.

Then, his gaze stopped at a small group of males. He watched intently as the males were singing and softly laughing with each other. He shrugged his shoulders and turned back to his computer. (Week 1, April 3, 2077).

The potency of a belief that one can accomplish the task appeared nested in an environment that was safe and risk-free. Regardless of the music teacher's attempts to be supportive and understanding of the students' apprehension, there was a need to feel this task was socially acceptable by the group. It was only after observing their peers that the early adolescents would feel safe enough to take risks and partake in the activity.

Peer Observation. When the students entered the computer lab each session, they picked up their folders and took a seat at a computer station. Although the computer program generated scores for the students' pitch accuracy, the music teacher had the students write their individual scores in their folders. This was because she wanted to use their folders to dialogue with the individual students about their progress.

There was no assigned seating during the treatment sessions. The early adolescents came into the lab and sat next to their friends. Interestingly, not only would the peers sit with their friends, but they also separated themselves by gender. The females in the group sat in each of the four corners of the computer lab; whereas, the seventh grade students sat wherever there was an open seat, although they usually found a seat that corresponded with their own gender. In addition, the seventh grade students only sat down and took their seats after the eighth graders were seated. The center of the lab was

taken over by the eighth grade males. There was one small group of eighth grade males in particular that placed themselves at the center table in the middle of the computer lab.

Once the students settled down and were seated, the music teacher would give the students directions for the session. She reminded them of the procedures for using the program and to record their pitch scores in their folders. After answering any questions, she told them to begin. Although all the students would put their headsets on and select their songs, very few actually started to use the program. In fact, during the first eight minutes each and every student in the lab secretly glanced around the room and over their shoulders looking at the other peers. The following is an example taken from observational notes of the students who were supposed to have started using the program:

In the four corners of the computer lab, small groups of females look at one another and start to laugh softly, as they secretly glanced around the room. A seventh grade male turned his head to the left looking over his shoulder and then to the right. Two males look at each other and then behind where they were sitting, to the middle of the lab. A female bends over and turns her body sideways in the chair. Then she scratches her leg and at the same time scans the room. Two males are slouched back in their seats with a blank vacant look on their faces and appear to stare into space; however, their eyes glance sideways without moving their heads to observe the peers around them.

(Week 1, April 5, 2007)

The students appeared to comply with the music teacher's direction to start using the program. They would face their computers, put their headsets on, and open up the program to a song on the computer. However, it was only after they observed one another

that they actually started to use the program. Instead, they looked at one another to see if it was acceptable for them to begin. The focus of the entire group was on a small group of eighth grader males that sat in the middle of the computer lab, at the center table. This group of three male students appeared not to be cognizant of the rest of the group's observations. Their conversations and attention remained within the group with the other males they sat with at the table.

Peer Hierarchy. This small group of eighth grade males sat directly in the middle table of the computer lab. This group was unlike the other peers in the lab. They were not quiet, and they did not look around to see what the other peers were doing. Instead they would talk and laugh with one another. The other students watched and listened to this small group of eighth grade males intently, and when they spoke and made comments about the task, the rest of the group would stop, listen, and follow their lead. The following is an example of this group of peers interacting and how the rest of the students responded:

Male 1: This is pretty cool.

Male 2: I stink at singing.

[Other peers around the room shake their heads in agreement]

Male 1: It's better than doing work.

Male 3: Yeah... not like real work.

[Three females look at one another and shrug their shoulders]

Male 2: Okay, let's do it!

[Male 1: holds up his hand and makes a rock-roll sign]

(Week 1, April 5, 2007)

This group of males appeared to be the dominant characters within the group. There were no interactions with the other peer members in the lab; however, once this group of males settled into the task of rereading through singing, a domino effect occurred within the lab. When the dominant males started to use the interactive sing-to-read program, all the students in the lab simultaneously turned to their computers and started working, as if on a silent cue. This occurred at the beginning of the intervention and continued across the entire seven weeks.

Peer Support. Once the peers settled in and felt comfortable with the interactive program they would come into the computer lab and get right to work, following the lead of the dominant males. This silent cueing system remained intact throughout the seven week intervention as the peers would listen and model their behaviors and actions according to what they saw and heard from the group of eighth grade males. The interactions among the group of males appeared to be supportive and collaborate. They modeled for the other students this support system within the group as described in the following example:

Male 1: How do you slow this music down?

Male 3: Click on the tempo key and that will do it.

Male 2: Yeah... like this.

[Male 2 shows Male 1 how to do it]

Male1: Thanks. (Week 2, April 11, 2007)

Although the eighth grade males primarily conversed amongst themselves, they modeled a support system for the rest of students in the lab. Once they heard the males being supportive and collaborating with each another, several small clusters within the

lab did the same. The climate of peer support was important in order to keep the students continually using the sing-to-read program. However, regardless of this supportive behavior, in order to maintain sustainability, the program itself had to be interesting and motivating for the entire group.

Motivation

Extrinsic Motivation. Often throughout the intervention the peers would be heard talking about their pitch accuracy scores. The scores for pitch accuracy ranged from 0-100, and the students would often compare their scores with one another. These scores measured how accurately the students could sing and record themselves within a given pitch. The students were able to see their voice through real time pitch tracking frequency lines as they recorded themselves singing. The objective was to keep these lines within the pitch box above the words in the song. When they completed recording themselves singing, a score would pop-up on the screen. This would show the students how accurately they met the pitch and rhythm of the song. This game-like quality was interesting and motivating for the students. Furthermore, the music teacher had not given the students one particular score that they should get (e.g., 80%) with pitch accuracy. Instead, she wanted them to work to their individual highest potential by trying their best. Regardless, the peers would often be heard challenging each other as they got their scores as in the following example:

Peer1: What did you get?

Peer 2: I got a 60.

Peer 3: You did better then me...I got a 53.

Peer 1: That's good... that's a hard song. (Week 2, April 9, 2007)

The peers were excited and motivated by the scores they received. However, it was interesting to note that even though the conversation by the peers was competitive, it was at the same time also supportive. This may be as a result of the music teacher's reinforcement of trying their best instead of getting only one score that was acceptable. In addition, the climate of the class that was set by the dominant group of males modeled support and cooperation.

Autonomy. The music teacher would place several new songs in the students' folders contained on the program each week. The interactions amongst the peers focused on what new songs they had and which song they were going to sing first. There was a lively discussion each week with the new songs the students could choose from in their folders. Some students would have similar songs; however, the music teacher tried to keep it interesting by varying songs within the individual students' instructional reading levels. Although the peers were often heard discussing what songs they got with one another, when it came to deciding on which song to sing the choice was individual. The following is an example of two peers from the observation notes:

Peer 1: What songs did ya get?

[Peer 2 shows the list of songs]

Peer1: I got that one too!

Peer 2: I am gonna do Home On The Range first.

Peer 1: Not me I'm gonna do this one.

[Peer 1 points to a song -then they both turn to the computers]

(Week 3, April 16, 2007)

The peers appeared to be motivated by the different songs they got each week. Although this choice of material was controlled (the teacher placed songs in their folders), it appeared sufficient to keep the students interested and motivated in the interactive program. In addition, even though they would discuss what songs they each got, they appeared to be comfortable with choosing what song they wanted to work on individually. The opportunity for choice appeared to contribute to holding their interest and keeping the students motivated.

Intrinsic Motivation. The game-like quality and different materials afforded to the students were contributing factors to their continued use of the interactive sing-to-read program. There was a shift, however, that occurred around the end of the third week of the intervention, when these students' motivation became internalized. No longer were the discussions about what score they had compared to their peers, instead they became individually focused, self-regulated, and engaged in their own achievement. They would be completely engrossed with their own songs and pitch scores regardless of what was happening around them. One example of this is cited below:

He was focused on rereading the song.

This was the fourth time he recorded himself singing.

He had the screen up that displayed his vocal tract.

He used his finger to align where he was off pitch.

He went back again and reread the song

[He moved back and forth in his seat, nodding his head to the beat, and tapping his foot to the music]

Finally, [he sub vocalized] I got a 90! (Week 4, April 25, 2007).

The above example was representative of the students in this intervention. The shift from motivation that was extrinsic, to getting a high score and competing with the other peers, shifted to a form of internal competition. The students became focused on the task and interactive with their own learning. As noted in the above example, this student was rereading, applying strategic processes, and regulating his learning for his own purpose. Interestingly, his focus appeared to be on comprehending the rhythm and beat of the song.

Singability vs. Readability

Twice a week for seven weeks, the students would enter the computer lab to use the interactive sing-to-read program, Tune Into Reading. In the beginning of the intervention, the music teacher worked with the whole group, giving direct instruction on how they were to use the interactive program. Along with discussing the protocol use of the program, she made only one song accessible for all students during the first two sessions of the intervention. She chose Hot Cross Buns because as she explained to the students:

I have put only one song in each of your folders on the program.

It is the same song for everyone to try. I picked this song because it has a steady beat and the words repeat themselves. Therefore, you will be able to feel comfortable while you are learning to use the program. (Week 1, April 2, 2007)

It appeared the music teacher felt that by using this song because of its low readability level (second grade) and limited change in octaves, pitch, and rhythms, the students would be able to concentrate on learning how to use the program and not have difficulty with singing and recording the song. In addition, the students would have

success in their pitch accuracy scores when they sang and recorded the song because of its easy accessibility for the diverse reading levels within this group.

Students Perspectives of the Alternative Text. While watching the students use the sing-to-read program, it appeared the students were adept at using the computer. They could easily manipulate this digital text, by adjusting the songs' speed, page size format, and much more. They would show one another some of the different difficulties that they had encountered and how to work around them. They were less sure, however, about the genre of rereading through singing, particularly matching the rhythms, pitch, tempo, and beat of the songs to the words they were singing.

Once they were comfortable with using the program, a little more than half of the students (approximately 59%) would skip the procedure for listening to the background music and rereading the text silently. Then, they would complain to the music teacher about their low pitch accuracy scores. One session during the third week of the intervention, the music teacher went over the importance of listening to the background music and rereading the song silently several times. She told the students that whether you are reading a book, or singing a song, you have to have the beat in you head. The only way you get the sounds in your head is by practicing. This helps you to know when you stop and take a breath, or stress certain words. In addition, you would know how fast or slow you should read the words, followed by the sentences, and finally sing the song in its entirety. This will help you understand what you are reading or singing, and it will improve your pitch accuracy scores. However, the students were not convinced that this was related to reading a book as shown in the following excerpt:

Peer 1: Yeah but singing is not like real reading

[The group shakes their heads in agreement]

Peer 2: Real reading is like a text book

Peer 3: Yeah you read the book and answer the questions

Peer 4: This is good for singing-but not reading a book

(Week 3, April 16, 2007)

It appears the students' view of music as an alternative text was not a task related to reading. Their perception of what constitutes real reading is ingrained in their school experience. However, the music teacher continued to make a point about the importance of prosody and rereading, whether you are reading a book or singing a song.

She suggested to the students they conduct an experiment. She asked the students to record Hot Cross Buns (a song they were all very familiar with) first with the background music and then without it. Then, she asked them to reread the song first three times while you listen to the background music, and then record themselves with the music and then without. After you record yourself, the music teacher told the students to write down their scores with the background music and then without. When the students finished, she wrote the scores up on the board and averaged the scores with background music and the scores without. What they found was that the average pitch scores without background music was $M=25$; whereas, the average score with the background music was $M=73$ for the students. After the students finished, she asked them to make comments on what they found. The following is an example of the comments made by the students:

Peer 1: I guess we need to reread stuff to remember how it sounds.

Peer 2: The beats help us sing the song and read the words.

Peer 3: I guess I can see how this could help... when you are reading books too.

Peer 4: Like if you're reading and you don't stop you like can't remember.

Peer 5: The music in the background kinda helps you to reach the highs and lows.

(Week 3, April 16, 2007)

The students were provided with an opportunity to see the importance of rereading and the prosodic features of text. They recognized the purpose of having the rhythm and beat in their heads helped them not only with their pitch scores but also when reading for meaning. In addition, their perceptions of what constituted a real reading task, regardless of its alternative format, were brought to the forefront. This was accomplished because the music teacher took the time to show the students why it was important to reread text and listen to the background music. She provided clear goals and objectives explaining why they were following the procedures for the literacy task rather than just assigning and telling the students to just follow the directions. As a result, the students understood the purpose and could adjust their view of what and why they were being asked to do during the task.

Disequilibrium. As the songs became more difficult the students often complained to the music teacher about how hard it was for them to sing the songs and get a good score. The music teacher sat with the individual students and had them sing the songs so that she could provide assistance. Often the students would discuss with one another how hard it was to sing the songs. However, no matter how difficult the singing got the

students would persevere. The following is an excerpt taken from the observational notes between two peers:

Peer1: I can't get this song.

Peer 2: Yeah it is really long.

Peer 1: I need to read it a ton of times before I record it.

Peer 2: Yeah... you could slow it down too.

Peer 1: I'm gonna try it again

Peer 2: Go for it! (Week 4, May 3, 2007)

As with any new learning, in reading as the task became more difficult the students need the stamina to continue. The disequilibrium that occurs with all new learning and the perseverance to continue is what makes learning successful. The example above shows the students were aware that the material was getting more difficult, yet they still opted to continue with the task of rereading through singing.

Summary of Results for Treatment Condition

Many of the constructs that evolved from the observational data are well documented in the literacy research (e.g., Guthrie & Wigfield, 2000) as effective practices to meet the needs for the early adolescent learner. Specifically, the students should be motivated and engaged in the literacy task presented, so they could achieve academically. Interestingly, the shift from an extrinsic form of motivation (motivated because of reward or punishment) to an intrinsic motivation (motivated because they want to do this above anything else) was when it appeared that the students really became engaged in the task.

The opportunity to make a choice of text also supported the student's motivation. Each week the choice of songs provided autonomy for the students and appeared to increase their motivation to continue using the program. In addition, although this was an independent task, there was considerable student collaboration and support in a safe risk-free environment. It was interesting to note, however, the peers' social system that was in place.

The peers' hierarchy and the passive aggressive stance taken by the students, as to whether or not they should partake in the task, are not as well documented in the research (Ryan, 2001). The students "buy-in" was nested in whether the dominant characters supported or rejected the task assigned. This silent cueing system should be considered when instructing this group of literacy learners. The outcomes could have been very different if the students influenced by the dominant characters had rejected the sing-to-read program.

The use of the alternative textual format and the genre it delivered was interesting to see and hear the student's perceptions as to what constitutes real reading. It was just as interesting to see the perspectives change about the task of rereading and the place that prosody has in understanding text. After the music teacher showed them how prosody and rereading affects their reading and singing, it was observed the students would constantly go back and reread the text before recording. Interestingly, when the students were provided with clear objectives for the task by the music teacher, there was a mutual understanding of the expected outcomes. In addition, the definition for alternative text in this study changed to include not only the format (digital), the genre (songs), but also the perspectives of the students as explained in more detail in Chapter 5.

These descriptive findings substantiate the statistical results previously reported in this chapter. That is when the students are motivated, have choice of text, have diverse and interesting textual formats, opportunities for peer collaboration, and understand why they are doing the literacy task, their academic achievement will improve (Guthrie & Wigfield, 2000). Specifically the treatment group of students increased significantly from pretest to posttest in fluency (WPM) $p < .001$, word recognition (WR) $p = .009$, reading comprehension (COMP) $p < .001$ at the same instructional level attained at pretest. In addition, at the increased reading level from pretest to posttest the treatment students increased in their instructional reading level $M = 1.13$ years within this seven week intervention.

Interestingly, at this increased reading level as disequilibrium occurred and the students were building the stamina to sing more difficult songs, their mean scores in all areas (WPM, WR, and COMP) declined. Although the students had increased in their instructional reading level from pretest to posttest, when comparing scores on the same instructional reading level attained at pretest to posttest scores at the highest reading level their scores decreased. Specifically in: (a) fluency (WPM), at the initial posttest $M = 160$ wpm to $M = 147$, at the increased reading level posttest, (b) word recognition (WR), initial posttest $M = .99$, to $M = .98$ at the increased reading level posttest, and (c) comprehension (COMP) at the initial posttest $M = .85$, to $M = .75$ at the increased reading level posttest. This suggested that, as the early adolescents in the treatment condition increased in text difficulty, their fluency (WPM), word recognition (WR), and comprehension, (COMP) shifted from a fluent expert reader at one level to a dysfluent reader (e.g. Topping, 2006) at a higher level.

Control Group

Description of Classroom Routine

There were 32 students in the control group during this study. Students would enter the classroom and choose their drum or other instrument that they used for the lesson. Then, they would get a chair and place it in one of the three semi-circle stadium steps in the classroom (Chapter 3 room description). The music teacher would bring the class together with a beat of her drum in the center stage of the music classroom. The students would echo back the beat and class would begin.

Data Analysis

The data collected from control group came from 14 classroom observations. The classroom observations took place in the music classroom during the fourth quarter of the 2006-2007 school year (April 2nd - May 15th, 2007) over the 7-week experimental treatment period. As previously noted, there were two classes randomly assigned by class to control conditions. Two classes were combined and became the control group. Observation occurred twice a week for the control group from the week of April 2, 2007- May 15, 2007.

Observational field notes were taken during each class session twice a week, during the 50 minute class periods for each of the 2 classes assigned to the control condition. The field notes were taken on a pad of paper during the Wheel Music Class periods noting time, place, attendance, and all the major character interactions during the observations. The focus of these observations was to describe the relationship, if any, between the literacy task the music teacher assigns (rereading through singing) and the

peer interactions (e.g., peer talk, peer modeling, and peer social reinforcement) among students who are singing in the regular music classroom..

There were 15 constructs that emerged from the control group’s observational data. The constructs were tabulated to determine the frequency of each construct in the data. Constructs that had a frequency count less than five were not included. Table 26 specifies the 6 constructs that emerged from these data with the highest frequency. Refer to Appendix C for a description of the constructs.

Table 26

Constructs from the Control Group Observational Data

Construct	Frequency
Extrinsic Motivation	34
Alternative Approaches to Singing	24
Dominant and Vulnerable Peer	21
Fake Rereading	18
Disengaged	13
Peer Leaders	8

There were three themes that emerged from the control group’s classroom observational data. They were developed after a thorough analysis of the data, which included reading through the data at least three times for a holistic sense of the data, analyzing the data for meaningful units, developing constructs from the emerging meaningful units, and tallying the constructs for frequency. The themes that emerged were Engagement, Group Formats and Reading Strategies. These themes encompassed the essence of peer interactions during the control groups’ singing during their regular music period. Table 27 presents these themes and the frequency with which they occurred in the data collected from the control group observational data.

Table 27

Themes from the Control Groups' Observational Data

Theme	Frequency
Engagement	58
Group Formats	29
Reading Strategies	33

Engagement

Alternative Approach to Singing. The music teacher wanted her students to be engaged and involved in the Music Wheel Class. She felt learning music theory and different aspects of singing for this group of students would not hold their interest. Specifically, this was because this group of students had not chosen singing as an elective, instead they were assigned to this elective. Therefore, she decided that a hands-on interactive alternative format would be more successful. The music teacher decided to use a drum circle. Not only would it build a sense of community for the students, using drums would also involve them in singing and creating their own music.

Initially, a simple drumming sequence was taught to the students, and then this was followed by learning the multiple stanzas of the three songs during the seven week intervention. In order for the students to learn the drumming sequence, the music teacher modeled the pattern of beats, and in turn, the students would echo in response the same pattern. She started very slowly at first and then would increase in speed. The students appeared to be engaged as they would follow her every beat. The following is an example from the observational notes:

The students would hit their drums echoing the music teacher.

Their backs were arched up straight. If they lost their place, they would stop and tap their foot until they caught the beat. Their faces were serious and intent on following the lead of the teacher. (Week 1, April 5, 2007)

The students were engaged and on task as they were echoing the music teacher's drumming sequence. They were focused and appeared to comply with the music teacher directions. The use of the alternative approach to teaching singing appeared to hold their interest and keep them engaged and focused on the task. There were no interactions among the peers while they were drumming, instead the group had remained serious and focused. Their attention was on concentrating, observing, and listening to the music of the drum sequence played by the teacher. Only after they finished did you hear the students laugh or make comments to one another.

Extrinsic Motivation. Once the students were able to perform the simple drumming sequence, they were taught the songs to accompany it. The goal was to keep the rhythm and beat of the song while singing by drumming. In order to assess if the students had accessed the song and its corresponding drum sequence, the music teacher had the students take turns in small groups and perform for the rest of the class. This round robin routine by the small groups was motivating and highly competitive for the students. When all the groups had a turn, the peer interactions were often comments regarding who performed the best during the challenge. The following excerpt between two groups was an example of this:

Group 1 Peer: We did better than you!

[Group 1 members cheered]

Group 2 Peer: No way... you guys messed up big time!

[Group 2 members shout yeah yeah]

Group 1 Peer: How did we mess up... your crazy

[Music teacher stops the interactions]

(Week 2, April 9, 2007)

The students were motivated by the challenge of competing with one another. The prize of being the best was the goal. During this competition, each group of students were actively engaged and tried to do their best. This was a motivating activity for all of the groups, and the reward was to perform, be the best, and get it right.

Group Formats

Dominant and Vulnerable Peers. Instructional delivery for this group of students was primarily accomplished through a whole group format. Each session the students would follow the same procedures. They would sing and play the drums echoing the music teacher. During the session, the music teacher stopped on occasion and addressed the students if she heard the drumming or singing off key. The students often made comments to one another, blaming them for making a mistake. It was the more dominate students in the group who addressed the more vulnerable students. However, these exchanges were not loud enough for the music teacher to hear. Instead the exchanges were accomplished secretly and critically as they blamed one another for making mistakes. The following is an example of one dominant peer admonishing a more vulnerable student for making a mistake while drumming and singing:

Peer1: You made the mistake.

Peer 2: No I did not... shut-up.

[Peer 2 makes a nasty face at Peer 1]

Peer 1: Yes you did I heard it you...jerk

[Peer 2 turns red and puts her head down] (Week 3, April 16, 2007).

Although the teacher did not single out any one student for making a mistake, the peers would blame one another. The criticism occurred often throughout the sessions between dominant and vulnerable peers. However, because it was done secretly, the teacher was not aware of what was happening. In addition, when the vulnerable peer was admonished, the other peers seated near or around the student who had been blamed for making the mistake did not say a word. Instead, they would look at one another or look away when this happened.

Peer Leaders. There was two occasions during the seven week intervention that the peers broke-up into small cooperative groups. The task was to create a new drum sequence that would accompany the song they learned in class. The music teacher selected the students for each group and then told them they had 30 minutes to complete the task. As the students were getting ready to join their groups, they were told that when they were done they would perform their creation for the rest of the class.

Once the students were in their groups, there was one dominant peer who would take control and lead the rest of the group. The peer leaders were self-designated eighth graders; however, they were not of a particular gender. When the peer leaders were in groups, they would take over and direct the other group members. They organized and managed the other peers, so the task assigned was accomplished. In addition, students in the group who did not cooperate were reprimanded because the goal was to complete the

task assigned. The following is an example of the peer interactions within the small groups:

Peer Leader: Okay, let's start with a high drum beat.

Peer Member 1: Let's hit it twice on high.

Peer Member 2: Sounds good, let's try it.

Peer Leader: M are you with us?

Peer Member 3: What if we hit the side like this.

[Peer 3 demonstrates for the group]

Peer Leader: Okay let's try it.

[She stops and speaks to M again]

Peer Member 4: That sounds good.

Peer Leader: Okay let's sing it with the song...Go.

(Week 4, May 1, 2007).

When the students were in the small cooperative group formats, there was a peer that assumed the role of leader. This was not an assigned position, instead it was allowed position by the rest of the group. The leader managed, organized, and kept the group focused to complete the task. Although the leader was either male or female, they were the dominant characters within each of the groups. In addition, even though the decision making appeared to be collaborative pertaining to the creation of the drum sequence, the final approval of what and how they would perform the drum piece was made for the group by the peer leader.

Reading Strategies

Fake Rereading. During the intervention there were three songs taught to the group. Initially, the music teacher put the song in its entirety on the overhead projector for the group to read. Then, she went over isolated vocabulary words she felt the students needed explained, so they would understand the song. This was followed by the music teacher's use of modeling each of the stanzas of the song, and in turn, the students would chorally sing and echo back what she sang. Finally when the music teacher felt the students were able to sing the song, she would have them play the accompany drum piece to go along with their singing.

For each session, the students would reread the song by singing each stanza and playing their drum sequence. The drumming supported the students' singing by providing the prosody needed to keep the rhythm, volume, and pitch of the song. However, even though the students knew the word to the songs as they reread through singing each session, often the teacher would stop the group and state she could not hear their voices. The following example is taken from the observation notes as the peers were charged with rereading (re-singing) the song:

The students were playing the drum sequence and moving their lips as if they were singing. Of the 20 students in the group, only about 6 were actually singing. The music teacher would stop the students and address the group reminding them to sing. However, this fake rereading through singing continued.
(Week 4, May 1, 2007)

However, this fake rereading through singing continued. Although the students knew the words to the song and the corresponding drum pattern, when it came time to

reread (re-sing) the song, they would play the drums but not sing. The more comfortable they became with what was required, the more they did not have to focus on what was happening. Even with addition of new songs or drum patterns, the learning became routine, and the students faked rereading (re-singing) the songs (Tovani, 2000).

Disengaged. The songs and drum patterns were taught to the student by having the students echo what the music teacher played and sang. The students would first listen to the teacher and then chorally sing and echo the back the stanza and the corresponding drum pattern. Towards the end of the seven week intervention, often the students would be seen daydreaming as they went through the motion of singing and playing their drums. Some students would be silently whispering to one another, and still others would be playing around and making up their own drum patterns. The following is an example from the observational notes on how the students were performing during the daily session towards the end of the intervention:

He would hit the drum while looking at the door. She was singing and looking straight ahead however, when the teacher spoke she became startled. They were whispering to one another exchanging ideas about what to wear to the dance. The two young males were laughing softly pretending to play their drums with another pattern. (Week 5, May 3, 2007)

It appeared the students had shifted from being very focused and engaged initially to more automatic in their response to the music teacher rereading through singing. They disengaged from the task they were performing. The learning became routine and presented the students with no struggle, challenge, or motivation to continue. They were off task and not engaged with the task the music teacher had them perform.

Summary of Results for the Control Group

The students in the control group appeared to be motivated and on task when they were using the alternative approach to rereading through singing, initially. The use of the drums appeared to hold the students interest. In addition, the opportunity to create their own drum sequence was extrinsically motivating. They were motivated to be the best and to sing the songs and reproduce the drum patterns correctly.

This was apparent with the light-hearted interactions as they competed with each other during the round robin performances as the music teacher assessed their learning. This competitive banter appeared to be extrinsically motivating and engaging for the students. It kept them focused and aligned with the objectives of the lesson. However, what started out as light hearted competition soon turned to critical analysis between the peers. They would blame one another for mistakes made during the performance. This was done secretly without the music teacher's knowledge of the interactions.

The interactions became uncaring and unsupportive as they would blame each other for not performing correctly. Although the music teacher did not single out a peer for making a mistake, through her actions she reinforced there was only one right way to perform the drum pattern and sing the song. This was done by stopping the class and trying the procedure again, often. The dominant peers blamed the more vulnerable peers while the other students would avoid becoming involved or supporting the peer needing support. These dynamics occurred consistently during the interventions, and even though the music teacher wanted the students to have a sense of community by using the drum circle, this appeared not to be the case.

There were students in the control group that immediately took on the role as peer leaders, even without that role being designated to them. It appeared to be assumed by the other members in the group that this person was in charge, and there was no questioning this position. In addition, decisions made in the small group format did appear to be collaborative; however, the final decision of what to do and how it was to be accomplished was determined by the peer leader. This position was an expected and allowed by the rest of the peers.

The use of rereading through singing was accomplished by having the students echo chorally back the drum patterns and the song lead by the music teacher. Although the students were engaged and motivated initially, they became disengaged with this routine. The data suggested the students became complacent in the task when the learning became unchallenging.

These descriptive findings concur with the statistical findings previously reported in this chapter. That is when the students become unmotivated, and disengaged, and the classroom environment does not provide opportunities for peer collaboration, their academic achievement will not improve (Guthrie & Wigfield, 2000). Specifically, the control group of students did not increase significantly from pretest to posttest in fluency WPM; $p=.219$, word recognition WR; $p=.379$, reading comprehension COMP; $p=.170$ at the same instructional level attained at pretest. In addition, at the highest reading from pretest to posttest, the control students did not increase in their instruction reading level from $M=5.58$ at pretest to $M= 5.77$ at posttest within the seven week intervention.

Cross Case Analysis

After a thorough analysis of the data for each of the cases, there were similarities and differences across cases. In order to capture the multi-dimensional and complex nature of peer interactions, these similarities and differences were described through the theme *Social Systems* that appear to capture the essence of peer interactions for the treatment and control groups. Specifically, the social structure of the peers through their interactions appeared to influence the task of rereading through singing. Within this theme, there were four constructs embedded. They were: (a) *Peer Positions*, (b) *Instruction Expectations*, (c) *Alternative Approaches to Tasks*, and (d) *Reading Strategies*. Therefore, analysis for this section will describe how the two groups displayed similar and dissimilar characteristics of peer interactions within this theme and across these constructs.

Social Systems

Peer Positions. The treatment and control group had in place a social system that positioned some of its peer members in the role of dominance over the other peers. These dominant characters held this position, and the other peer members allowed them to assume it. Both groups contained this two class system where a small group or individuals lead the rest of the group pertaining to acceptable social behavior. Interestingly, however, within the treatment group and the control group, the interactions from the dominant peers with the other peer members were accomplished very differently.

Within the treatment group, the peers modeled behaviors or talk that resulted in the rest of the peers imitating their behaviors. There were no discussions with the other

peer members only among the small group of eighth grade males. Whereas, within the control group, the dominant peers of various gender (male or female) specifically directed the other members to conform to a certain behavior they deemed socially acceptable. This may have been as a result as to how these dominant peers interpreted what was expected of them through the instructional delivery provided.

Instructional Expectations. The role of dominant peers remained constant within the treatment and control groups, during the intervention. However, the instructional expectations afforded to the groups by the music teacher were very different. These expectations appeared to be interpreted by the dominate peers and then reinforced through their interactions with the rest of the peer group.

What appeared to be expected of the peers in the control group was that there was only one right way to perform the singing and drumming. These expectations were modeled to the peers as the music teacher would stop the singing and drumming several times daily during each session and tell the students that some people are off key, try it again. In turn, the dominant peer would admonish the vulnerable peer for making a mistake. However, the treatment group was expected to try their best. The music teacher would often remind the students to try their best and not to worry about their pitch scores. The dominant males in the treatment group would encourage and support one another, modeling collaboration to the other peers. Therefore, it appeared that within both groups the dominant peers interpreted what was expected of them and then in turn reacted to these expectations through their interactions.

The control group peers were expected to perform correctly and accurately as modeled by the music teacher. Since there was only the right or wrong way to perform

the song and drum sequence, the interactions might have been interpreted by the dominant peers and conveyed to the rest of the group in this manner through their interaction. However, the treatment group was allowed differentiation through the instructional delivery. Therefore, the dominant peers appeared not to be compelled to take on the task of reinforcing group accuracy; instead they became a group member while still maintaining their position among the group.

Alternative Approaches. Both groups found the alternative approaches to learning motivating and engaging. This was apparent during the interactions within the groups. The light-hearted competitions through interactions were documented in the data as the groups led by the dominant peers would either through discussion or modeling set the climate of motivation for rest of the group. However, as the intervention continued, a shift occurred within both of the groups as to their motivation for these alternative approaches to the task of rereading through singing, as did the role of the dominant peers.

The peers within the control group became disengaged towards the task, including the dominant peers within the group as the sessions progressed in time. The data suggested that around the fourth week as the sessions continued the peers would daydream, talk, and entertain each other during the sessions. The motivation levels shifted from highly motivated to complacent. In contrast, the treatment group of students' motivation shifted from external motivation to internal for all the peers, including the dominant peers. They became engaged in the task and self-regulated in their learning. This might have occurred as a result of how the strategic process for reading unfolded.

Reading Strategies. Fluency instruction for the students was the same in both groups. The music teacher used repeated readings of the songs, while embedding

prosodic features of text. The students reread (re-sang) their songs three or more times each session and each group was supported with the prosodic features of songs through background music, rhythm, tempo, pace, and volume. The control group used their drums and followed a modeled example of prosody from the music teacher; whereas, the treatment group prosodic elements were contained in the background music from the *Tune Into Reading* program. This process as reported in the literature (e.g., Samuels & Farstrup, 2006) should improve the students' automaticity (WPM), accuracy (WR), reading comprehension (COMP), and instructional reading levels (RL) for both of the groups. However, this was not the case for these two groups.

Summary Cross Case Analysis

The descriptive findings did support the statistical results previously reported; however, these findings did not concur with the findings cited in the literature in entirety. Within the groups, the treatment group displayed a statistically significant difference in fluency (WPM) $p < .001$, word recognition (WR) $p = .009$, and reading comprehension (COMP) $p < .001$; whereas within the control group, they did not in fluency (WPM) $p = .219$, word recognition (WR) $p = .379$, or reading comprehension (COMP) $p = .170$ on the same instructional level attained at pretest. This suggests that there may be other contributing factors for the students to be fluent readers. One potential factor, as suggested in the descriptive findings, is an environment that is safe and supportive and instruction differentiated to meet the needs of all of the students.

In addition, the statistical findings show that when the treatment and control groups were compared the treatment, students had a significant increase in reading comprehension (COMP) at $p < .001$ and instructional reading level (RL) $p < .001$ as

compared to the control group. According to Rasinski (2004), it is the prosodic features of text genre that assist in reading comprehension; therefore, these findings suggest prosody was a contributing factor in the increase of reading comprehension for the treatment students. Yet, this appeared not to be factor for the control group. The descriptive findings suggest the treatment group internalized their learning as the sessions continued; however, the control students disengaged from task. This could be interpreted to mean that prosody needs to interact with the learning and not be passive so that reading comprehension can occur. Therefore, although both groups were following the protocol for reading fluency improvement, the students singing through rereading alone, as in the case of the control, did not improve reading comprehension, which is the goal of fluency instruction.

Chapter Summary

In this chapter, I answered the three research questions after an in-depth analysis of the statistical and observational data from students in the treatment condition, using the interactive sing-to-read program, *Tune Into Reading*, and their counterparts singing as part of their regular music program. The statistical analysis was conducted on the first two questions initially investigating the difference in reading outcomes in fluency (WPM), word recognition (WR), reading comprehension (COMP), and instructional reading levels (RL). I administered pretest and a posttest measured by the *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006) and compared the posttest scores with the pretest scores to determine if students in the experimental group gained significantly over their counterparts in the control group. Initially, the students were assessed at posttest with a reading passage on the same instructional level attained during

the pretest. This was followed with statistical analysis at the highest instructional reading level attained by the students.

Then, the students were grouped according to their FCAT reading level scores. FCAT level reading scores (level 1-5) range from highest level (5) to lowest level (1). The treatment and control groups were stratified according to their FCAT level as: (a) Level 4 or 5 *Above* grade level, (b) Level 3 *At* grade level, and (c) Level 1 or 2 *Below* grade level. Once the students were grouped, percentages of students for each group were calculated for the students at each level. The percentages showed an equal distribution of FCAT level reading scores between the two groups; however, eight students were missing FCAT level reading scores (four treatment and four control). Therefore, only 56 out of 64 students' data were analyzed.

Finally, question three investigated the peer interactions that occurred during the study intervention. There were 14 rereading through singing sessions for the students in the treatment and control groups that occurred for 50 minutes each, twice a week, over a seven-week period. During these sessions, observational notes were taken on the peer interactions that occurred during these sessions. These observations focused on describing the relationship, if any, between the literacy task the music teacher assigned (rereading through singing) and the peers' interactions during the task. There were two cases in this study. The treatment students were singing using the interactive program *Tune Into Reading*, and the control students were reading through singing in the traditional music class.

Priority was given to the quantitative approach because it looked at the statistical relationship between the sing-to-read program, *Tune Into Reading*, used by the treatment

group compared to their counterparts in the control group rereading through singing in the regular music class. The reading outcomes from the students of varying reading abilities were measured by the QRI-4. The analysis for this approach was executed first to answer the first two questions of this study. However, concurrently qualitative case study methods were used to better understand and describe the peer interactions occurring during the literacy task assigned by their teacher. The integration of the two types of data occurred within the qualitative findings section of this chapter and used a triangulation strategy to interpret the findings. This integrated the statistical results with the descriptive findings in order to answer the research questions of the study.

The study findings indicated that the middle school students of varying reading levels significantly improved in their reading fluency scores through the use of the interactive sing-to-read program *Tune Into Reading*, compared to the group who were rereading through singing in the regular music classroom. In addition, prosody appeared to have a direct connection to reading comprehension. Furthermore, the use of the interactive program provided opportunities for differentiated reading level achievement. Finally, group dynamics highly influenced the early adolescent's motivation, engagement, participation, and successful outcomes in reading fluency.

CHAPTER FIVE: DISCUSSION

Chapter Five provided a discussion of the study results. There were five sections within this chapter. The first section summarized the study. The second section described the conclusions and implications derived from the research findings. The third section discussed the contributions this study makes to the existing body of knowledge on reading fluency with middle school students of varying reading abilities. Along with the discussion on reading fluency, a discussion of the findings related to the sociocultural interactions during the literacy task between the peers was included. Recommendations for practice derived from the research findings and the study's conclusions and implications were in the fourth section. Finally, the fifth section provided suggested recommendations for future research.

Summary of the Study

Fluency research suggests a fluent reader is one who can read a text with automaticity, accuracy, and proper expression, while viewing comprehension of text as the ultimate goal (LaBerge & Samuels, 1979). The methodology most noted in the literature to support fluency instruction is the process of rereading text, three or four times. Rereading affords the students quicker, more accurate, and better sounding reading. The literature on fluency also suggests a fluency model should be provided so students can hear proficient oral reading that captures all the elements of what fluent reading sounds like. Rasinski (2004) contends that utilizing a text with naturally

embedded features of prosody, such as poetry, speeches or singing will assist with building fluency in the readers. However, the assumption is often made that by the time most students enter middle school, they are fluent readers and comprehenders across a variety of texts (Alvermann, 2001). This is especially true of those students deemed proficient readers, determined by their yearly standardized test results. As a result, fluency instruction is often only provided to the students deemed less than proficient in their reading, according to the high-stakes test results.

The purpose of this concurrent mixed methods study was to investigate rereading through singing with two groups of heterogeneously grouped middle school students within a music classroom. The two groups were randomly assigned by class to a treatment group (n=32), that used an interactive sing-to-read program, *Tune Into Reading* (Electronic Learning Products, 2006), or to a control group (n=32) that were rereading through singing as part of their regular music program. All 64 participants were members of an assigned elective Wheel Music Class classroom during the fourth quarter of the 2006-2007 school year (April 2nd - May 15th, 2007) over a seven week experimental treatment period.

The *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006) was utilized to measure from pretest to posttest the performance in fluency, WPM (measured by words per minute), word recognition, WR (measured by oral reading accuracy), reading comprehension COMP (measured by implicit and explicit questions after the reading), and instructional reading level, RL (measured by combining scores from word recognition and comprehension questions) before implementation. Initially, the students were assessed at posttest with a reading passage on the same instructional levels attained

during the pretest. This was followed by analysis at the students' highest instructional reading level.

Concurrently, this investigation provided a description of the peers' interactions in both groups during the literacy task assigned by the music teacher. The intent of this study was to address the following research questions:

Quantitative Research Questions

1. To what extent is the reading performance of word recognition, fluency, comprehension, and instructional reading level, as measured by the QRI-4, of students using the *Tune Into Reading* program, different from their regular music curriculum counterparts?
2. To what extent does the *Tune Into Reading* program differently impact the reading scores of students who are “below, at, or above” grade level as determined by the Florida Comprehensive Assessment Test (FCAT) reading scores?

Qualitative Reading Question

1. How do middle school readers interact with their peers within the context of their music classroom?

Question one addressed the differences in reading performance for the students using the interactive sing-to-read program, *Tune Into Reading*, compared to the students who were rereading through singing in their regular music class. This comparison measured the students in their fluency, WPM (measured by words per minute), word recognition, WR (measured by oral reading accuracy), reading comprehension COMP (measured by implicit and explicit questions after the reading), and instructional reading

level, RL (measured by combining scores from word recognition and comprehension questions).

Question one findings revealed the treatment students of varying reading abilities that used the interactive singing program, *Tune Into Reading*, illustrated a significant increase in their Fluency (WPM), Reading Comprehension (COMP), and Instructional Reading level (RL) as compared to their counterparts who were singing in the regular music class. In addition, for the treatment students, Word Recognition (WR) indicated a larger effect from pretest to posttest than the control group. Specifically, this suggests that rereading through singing, using the interactive singing program, *Tune Into Reading*, was more effective regardless of the reading levels for treatment students compared to control students. These results can also be interpreted as rereading through singing in the music classroom alone, as was the case for the control students, does not improve WPM, WR, COMP, and RL for the students of varying reading abilities. Therefore, these findings suggest that regardless of their reading levels early adolescents benefited from fluency instruction.

Furthermore, at the highest reading level reported at posttest, although the treatment group had a significant increase in their instructional reading level (RL), it was reported there was no significant difference between the groups in WPM, WR, or COMP. The descriptive findings suggested the treatment students were interactive with their learning. They appeared to assimilate and accommodate the new learning from the text while they were rereading through singing. However, as the new material became harder they shifted to a state of disequilibrium (Piaget, 1964). This might be interpreted as reading fluency is not a static condition; instead, it is fluid and continually developing.

Toppings (2006) suggests an early adolescent can be a fluent reader at one level and yet display dysfluency at a higher level. Thinking about reading fluency using Toppings theory, it then might be inferred for this group of literacy learners that reading fluency should be thought of as a strategic process, rather than a skill acquired through repeated practice alone as reported in the literature (e.g., Samuels, 2006).

Finally, it was reported that within the groups, the treatment group illustrated a significant increase in fluency (WPM), word recognition (WR), reading comprehension (COMP), and instructional reading level (RL); whereas, within the control group, there was no significant increase from pretest to posttest in any of these areas. This suggests that within the groups, during the literacy task of rereading through singing, something happened within the classroom culture of the treatment group that was different from the control group. Further analysis revealed the peers' social interactions within the treatment group's classroom culture might have contributed to the significant increases in all variables. Particularly, the peer interactions appeared to be supportive and collaborative. These results suggested the act of literacy was embedded within this network of social relations. Moje (1996) contends that in the secondary content classroom, it is the social context that shapes the literacy practices for the early adolescent.

Question two used the same scores from the QRI-4; however, the students were grouped by their 2006 FCAT Reading Level achievement scores. The FCAT reading scores (levels 1-5) range from highest level (5) to lowest level (1). The treatment and control groups were stratified according to their FCAT Levels as: (a) Level 4 or 5 *Above* grade level, (b) Level 3 *At* grade level, and (c) Level 1 or 2 *Below* grade level. Once the

students were grouped, a comparison was made between the groups on each dependent variable, looking at how each FCAT Level was differently impacted.

The results reported the intervention was more effective for the treatment students that used the sing-to-read program, *Tune Into Reading* program, compared to the control group. Interestingly, it was also noted that for the treatment students at the various FCAT Levels, the program used afforded them opportunities to improve differently in the reading components each level individually needed. Specifically, for the students grouped as *Below* grade level in their FCAT scores, the intervention was more effective in improving reading rate WPM and word accuracy WR. However, for the students grouped as *Above* grade level, the results reported reading comprehension COMP was more effective. These findings suggest for the treatment students that used the sing-to-read program, *Tune Into Reading*, this interactive sing-to-read program was effective in meeting the differentiated needs for each level.

However, when the FCAT Levels were used as benchmarks for the initial pretest of the groups, there was a discrepancy between the reported FCAT Levels and results of pretest scores from the QRI-4. Specifically, when all the participants were given the QRI-4 pretest, there was no significant difference between the groups. Conversely, for the students stratified by their FCAT levels 3 (*At* level) and 4 and 5 (*Above* level), it was reported the groups had different pretest scores on their instructional reading levels (RL) than the control group. Specifically, showing for the treatment group on Level 3 (*At*), they had significantly higher pretest scores than the control; whereas for control group Level 4 and 5 (*Above*), they had significantly higher pretest scores than the treatment students. This suggests using FCAT reading level scores as benchmarks to determine

instructional reading level does not appear to correlate with the scores from the QRI-4 assessment.

Question three investigated peer interactions that occurred during the study intervention. There were 14 rereading through singing sessions for students in treatment and control groups that occurred for 50 minutes each, twice a week, over a seven-week period. During these sessions, observational notes were taken on the peer interactions that occurred during these sessions. These observations focused on describing the relationship, if any, between the literacy task the music teacher assigned (rereading through singing) and the peers' interactions during the task. There were two cases in this study. The treatment students were singing using the interactive program, *Tune Into Reading*, and the control students were reading through singing in the traditional music class.

As noted previously in the findings, it was suggested that during the literacy task of rereading through singing, the classroom culture and the occurrence of social interactions might have contributed to the significant increases in all variables within the treatment group; however, this appeared not to be the case within the control group. Suggesting that within the treatment group during these sociocultural interactions, the classroom culture supported academic improvement

The findings suggested the treatment groups' classroom culture appeared to be safe, risk-free, motivating, and collaborative; whereas within the control group, the classroom culture was initially motivating, engaging, and competitive. In addition, it was found dominant peers within the treatment group had no direct discussions with the other peers. Instead, they modeled support and collaboration with one another, and other peer

members followed. The dominant peer interactions with other peers within the control group however were through direct discussion as they told the peers to conform to the literacy task assigned. This suggested the data revealed learning for the students in the treatment group progressed from engagement to assimilation, followed by self-regulation and interaction with text. The control group findings revealed their learning shifted from initial engagement to fake reading to disengagement. These findings suggest these sociocultural interactions played an important role in improving fluent reading performance as noted in the treatment group scores.

Discussion: Conclusions and Implications

Addressing Early Adolescents Differing Fluency Development

Biancarosa and Snow (2006) reported to the Carnegie Corporation, over 70% of adolescents struggle with their reading in some manner, and therefore, require differentiated and strategic instruction. Furthermore, they contend that when thinking about reading fluency for the early adolescents, there are a range of literacy needs to be met for this population. Some students may still need support with reading the words; whereas other students can read the words accurately but need support with comprehension. Still, other adolescents may know the strategies but not have had sufficient practice within the classroom. What they need is instruction and support that addresses the differing literacy needs for all students.

As previously noted in the findings, for the treatment students that used the sing-to-read program, *Tune Into Reading*, the program was effective in meeting differentiated needs for each level. For students grouped as *Below* grade level in their reading, the intervention was more effective in the reading areas of Fluency (WPM) and Word

Recognition (WR); whereas, for students grouped as *Above* grade level, the intervention was more effective on reading comprehension (COMP). This suggested the sing-to read program was effective for the reading areas each group of students needed and therefore addressed the range of differing needs.

However, even when given the unique individual differences among early adolescent literacy learners, curriculum delivery is often a one-size-fits-all practice (Alvermann, 2001; Ivey, 1999; Moore, 2000). Therefore, the integration of content literacy to meet the diverse needs for this population is challenged through the contextual structure and curriculum delivery. This was evident in the classroom structure and curriculum delivery of the treatment and control groups.

Small groups provided classroom structure to the treatment group. The students worked in the computer lab in small group communities, and curriculum delivery was accomplished through individual computer usage. However, the classroom structure for the control group was through a whole group format, and curriculum delivery was provided to the entire body of students present.

The findings suggested the treatment group using the interactive sing-to-read program, *Tune Into Reading*, individually had higher reading outcomes compared to the control group who were singing within a whole group setting. The inference is that in order to meet and address the reading fluency needs for this population of literacy learners, instruction and delivery of curriculum needs to meet the individual needs of the students. This suggests that for fluency instruction to be successful, the curriculum delivery should provide opportunities for individual work.

Assessing Fluent Readers in the Middle School

Fluency is a necessary aspect of successful reading, as it allows the readers to read with speed, accuracy, and proper expression (National Reading Panel, 2000; Rasinski, 2004). The National Reading Panel (2000) reported they found sufficient evidence that guided oral reading through repeated reading will have a positive impact on fluency and comprehension. However, the literature on reading fluency often focuses on the beginning reader's initial stage of literacy acquisition or on the older adolescent reader who has difficulty learning to read. This focus has placed reading fluency in a deficit view, which focuses on remediation at the decoding level, rather than creating a direct link to comprehension (Clay, 1985). Stayter and Allington (1991) suggest that "we have failed to consider some of the broader ramifications of an emphasis on fluency, especially with older and more developed readers" (pp.143-144). This appears to be true when fluency instruction could support both the struggling and more developed reader's, as was found in this study with the increased reading outcomes for all of the students of varying reading levels, using the interactive sing-to-read program, *Tune Into Reading*.

Assessing proficient fluent reading for this group of literacy learners proves to be a difficult task, as little has been addressed about needs of early adolescent middle school reading fluency of varying reading abilities. Therefore, assessments that are used with beginning readers (e.g., Dynamic Indicators of Basic Early Literacy Skills DIBELS; Good & Kaminski, 2002) or high-stake test scores (FCAT) are utilized to determine proficient fluent adolescent readers.

The ORF (oral reading fluency) and RF (retell fluency) are assessments currently used in this middle school. These are subtests and a part of the DIBLES assessments used with older students. The concern with the ORF and RF tests according to Allington, (2006) is:

During the ORF test the student is given one minute to orally read a passage, while the examiner counts the number of words correctly read within the minute. During the RF test the students reads orally for one minute and then the student is asked to retell what he or she can recall from the passage. While the student is retelling the story, the teacher counts the number of words uttered by the student (p.40).

This might explain why the students in this study equated fluent reading with speed and not to comprehending text while decoding. As noted during the pretest assessment, 75% of the students asked prior to reading, "*How fast do you want me to read?*" or "*Do I need to read this fast?*" (Assessment Notes April 2, 2007). Although they were told to read at a pace that they would be able to understand and answer the comprehension questions, after reading the passage, the students still read quickly with no expression and no pauses or stopping at punctuation during their reading at the pretest. Suggesting, for these students, their understanding of what it meant to be a fluent reader was equated to reading the words with speed, instead of reading for meaning. A reasonable conclusion reached is assessment that does not take into consideration deep comprehension (internalizing material), but only surface comprehension (word level speed and accuracy), can prove to be problematic for determining proficient reading of early adolescents of varying reading abilities.

High-stakes testing scores are also used to determine a fluent and proficient reader. The FCAT levels 1- 5 in reading are used as indicators of reading ability and performance, according to the State of Florida. Levels 1 or 2 are considered below grade level; whereas, Levels 3 through 5 are considered at or above grade level in reading. The results of this yearly assessment can have a dramatic impact on the early adolescent literacy learner with the possibility of retention, class placement, and specifically instructional practices provided to the students. The score obtained from this high-stakes test place early adolescents below, at, or above their classmates in reading, and it is assumed early adolescent students who may or may not have passed the test will receive instructional strategies needed to prepare them to be fluent readers and comprehenders.

The FCAT reading level scores were used in this study for two reasons: (a) to address the second research question of this study concerning the comparison of the relationship with reading performance and FCAT levels, and (b) to approximate the appropriate beginning reading levels prior to the QRI-4 pretest assessment.

The primary purpose of the FCAT in reading is to assess student achievement of higher-order thinking skills (Florida Department of Education, 2005); therefore, it was assumed a student who attained higher FCAT level scores in reading (levels 3-5) would be at or above grade-level in reading. However, when the FCAT level reading scores were used to determine the benchmarks for administering the QRI-4 at pretest and posttest, there appeared to be a much lower than anticipated relationship between the FCAT level reading scores and scores obtained during the QRI-4 assessments.

In particular, when the students were stratified by FCAT reading levels 3-5 (as *At* or *Above* grade level in reading), it was found that FCAT reading level scores reported

71% (40 out of 56) of the students were meeting grade level or above in their reading. However, when these same students were given the QRI-4 reading assessments, the results demonstrated only 19% of the students (11 out of 56) at pretest were on grade level or above in reading, and at posttest only 27% of the students (15 out of 56) were on grade level or above in reading. This suggested only 15 students out of the 40 students determined by their FCAT level reading scores were in fact meeting grade level proficiency, according to their QRI-4 scores in reading at posttest.

These findings suggest the correlation assumed was not found between scores on the FCAT and scores from the QRI-4 used to determine proficient fluent readers. Therefore, it might be inferred that the use of a high-stake test scores can not account for the many variables associated with understanding the reading process when relating that to the characteristics of this group of early adolescent literacy learners and their fluent reading behavior (McCombs, Kirby, Barney, Darilek, & Magee, 2005; Rothstein, 2000). Amrein and Berliner (2002), overall, contend that “there is no compelling evidence from a set of states with high-stakes testing policies that those policies result in transfer to the broader domains of knowledge and skill for which high-stakes test scores must be indicators” (p.54).

Furthermore, the findings revealed that for the 56 students in this middle school, 27% were proficient readers, and 73% are reading below grade level, according to their QRI-4 scores. These results align with the report from The National Assessment of Educational Progress (NAEP) (2005) that reports 73 % of eighth grade students perform below or at a basic level in their reading achievement.

In addition, Rothstein (2000) questions whether an annual test of a students' knowledge, at just one point in time, could provide an accurate assessment of fluent reading for this population of literacy learners. This was particularly true in this study, as the treatment students shifted from a fluent reader on one level to a surface fluent reader on a higher level. The findings reported suggest treatment students were in a state of disequilibrium that mirrored Piaget's Theory of Cognitive Development (1964).

Based on Piaget's theory (1964), as the students assimilated the higher level reading material and were building the schema for this new information, so they could accommodate it, they were in a state of disequilibrium. Suggesting their fluency growth in reading was fluid and changing as each new cognitive task presented itself and required students to build the cognitive stamina for the new more difficult reading tasks. This is interpreted as the use of an annual assessment to determine a fluent proficient reading for the early adolescent is problematic because it does not allow for the ever changing state of fluent reading as found in this study.

Toppings (2006) contends reading fluency is not "an entity, or a benchmarkable competence, or a static condition" (p.106). In addition he adds, "Even expert readers will show dysfluency when confronted with an unfamiliar topic that provides challenge greatly beyond the students' independent reading level" (Toppings, 2006, p.106). This appears to contradict some of the literature on reading fluency assessment, particularly when fluency is measured as a discrete skill (reading rate and word accuracy). However, it appeared to be a strategic process for the students in this study.

The Role of Prosody in Reading Fluency

Reading prosody is the music and rhythm of oral language. Specifically, when a student demonstrates expressive oral reading by using pace, volume, pitch, and rhythm, this is indicating behaviors of prosodic reading. However, there is not a consensus in the field concerning the role prosody plays in reading fluency. The reading literature suggests fluent readers exhibit behaviors that blend reading accuracy, automaticity, and prosody (Samuels, 1979). Some scholars contend it is the prosodic elements in reading that has a direct connection to reading comprehension (e.g., Allington, 2006; Rasinski, 2004); whereas, other scholars (e.g., Torgesen & Hudson, 2006) view reading prosody as not having any direct relationship to comprehending text. Instead, they suggest decoding (word accuracy) with automaticity (reading rate) are the direct connection to comprehension. While there is no debate amongst the reading community as to the need for fluent readers to be efficient decoders, in order to comprehend text, the stance taken that word level reading with speed alone improves comprehension can be problematic for the fluent middle school decoders. Biancarosa and Snow (2006) contend most early adolescents do not have difficulty reading fluently at the word level; instead, the difficulty arises with their reading comprehension.

The *Qualitative Reading Inventory-4* (QRI-4) (Leslie & Caldwell, 2006) was used in this study as a pretest and posttest measure to determine the students' fluency, word recognition, comprehension, and instructional reading levels for the groups. An instructional reading level is calculated by using the combination of a score in word accuracy and reading comprehension. Therefore, to determine the instructional reading level for the students in this study, it combined their word accuracy scores with their

comprehension scores. As previously noted by Biancarosa and Snow, the students in this study appeared not to have difficulty with reading the words; the difficulty arose with comprehending what they read.

The word accuracy (WR) scores reported showed no statistically significant difference between the groups from pretest to posttest. In addition, the WR scores indicated both groups were at independent level in how accurately they could read the words in the text. In fact, when the students were group by FCAT Levels for the students *At* and *Above* grade level in reading, they had reached a “ceiling effect” (Stevens, 2002). This suggested that for these students WR had gone as high as it could go at this level. However, because their comprehension scores were not as high, the students could not be moved to a higher instructional reading level. In addition, the reading rate WPM for both groups met an acceptable criterion (140 WPM) for the students of this age group (Rasinski, 2004). This might suggest the variables of word accuracy in reading (WR) and the reading fluency rate of speed (WPM) may not be contributing factors for early adolescents when thinking about important components for fluent reading leading to comprehension.

Based on The Automaticity Theory, LaBerge and Samuels (1974) define fluent reading as the ability to decode and comprehend text at the same time. Their theory suggests cognition has only a limited capacity to process information. Therefore, decoding (at the word level) can become automatic, and the focus cognitively can be on the complex process of comprehending text. Through guided and repeated reading, both decoding (automaticity and accuracy) in word recognition and comprehension are developed.

Samuels (1979) further defines repeated reading as a fluency-building strategy that consists of timed rereading of a short passage several times (at least 3 times), checking for accuracy (word recognition), automaticity (words per minute) and with prosody (expression). Furthermore, the steps recommended for an effective fluency instructional model are: (a) to provide a model for student's expressive fluent reading, (b) to give the students a passage to read (approximately 150 words) 3 times, and (c) to have the students orally read the passage assessing for accuracy, automaticity, and expression (Rasinski, 2004).

Repeated reading is most authentic when the practiced material is eventually performed orally, such as plays, poetry recitation, or in this study singing lyrics to songs (Rasinski, 2004; Stayter & Allington, 1991). This form of repeated exposure through singing assists the reader with fluency through prosodic reading. The singing performed by the students appears to exaggerate the language of reading, as the students find their voice in the rhythm and the bounce of the music. The reader uses appropriate volume, rhythm, pitch, tone, and phrasing (prosody), while singing the song lyrics; therefore, they give evidence of actively constructing meaning from the passage (Rasinski, 2004).

The findings of this study concur with Rasinski (2004), in part. Distinctively, prosody when rereading through singing appeared to have a direct connection to reading comprehension (COMP) and increasing the instructional reading level (RL). However, the practice of rereading, through singing by following the protocol recommended in the literature alone, did not produce the same findings as what has been previously reported (Samuels, 1979). If that were indeed the case, then both groups should have increased in their reading comprehension and instructional reading level because both groups

followed the recommended procedures for fluency instruction. Nevertheless, the treatment students significantly outperformed the control group in reading comprehension (COMP) and instructional reading level (RL). This suggests repeated practice of rereading through singing by the control group of varying reading levels did not improve their reading comprehension or instructional reading levels.

In addition, the treatment group appeared to interact with the prosodic elements of text, rather than just being passively immersed in the prosodic elements through repeated practice as noted in the control group. Specifically, the treatment group applied reading strategies to comprehend the prosody of the songs which resulted in an increase in their reading comprehension over the control group. As noted in the excerpt from the data:

He was focused on rereading the song.

This was the fourth time he recorded himself singing.

He had the screen up that displayed his vocal tract.

He used his finger to align where he was off pitch.

He went back again and reread the song again.

[He moved back and forth in his seat, nodding his head to the beat, and tapping his foot to the music]

Finally, [he sub vocalized] I got a 90!

(Week 4, April 25, 2007).

The student (representative of his peers) was being strategic and metacognitive as he interacted with the text. He was interactive with the text as he applied strategies for effective comprehension of text. However, he appeared to be interacting and applying strategic processes to the prosodic elements of the text. As noted, he would trace his

finger on the vocal tract line and then reread and re-sing the song as his body moved to the rhythm and beat of the music. The above example exemplifies an interaction with the prosodic elements of the text. Therefore, based on The Automaticity Theory (LaBerge & Samuels, 1974), it appeared that not only was the student decoding automatically, the strategic processes appeared to focus on comprehending the prosody of the text. This suggested that for reading comprehension to increase and see transfer effects for reading comprehension to other reading material (e.g., QRI-4 assessments), these middle school students needed to be interactive and comprehend the song lyrics and the prosodic elements of the text they were reading through singing.

The individual interactive sing-to-read program, *Tune Into Reading*, was used individually, and as noted in the above excerpt, the student manipulated the text to understand the song and its prosodic elements. He used various reading strategies to see where he could improve, as he traced the voice frequency lines, reread the song and then recorded himself again until he reached his goal. The practicality of this alternative format assisted him in comprehending text, unlike a linear text that can not be stopped, started, or slowed down. In addition, the continuous background music assisted him because he did not have to use his cognitive capacity to remember the rhythm or beat of the song that was being automatically supplied. Therefore, he could focus on comprehending the prosodic elements while being guided automatically by the background music.

In addition to the significant increase in reading comprehension (COMP) scores, the treatment group significantly outperformed the control group in their instructional reading level (RL). As previously noted, instructional reading level is calculated by using

the combination of a score in word accuracy and a score in reading comprehension.

Mariotti and Homan (2005) suggest that to determine the percent correct for word recognition, the teacher counts the errors and subtracts it from the total number of words in the passage, then divides by the total number of words contained in the passage (p.76).

The formula is noted as:

$$\frac{\text{total number of words in the passage} - \text{errors}}{\text{total number of words in the passage}} = \text{word recognition percent correct}$$

To determine the comprehension percent correct, the teacher subtracts the errors from the total number of questions, and then divides that number by the total number of questions.

The formula is noted as:

$$\frac{\text{total number of questions} - \text{errors}}{\text{total number of questions}} = \text{comprehension percent correct.}$$

Once this is accomplished, Mariotti and Homan (2005) suggest a criterion is used to indicate instructional reading levels of the students. Two well known scholars developed criteria for determining instructional reading levels, Betts (1946) and Powell (1971). Betts criteria suggest that there is a standard baseline of scores across grades that can be interpreted descriptively incorporating the prosodic elements of oral text reading. Powell criteria adjust the baseline in word recognition and comprehension for passage difficulty by passage reading levels (Mariotti & Homan, 2005).

When looking at the criteria separately, it does not appear to totally address the needs of interpreting instructional reading level for these early adolescents of varying reading ability. However, possibly combining the criteria might address the elements

necessary to capture the behaviors for instructional reading levels for early adolescent readers. Specifically, this could be accomplished by using Betts criteria that descriptively captures the prosodic elements and Powell's criteria that adjusts the baseline for word recognition. Therefore, it would address what was found in this study; the prosodic elements in reading played a significant role in increasing instructional reading level and comprehension for the treatment group.

In addition, I concur with Mariotti and Homan when they state the most important function of and Informal Reading Inventory (IRI) is qualitative descriptive interpretations of behaviors in reading, along with the quantitative criteria that need to be taken into consideration when determining instructional reading levels. These behaviors such as pausing at sentence, self-correcting, using tone, and other prosodic elements were found as indicators of comprehending text. In particular, as found in this study at pretest during the reading assessments, both groups of readers read their assessment passages orally with speed and a high level of accuracy in word recognition, yet they struggled with comprehension during the pretest assessment. Their oral reading was absent of volume, tone, pitch or any expression. There was no pausing at punctuation, rereading for clarification, or self-corrections made in 53 out of 64 students or 83% of the groups.

However, at posttest, the treatment group of students outperformed their counterparts significantly from pretest to posttest in reading comprehension COMP and instructional reading level RL. The oral reading of the students in the treatment group, although fast (180 wpm), had expression, pitch, and volume, unlike their counterparts. Specifically, 81% of the treatment students or 26 out of 32 read their passage making self-correction, pausing at punctuation, and rereading phrases or sentences. Whereas, in

the control group of students, only 28% or 9 out of 32 of these students incorporated these prosodic elements in their reading. This suggests prosodic elements of reading appear to have a direct connection to reading comprehension.

Sociocultural Interactions

Vygotsky (1978) contributed to the conception that literacy is a social construction, specifically, viewing cognition as a profound social phenomenon. Initially, learning is socially constructed, and then as the higher mental processes take shape, learning becomes internalized. If this perspective is embraced, it could be interpreted as social experiences through sociocultural interactions shape thinking and interpretations of the world.

The treatment and control groups had a social system in place that positioned some of its peer members in the role of dominance over other peers. These dominant peers took this position, and the other peer members allowed them to assume it. Both groups appeared to have this two class system, where a small group or a few individuals lead the rest of the group, determining what was considered acceptable social behavior. Interestingly, however, within the groups the interactions from the dominant peers with the other peer members was accomplished very differently.

A small group of eighth grade males were the dominant peers within the treatment group. They modeled behaviors or talk that resulted in the rest of the peers imitating their behaviors. Their talk was supportive and collaborative with one another; however, there were no discussions with other peer members only amongst this small group. This is consistent with Ryan's (2000) definition; modeling is a form of adolescent peer interaction. This interaction refers to individual changes in cognition, beliefs, or affect,

which are a result of adolescents observing their peers. Observing a specific behavior a peer performs or listening to a peer voice, a certain belief can induce an adolescent to change their stance or adopt their peers' behaviors or beliefs. Schunk and Zimmerman (1996) reported peer modeling influenced self-efficacy beliefs, as was found in this study. The students, after observing the dominant peers in the treatment group, initially used the sing-to-read program, *Tune Into Reading*, and continued its use while showing support and cooperation with one another as modeled by these dominant peers.

Whereas within the control group, the dominant peers were male or female individuals and their interactions were direct discussion with the other peer members. The dominant peers directed the other members to conform to certain behaviors they deemed socially acceptable. This might have been a result of how these dominant peers interpreted what was expected of them through the instructional expectations provided by the music teacher's modeling. Information exchange occurs when adolescents have a discussion with their peers (Berndt, 1999). This form of interaction could influence the early adolescent's choice to partake in the literacy task presented by the teacher if it was used effectively.

However, Ryan (2000) contends it also has an adverse influence if the peers use this form of interactions to control other peers to conform to socially acceptable behavior. As noted in the findings, the dominant peers within the control group directed and tried to intimidate and control the more vulnerable peers into conforming to the instructional expectations. This appeared to have an adverse effect on the other peer members. As noted in the findings, the other vulnerable peers did not come to aid of the peer that was being admonished; instead, they would look at one another or look away.

Peer pressure can also take on the role of social reinforcement (Ryan, 2000). Brown, Lohr, and Eicher (1986) found beliefs and behaviors that are discouraged by the groups are not likely to be displayed; whereas, beliefs and behaviors positively received by the group are more likely to surface. Participation in the literacy tasks involving the treatment peer group positively modeled through the dominant peer interactions had a positive effect on the group's beliefs and decisions to participate by all of the group members. Whereas, within the control group setting, what appeared to happen was that peer pressure was applied by the dominant peers, and it was not positively received. Therefore, they disengaged from the task while trying to escape the pressure. These findings suggest the role of the dominant peers and sociocultural interactions have a significant influence in the reading performances of the group. Specifically, it was found the treatment group showed a significant increase in all areas of reading fluency; however, the control group did not. This might be interpreted as the sociocultural interactions modeled through the dominant peers in the treatment group of support and collaboration was positively interpreted by their peers, and the results were higher in performance of reading within the group.

Contributions of the Study

Although previous research has identified characteristics of effective reading fluency instruction, the focus has been on beginning readers or older struggling readers. This focus has involved interpreting fluent reading as having a connection to reading comprehension at the word level (LaBerge & Samuels, 1974). As found in this study, that was not the case for middle school readers in the treatment group. Instead, it appeared to be the prosodic elements of text that had the direct connect to reading comprehension. As

future fluency studies at the middle school level are completed, the results from this study will provide additional information to the role prosody plays for fluent readers of varying reading ability in the middle school.

In addition, there have been very few studies conducted on middle school readers of varying reading ability, following the protocol for effective fluency instruction. Furthermore, few studies have been conducted that embed literacy strategies naturally in a content class, while looking at transfer effects of comprehension to other reading material. In this study, the literacy task of rereading through singing maintained the integrity of the subject matter of the music class, while embedding the literacy elements for fluency instruction. The results revealed rereading through singing for treatment students transferred to a reading assessment, showing a significant increase in reading comprehension. These results contribute to the concept of embedded literacy instruction and transferability of reading performance for other reading tasks.

Finally, very few studies have integrated the role of peer interactions during a specific literacy task, while measuring their reading performance in reading. The findings revealed for both treatment and control group a two class system within the classroom settings. These social systems were led by the dominant peers, and the rest of the members allowed this and followed their lead. However, how the dominant peers interacted influenced the other peer members. The results suggested these interactions determined how all the peers responded to the literacy task. These findings help to clarify the role that peer interactions might have in the middle school.

Recommendation for Practice

Fluency Instruction

As noted in the findings, the role of prosody appears to have a direct connection to reading comprehension for treatment students. However, the instructional emphasis of expressive reading tends to decrease for students once they leave primary grade (Allington, 2006). Therefore, there is a need to incorporate models and practice of prosodic elements of reading text for students of all levels, especially within the middle school environment. This can be accomplished by having teachers have more read alouds and provide students with more opportunities to orally read so they are able to practice the prosodic elements of text.

In addition, fluency instruction needs to be differentiated to meet the developing needs of these students. As noted in the findings, the treatment group using the interactive singing software, *Tune Into Reading*, when grouped by FCAT Levels was effective in meeting the differentiated needs for each level. Additionally, when the treatment students reached a higher level in their reading, their fluency decreased. This suggests fluency in reading is not stagnant; it is instead fluid and ever changing with the different tasks middle school students face (Topping, 2006). Suggesting, fluency is a strategic process rather than a skill. As well, the expectations that students in the middle school enter the context of the school environment as fluent readers should be reevaluated, as this was not the case with this group of early adolescents

Furthermore, opportunities for individual practice, rather than a whole group one-size-fits-all model, should be considered. It was found that students in the treatment group made a cognitive shift from assimilating the reading information to interacting and

internalizing their learning. This, in part, appeared to be because they had opportunities for individual practice. In addition, the alternative text students used in the treatment added to their comprehension improvement.

Both groups found the alternative approaches to learning motivating and engaging. This was apparent during the interactions within the groups. The light-hearted competitions through the peer interactions were documented in the data. The groups led by the dominant peers would, either through discussion or modeling, set the climate of motivation for rest of the group. However, as the intervention continued, a shift occurred within both of the groups as to their motivation for these alternative approaches to the task of rereading through singing, as did the role of the dominant peers.

For the alternative textual format, the *Tune Into Reading* program that the treatment group used was not only motivating and engaging but easily manipulated. The students could adjust the program, and this appeared to assist them in comprehending the prosodic elements of the text. In addition, the perception of what alternative was changed during the course of the interaction. The treatment students appeared to perceive this musical textual format as one that assisted them in their learning. This suggests defining and using alternative textual formats should include, not only the delivery of the text and the genre it provides, but the perceptions it develops. This perception changed from a fun game-like alternative text to a text the student could use to comprehend the reading material.

Assessing Reading Fluency

Since it appeared prosody had a direct connection to reading comprehension, assessments should assess the students with this element in mind. I concur totally with Mariotti and Homan (2005) when they suggest qualitative descriptive interpretations of behaviors in reading, along with the quantitative criteria, need to be taken into consideration when determining instructional reading levels and fluent reading behaviors. These behaviors such as pausing at sentence, self-correcting, using tone, and other prosodic elements were found as indicators of comprehending text in this study.

As well, the measurement tools currently being used (e.g., ORF and RT) appear detrimental to interpreting reading fluency for these students, as the students and their teachers are interpreting fluency reading at a surface word level. Furthermore, using the FCAT reading levels scores appears problematic, as the use of these scores as benchmarks did not correlate to the QRI-4 reading levels. Therefore, not all the students that might benefit from further instruction in their reading fluency are actually getting instruction they need. In addition, the findings revealed all the students in the treatment group benefited from fluency instruction. However, when using these scores within the school setting, only those students suggested by these scores are receiving fluency instruction, when all could benefit.

Sociocultural Interactions and Influences on Instruction

The role of dominant peers remained constant within the treatment and control groups, during the intervention. Nonetheless, instructional expectations afforded to the groups by the music teacher were very different within each setting. These expectations

appeared to be interpreted by the dominate peers and then reinforced through their interactions with the rest of the peer group.

What appeared to be expected of the peers in the control group was only one right way to perform the singing and drumming. These expectations were modeled to the peers as the music teacher would stop the singing and drumming several times daily during each of the sessions and tell the students some people were off key, and to try it again. In turn, the findings revealed the dominant peer would admonish the vulnerable peers for making a mistake. Nevertheless, the treatment group was expected to try their best. The music teacher would often remind the students to try their best and not to worry about their pitch scores. The dominant males in the treatment group would encourage and support one another, modeling collaboration to the other peers. Therefore, it appeared that within both groups, the dominant peers interpreted what was expected of them, and then in turn, reacted to these expectations through their interactions.

The control group peers were expected to perform correctly and accurately as modeled by the music teacher. Since there was only the right or wrong way to perform the song and drum sequence, the interactions might have been interpreted by the dominant peers and conveyed to the rest of the group in this manner through these interactions. However, the treatment group was allowed differentiation through the instructional delivery. Therefore, the dominant peers appeared not to be compelled to take on the task of reinforcing group accuracy; instead, they became a group member while still maintaining their position among the group. This suggests instruction should meet the needs of the individual students, and individual accomplishments should be rewarded.

Recommendations for Future Research

The findings of this study reported the role of prosody appears to have a direct connection to reading comprehension. Conversely, the fact that there were only 64 students and the individual students of this study were not each randomly assigned to a treatment or control condition limits the generalizability of the findings to this group of students. In addition, the duration of the study was seven weeks in length. Future researchers may consider increasing the sample size and lengthening the study period to obtain additional data for reading fluency.

Additionally, as to random selection for this population, the sample characteristics were predominantly White eighth grade low SES males. There were no sixth grade students, a limited number of seventh graders, or students that required additional support in their learning. Future research should investigate a greater diversity in the classification characteristics of the students of this study.

The content class was a music class, and the strategies taught were appropriate for this content area. The findings suggested there was a transfer effect from the embedded literacy taught to another literacy task. Future research might investigate embedded literacy to see if this transfer effect holds between other content classes.

Finally, the mixed method design of this study was effective in capturing the reading performances and the descriptive findings. However, the case study used observational field notes only to capture the peer interactions but did not include videotapes or tape recordings of these interactions. Future researchers might want to utilize these in their research designs for the purposes of capturing more in-depth

understanding of peer interactions and how this relates to the literacy task that the students are involved

References

- Alfassi, M. (2004). Reading to learn: Effects of combined strategy instruction on high school students. *Journal of Educational Research, 97*, 171–84.
- Alexander, P. A. (1998). The nature of disciplinary and domain learning: The knowledge, interest, and strategic dimensions of learning from subject-matter text. In C. Hynd (Ed.), *Learning from text across conceptual domains* (pp. 263-287). Mahwah: Erlbaum.
- Alexander, W.M. (1968). *A survey of organizational patterns of reorganized middle schools*. (final report number 7-0-026). Gainesville, FL: University of Florida.
- Alexander, W.M., & George, P.S. (1981). *The exemplary middle school*. New York: Holt, Rinehart, and Winston.
- Alexander, P. A., & Jetton, T. L. (2000). Learning from traditional and alternative texts: New conceptualization for an information age. In A. Graesser, M. Gernsbacher, & S. Goldman (Eds.), *Handbook of discourse processes*. Mahwah: Erlbaum.
- Allington, R. L. (2006). Fluency: Still waiting after all these years. In S.J. Samuels, & A. E. Farstrup (Eds.), *What Research Has to Say About Fluency Instruction*. Delaware: International Reading Association.
- Allington, R. L. (2001). *What really matters for struggling readers: Designing research-based programs*. New York: Addison-Wesley.

- Alvermann, D. E. (2002). Effective literacy instruction for adolescents. *Journal Literacy Research*, 34, 189-208.
- Alvermann, D.E. (2003). *Seeing themselves as capable and engaged readers: Adolescents and re/mediated instruction*. Learning Point Associates. North Central Regional Laboratory, Institute of Education Sciences, U.S. Department of Education, contract number ED-01-CO-0011.
- Alvermann, D.E., & Moore, D. W. (1991). Secondary school reading. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of Reading Research*, 2, 951-983. New York: Longman.
- Alvermann, D.E., & Phelps, S. P. (2005). *Content reading and literacy: Succeeding in today's diverse classroom* (4th Ed.). Boston: Pearson Education.
- Amrein, A.L., & Berliner, D.C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved (11/11/2006) from <http://epaa.asu.edu/epaa/v10n18/>
- Ary, D., Jacobs. L.C., & Razavieh, A. (1996). *Introduction to research in education* (5th ed.). Fort Worth: Harcourt Brace College Publishers.
- Bandura, A. (1997). *Self-Efficacy: The exercise of control*. New York: Freeman.
- Barr, A.S. & Glifford, C. W. (1929). The vocabulary of American history. *Journal of Educational Research*, 20, 103-121.
- Beane, J., & Brodhagen, B. (1996). *Doing Curriculum Integration*. Madison: University of Wisconsin.

- Bean, T.W. (2000). Reading in the content areas: Social constructivist dimensions. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 629-644). Mahwah: Erlbaum.
- Becker, H. J. (1987). *Addressing the needs of different groups of early adolescents: Effects of varying school and classroom organizational practices on different social background and abilities*. (Report No. 16). Maryland: VSP Industries.
- Berndt, T. J. (1999). Friends' influence on students' adjustment to school. *Educational Psychologist*, 34, 15-29.
- Berndt, T. J., Laychak, A. E., & Park, K. (1990). Friends' influence on adolescents' academic achievement motivation: An experimental study. *Journal of Educational Psychology*, 82, 664-670.
- Biancarosa, C., & Snow, C. E. (2006). Reading next—A vision for action and research in middle and high school literacy: A Report to Carnegie Corporation of New York (2nd ed.). Washington, DC: Alliance for Excellent Education.
- Biggs, M.C., Homan, S.P., Dedrick, R., Minick, V., & Rasinski, T. A. (in press). Using and interactive singing program: A comparative study of middle school struggling readers. *Reading Psychology An International Quarterly*.
- Binet, A. (1904). New investigations upon the measure of the intellectual level among school children. In *The Development of Intelligence in Children*. Baltimore: Williams and Wilkins.
- Bogdan, R. & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theories and methods* (4th Ed.). Boston: Pearson Education Group.

- Bond, G. L. & Bond, E. (1941). *Developmental reading in high school*. New York: MacMillian.
- Bough, M. (1983). Middle school curriculum- does it meet the need of the transescent child? *Contemporary Education*, 54, 272-274.
- Brimm, R.P. (1963). *The junior high school*. Washington, D.C.: The Center for Applied Research in Education, Inc.
- Broudy, E. (1975). The trouble with textbooks. *Teachers College Record*, 77, 13-34.
- Brooks, K., & Edwards, F. (1978). The middle school in transition: A research report on the status of the middle school movement. Lexington: College of Education , University of Kentucky.
- Brown, B. B. (1990). Peer groups and peer culture. In S.S. Feldman & G.R. Elliot (Eds.). *At the threshold: The developing adolescent* (pp.171-196). Cambridge: Harvard University Press.
- Brown, B. B., Lohr, M.J., & McClenahan, E.L. (1986). Early adolescents, perceptions of peer pressure. *Journal of Early Adolescence*, 6, 139-154.
- Bulgren, J. A., Schumaker, J. B., Deshler, D. D., Lenz, B.K., & Marquis, J. (2002). The use and effectiveness of a comparison routine in diverse secondary content classrooms. *Journal of Educational Psychology*, 94, 356–71.
- Bushwell, G. T., & John, L. (1931). *The vocabulary of arithmetic*. (Supplemental Educational Monographs No. 38). Chicago: University of Chicago Press.
- Butzlaff, R. (2000). Can music be used to teach reading. *Journal of Aesthetic Education*, 34(3), 167-178.

- Callahan, R. (1962). *Education and the cult of efficiency*. Chicago: University of Chicago Press.
- Calkins, L. M. (2001). *The art of teaching reading*. New York: Addison-Wesley.
- Carnegie Council on Adolescent Development (1989). *Turning Points: Preparing American Youth for the 21st Century*, New York: Carnegie Corporation of New York.
- Chall, J.S. (1996). *Stages of reading development*. (2nd Ed.) Fort Worth: Harcourt Brace.
- Clay, M. (1985). *The early detections of reading difficulties*. Portsmouth: Heinemann.
- Cody, R. P. & Smith, J. K. (2006). *Applied statistics and the SAS programming language*. (5th Ed.). Upper Saddle River, New Jersey: Pearson.
- Conley, M. W. (1988). *Knowledge about pedagogical content: How practicing teachers come to understand content reading instruction*. Paper presented at the annual meeting of the National Reading Conference, Tucson, AZ.
- Cook-Gumperez, J. (1986). Introduction: The social construction of literacy. In J. Cook-Gumperez (Ed.), *The social construction of literacy*, (pp. 1-15). Cambridge, England: Cambridge University Press.
- Coser, L.A., Kadushin, C., & Powell, W.W. (1982). *Books: The culture and commerce of publishing*. New York: Basic Books.
- Cottle, T. J. (2001). *Mind fields-Adolescents consciousness in a culture of distraction*. New York: Peter Lang.
- Creswell, J. W. (2003). *Research design qualitative, quantitative, and mixed methods approaches*. (2nd Ed.) Thousand Oaks: Sage Publications.

- Creswell, J. W., Plano Clark, V. L., Guttman, M., Hanson, W. (2003). Advanced mixed methods research design. In A. Tashahhori & C. Teddlie (Eds.), *Handbook on mixed methods in the behavioral and social sciences*, (pp. 209-240). Thousand Oaks: Sage Publications.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Cross, D., & Paris, S. (1987). Assessment of reading comprehension: Matching test purposes and test properties. *Educational Psychologist*, 22, 313-332.
- Cuban, L. (1992). What happens to reforms that last? The case of the junior high school. *American Educational Research Journal*, 29, 227-251.
- Curtis, F. D. (1938). *Investigation of vocabulary in textbooks of science for secondary schools*. Boston: Ginn.
- Deci, E.L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- deLeon, A. G. (2002). Moving Beyond Storybooks: Teaching Our Children to Read to Learn. *Carnegie Reporter*. Fall, Vol. 2, No. 1.
- Dewey, J. (1910). *How we think*. Boston: D.C. Heath.
- Dillon, J. T. (1983). Cognitive complexity and duration of classroom speech. *Instructional Science*, 12, 59-66.
- Dolch, E. W. (1928). Combined word studies. *Journal of Educational Research*, 17, 11-19.

- Dolan, T., Harrison, C., Gardner, K. (1979). The incidence and context of reading in the classroom. In E. Lunzer & K. Gardner (Eds.). *The effective use of reading*, (pp. 108-138). London : Heinemann.
- Dole, J.A., Duffy, G.G., Roehler, L.R., & Pearson, P.D. (1991). Moving from old to new: Research on reading comprehension instruction. *Review of Educational Research*, 61, 239-264.
- Donahue, P.L., Voelkl, J.R., Campbell, J.R., & Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation and states*. Washington, DC: Department of Education
- Dowhower, D. L. (1987). Effects of repeated reading on second grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, 22, 389-406.
- Dowhower, D. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice*, 30, 158-164.
- Draper, R., Smith, L. K. Smith, Hall, K. M., & Siebert, D. (2005). What's more important-literacy or content? Confronting the literacy-content dualism. *Action in Teacher Education*, 27(2), 12-21.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170-177.
- Durkin, D. (1978-1979). What classroom instruction reveal about reading comprehension instruction. *Reading Research Quarterly*, 15, 481-533.
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to adolescents? The impact of stage–environment fit on adolescents' experiences in schools and families. *American Journal of Education*. 99, 90-101.

- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, goals. *Annual Review of Psychology*, 53, 109-132.
- Eichhorn, D.H. (1973). The Boyce medical study. *Educational dimensions of emerging adolescent learner*. Springfield: Educational Leadership Institute, Inc.
- Eichhorn, D.H. (1966). *The middle school*. New York, NY: Center for Applied Research.
- Electronic Learning Systems. (2006). Carry-a-Tune Learn to Sing Program. Tampa, Florida.
- Elish-Piper, L., & Tatum, A. W. (2006). Addressing the literacy needs of adolescent students: Listening to their voice. *The NERA Journal*, 42(1), 1-12.
- Englert, C. S., & Palincsar, A. S. (1991). Reconsidering instructional research in literacy from a sociocultural perspective. *Learning Disabilities Research & Practice*, 6, 225-29.
- Fisher, D., & Ivey, G. (2005). Literacy and language as learning in the content-area classes: A departure from “Every Teacher a Teacher of Reading”. *Action in Teacher Education*, 27(2). San Diego, CA.
- Flesch, R. F. (1940). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- French, D. (2003). A new vision of authentic assessment to overcome the flaws in high stakes testing. *Middle School Journal*, 35, 1-12.
- Fry, E. (1977). Fry’s Readability Graph: Validity, and Extension to Level 17. *Journal of Reading*, 21, 249.

- Gates, A.L. (1935). *The improvement of reading: A program of diagnostic and remedial methods*. New York: MacMillian.
- Gillet, J., Temple, C. & Crawford, A. (2004). *Understanding reading problems*, (6th Ed.). Boston: Pearson.
- Glass, G.V., & Hopkins, K. D. (1996). *Statistical Methods in Education and Psychology*. (3rd Ed.). Boston: Allyn and Bacon.
- Glaser, B.G. & Strauss, A.L. (1967). *The discovery of grounded theory*. Chicago: Aldine Publishing Company.
- Goetze, M., Cooper, N., & Brown, C. (1990). Recent research on singing in the general music classroom. *Bulletin of the Council for Research in Music Education*, 104, 16-37.
- Good, C.V. (1927). *The supplemental reading assignment*. Baltimore: Warwick & York.
- Gordon E.E. (2003). *Learning sequences in music: Skill, Content, and Patterns*. Chicago: GIA Publications.
- Gray, W.S. (1919). *The relationship between study and reading* (Proceedings of the Annual Meeting of the National Education Association). Washington, DC: National Education Association.
- Gray, W.S. (1937). The nature and organization of basic instruction in reading. In G. M. Whipple (Ed), *The teaching of reading*. (Thirty-sixth Yearbook of the National Society for the Study of Education Part I). Bloomington: Public Schools Publication Co.

- Gray, W.S. (1939). Reading. In G. M. Whipple (Ed), *Child development and curriculum*. (Thirty-eighth Yearbook of the National Society for the Study of Education Part I). Bloomington: Public Schools Publication Co.
- Green, J.C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed methods evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255-274.
- Greenberg, M.T., Siegel, J.M., & Leitch, C.J. (1982). The nature and importance of attachment relationships to parents and peers during adolescence. *Journal of Youth and Adolescence* 12(5), 373-386.
- Guthrie, J.T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. Kamil, P.B. Monsental, P.D. Pearson & R. Barr (Eds.) *Handbook of Reading Research*, 3, 403-422.
- Guthrie, J. T., & Davis, M. H. (2003). Motivating struggling readers in middle school through an engagement mode of classroom practice. *Reading and Writing Quarterly*, 19, 59-84.
- Hall, G. S. (1908). Social significance of adolescence. *Journal of Psychology*, 19, 149-212.
- Hall, L., Boone, N., Grashel, J. & Watkins, R., Eds. (1997). *Strategies for Teaching: Guide for Music Methods Classes*. Reston: Music Educators National Conference.
- Havighurst, R. J. (1972). *Developmental tasks and education*. New York: Davis McKay Co., Inc.
- Herber, H. L. (1970). *Teaching reading in the content area*. New Jersey: Prentice-Hall.

- Hodgkinson, H. (1986). The right school for the right kids. *Educational Leadership*, 45(5), 10-14.
- Huey, E. B. ((1908). *The psychology and pedagogy of reading*. New York: MacMillian.
- Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescents*. New York: Basic Books.
- Ivey, G. (1999). A multicase study in the middle school: Complexities among young adolescent readers. *Reading Research Quarterly*, 34, 172-192.
- Ivey, G., & Broaddus, K. (2001). "Just plain reading": A survey of what makes students want to read in middle school classrooms. *Reading Research Quarterly*, 36(4), 350-377.
- James, W. (1923). *Talks to teachers on psychology: And to students on some life's ideals*. New York: Henry Holt.
- Johnson, H.J. (1984). A synthesis of research findings on middle level education. In J. H. Lounsbury, (Eds.) *Perspectives : middle school education, 1964-1984*. OH: National Middle School Association.
- Johnson, B., & Turner, L.A. (2003). Data collection strategies in mixed methods research. In A. Tashahhori & C. Teddlie (Eds.), *Handbook on mixed methods in the behavioral and social sciences*,(pp.297-320). Thousand Oaks: Sage Publications.
- Johnston, J.H., Markle, G. C., & Stingley, L. (1982). What research says to the practitioner about peer relations in the classroom. *Middle School Journal*, 13(3), 22-26.

- Judd, C.H. & Buswell G. T. (1922). *Silent reading a study of various types* (Supplemental Education Monographs, No. 23). Chicago: University of Chicago Press.
- Juvonen, J., Vi-Nhuan, L., Kaganoff, T., Augustine, C., & Constant, L. (2004). Focus on the wonder years: Challenges facing the American middle school. A Report to the Rand Corporation. Santa Monica, CA.
- Kamil, M. (2002). Adolescence and literacy: Reading for the 21st century. *A Summary of the Research: National Assessment of Education progress.*
- Kilpatrick, W. H. (1919). *The project method*. New York: Teachers College Press.
- Kottmeyer, W. (1944). Improving reading instructions in St. Louis schools. *Elementary School Journal*, 45, 33-41.
- Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition & Instruction*, 18, 495-523.
- Kuhn, M.R., & Stahl, S. A. (2003). *Fluency: A review of development and remedial practices*. Ann Arbor: Center for the Improvement of Early Reading Achievement.
- LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.
- Lamb, S.J., & Gregory, A.H. (1993). The relationship between music and reading in beginning readers. *Educational Psychology*, 13, 19-27.
- Lee, C. D. (2001). Is October Brown Chinese? A cultural modeling activity system for underachieving students. *American Educational Research Journal*, 38, 97-141.

- Lenz, B. K., & Deshler, D. D. (2003). *Teaching content to all: Evidence-based inclusive practices in middle and secondary schools*. Boston, MA: Allyn and Bacon.
- Leslie, L., & Caldwell, J. (2006). *Qualitative Reading Inventory-4*. Boston, MA.: Allyn & Bacon.
- Leu, D.J., Jr. (2000). Literacy and technology: Deictic consequences for literacy education. In an information age. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research*, 3, 743-770. New York: Longman.
- Lipsitz, J. (1980). *Growing up forgotten: A review of research and program concerning early adolescence*. New Brunswick, NJ: Transaction Books.
- Lounsbury, J.H. (1984). *Perspectives: middle school education, 1964-1984*. Ohio: National Middle School Association.
- Lounsbury, J.H., & Vars, G.F. (1978). *A curriculum in the middle school years*. New York : Harper and Bow.
- Mac Iver, D. J. , & Epstein, J.L. (1991). Responsive practices in the middle grades: Teacher teams, advisory groups, remedial instruction, and school transition programs. *American Journal of Education*, 99(4), 587-622.
- Marger, R. F., (1968). *Developing attitudes towards learning*. Belmont: Fearon.
- McCallister, J. M. (1930). Reading difficulties in studying content subjects. *Elementary School Journal*, 31, 191-201.
- McGuire, Mary Sweatt (2004) Making the lyrics sing for struggling readers: An insider's view. Ed.D. dissertation, Boise State University, United States -- Idaho. Retrieved June 13, 2007, from ProQuest Digital Dissertations database.

- McCallister, J. M. (1932). Determining the types of reading in studying content subjects. *School Review*, 40, 115-123.
- McCombs, J. S., Kirby, S. N., Barney, H., & Magee, S. (2005). Achieving state and national literacy goals, a long uphill road: A Rand Report to Carnegie Corporation of New York (2nd Ed.). Washington, DC: Alliance for Excellent Education.
- McKee, P.G. (1934). *Reading and literature in the elementary school*. Boston: Houghton Mifflin.
- McPartland, J.M. (1987). *Balancing high quality subject-matter instruction with positive teacher-student relations in the middle grades: Effects of departmentalization, tracking and block scheduling on learning environments*. (Report No. 15), Maryland: VSO Industries.
- Merriam, S.B. & Associates (2002). *Qualitative research in practice: Examples for discussion and analysis*. California: Jossey-Bass.
- Mariotti, A.S., & Homan, S. P. (2005). *Linking reading assessment to instruction: An application worktext for elementary classroom teachers*. (5th Ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Miles, M. & Huberman, A.M. (1994). *Data analysis: An expanded sourcebook*. Thousand Oaks, California: Sage.
- Moje, E.B. (1996). "I teach students, not subjects": Teacher-student relationships as contexts for secondary literacy. *Reading Research Quarterly*, 31(2), 172-195.
- Moje, E.B., Young, J.P., Readence, J. E., & Moore, D. W. (2000). Reinventing adolescent literacy for new times: Perennial and millennial issues. *Journal of Adolescent & Adult Literacy*, 43(5), 400-410.

- Moll, L. C., Gonzalez, N. (1994). Critical issues: Lessons from the research with language minority children. *JRB: A Journal of Adolescents & Adult Literacy*, 26, 439-456.
- Moore, D. W. (1996). Contents for literacy in a secondary school. In D.J. Leu, C.K. Kinzer, & K.A. Hinchman (Eds.). *Literacies for the 21st century: Research and practice*. Chicago: National Reading Conference.
- Moore, D. W., Readence, J. E., & Rickleman, R. J. (1983). An historical exploration of content area reading instruction. *Reading Research Quarterly*, 18(4), 419-438.
- National Center for Education Statistics, National Assessment of Educational Progress: The Nation's Report Card, Web site, 2005. Retrieved January 7, 2007, from <http://nces.ed.gov/nationsreportcard>.
- National Reading Panel Report. (2000). *Teaching children to read: An evidence-based assessment of scientific-research literature on reading and its implications for reading instruction*. Washington, D.C.: National Institute of Child Health and Human Development.
- No Child Left Behind Act, Department of Education, Washington, D.C. Retrieved January 7, 2007 from <http://www.ed.gov/nclb/landing.jhtml>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd Ed.). New York: McGraw-Hill.
- O'Brien, D. G., Stewart, R. A., & Moje, E. B. (1995). Why content literacy is difficult to infuse into the secondary school: Complexities of curriculum, pedagogy, and school culture. *Reading Research Quarterly*, 30, 442-463.

- O'Shea, L.J., Sindelar, P.T., & O'Shea, D. J. (1995). The effects of repeated reading and attentional clues on reading fluency and comprehension. *Journal of Reading Behavior*, 17, 129-142.
- Onwuegbuzie, A. J. (2003). Expanding the framework of internal and external validity in quantitative research. *Research in the Schools*, 10(1), 71-89.
- Pearson, P.D. (1985). Changing the face of reading comprehension instruction. *Reading Teacher*, 38, 724-738.
- Pearson, P.D., & Gallagher, M.C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8, 317-344.
- Pearson, P.D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of Reading Research* 2, 815-860. New York: Longman.
- Pearson, P.D., Ferdig, R. E., Blomeyer, Jr, R. L. & Moran, J. (2005). The effects of technology on reading performance in the middle-school grades: A meta-analysis with recommendations for policy. Education, contract number ED-01-CO-0011.
- Phelps, S. (2005). Ten years of research on adolescent literacy, 1994-2004: A review. Learning Point Associates. North Central Regional Laboratory, Institute of Education Sciences, U.S. Department of Education, contract number ED-01-CO-0011.
- Piaget, J. (1964). Development and Learning. In R.E. Ripple and V. N. Rockcastle, *Piaget Rediscovered*. New York: Cornell University Press.

- Pikulski, J.J. (1991). The transitions years: Middle school. In J. Flood, J.M. Jensen, D. Lapp, & J. Squire (Eds.), *Handbook of Research on Teaching English Language Arts* (pp. 303-319). New York: Macmillan Publishing Company.
- Pikulski, J.J., & Chard, D.J. (2005). Fluency: Bridge between decoding and reading comprehension. *The Reading Teacher*, 58(6), 510-519.
- Pressley, L.C. (1923). *The technical vocabulary of public school subjects*. Bloomington: Public School Publishing Co.
- Rasinski, T.V. (2004). Creating fluent readers. *Educational Leadership*, 61(6), 46-51.
- Resnick , D. P. & Rensnick, L. B. (1977). The nature of literacy: An historical exploration. *Harvard Educational Review*, 47, 370-385.
- Rice, J.M. (1893). *The public-school system of the United States*. New York: Century.
- Ritter, B. T. , & Lofland, W. T. (1924). The relationship between reading ability as measured by certain standard tests and the ability required in interpretation of printed matter involving reason. *Elementary School Journal*, 24, 529-546.
- Rosenshine, B. (1980). How time is spent in elementary classrooms. In Denham and Leiberman (Eds.). *Time to learn*. Washington: Department of Health, Education and Welfare, National Institute of Education.
- Rothstein, R. (2000). How standardized test drop the ball. *New York Times*, September, 2000.
- Ryan, A. M. (2000). Peer group as a context for the socialization of adolescents' motivation, engagement, and achievement in school. *Educational Psychologist*, 35(2), 101-111.

- Sackes, P. (2000). *Standardized minds: The high price of America's testing culture and what we can do to change it*. Perseus Publishing: Cambridge, MA.
- Salvin, R.E (1988). Synthesis of research on grouping in elementary and secondary schools. *Educational Leadership*, 46(1), 67-77.
- Samuels, S.J. (1979). The method of repeated readings. *The Reading Teacher*, 41, 756-760.
- Samuels, S.J & Farstrup, A.E. (2006). *What research has to say about fluency instruction*. Delaware: International Reading Association.
- Sample, K.J. (2005). Promoting fluency in adolescents with reading difficulties. *Interventions in Schools and Clinics*, 40(4). 243-246.
- Schunk, D.H., & Zimmermann, B. J. (1916). Modeling and self-efficacy influences on children's development and self-regulation. In J. Juvonen & K. Wentzel (Eds.). *Social motivation: Understanding children's school adjustment*, (pp. 154-180).Cambridge, England: Cambridge University Press.
- Schreiber, P.A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior*. 12, 177-186.
- Schreiber, P.A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz and S. J. Samuels's (Eds.) *Comprehending oral and written language*. (pp. 243-270). New York: Academic Press.
- Schreiber, P.A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice*, 30, 158-164.
- Schreiber, P.S., & Read, C. (1980). Children's use of phonetic cues in spelling, parsing, and maybe- reading. *Bulletin of the Orton Society*, 30, 209-224.

- Schunk, D. H. & Rice, J. M. (1993). Strategy fading and progress feedback: Effects of self-efficacy and comprehension among students receiving remedial reading service. *Journal of Special Education, 27*, 257-276.
- Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
- Shaywitz, B.A., Pugh, K. R., Jenner, A. R., Fulbright, R. K., Fletcher, J. M., Gore, J. C., & Shaywitz, S. E. (2000). The neurobiology of reading disability (dyslexia). In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, R. Barr (Eds.). *Handbook of Reading Research, 3*, 141-151. Mahwah New York: Erlbaum.
- Simmons, R.G., & Blyth, D.A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. Hawthorn: Random House.
- Smith, N. B. (1919). Correlations of ability in reading with the general grades in high school. *School Review, 27*, 493-511.
- SPSS Inc. (2006). SPSS Base 15.0 for Windows. SPSS Inc., Chicago IL.
- Spring, J. (1986). *The American school*. New York: Longman.
- Stake, R.E. (2000). Case studies. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of Qualitative Research, (2nd Ed.)* (pp. 509-536). Thousand Oaks: Sage Publications.
- Stayter, F.Z., & Allington, R.L., (1991). Fluency and the understanding of text. *Theory into practice, 30*(3), 143-148.
- Stanovich, K.D. (1980). Effects of explicit teaching and peer tutoring on the reading achievement of learning disabled and low performing students in regular classrooms. *Reading Research Quarterly, 16*, 32-71.

- Stanovich, K.D. (1984) The interactive compensatory model of reading. A confluence of Developmental, experimental, and educational psychology. *Remedial and Special Education*, 5(3), 11-19
- State of Florida Department of Education (2001). Understanding the FCAT reports 2006. Retrieved January 7, 2007, from <http://www.firn.edu/doe/sas/fcat.htm>
- Stevens, J.P. (1999). *Intermediate statistics: A modern approach*. Mahwah, NJ: Erlbaum.
- Stevens, J.P. (2002). *Applied multivariate statistics for the social sciences*, (4th Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*, (2nd Ed.). CA: Harper Collins Publishers.
- Tanner, J. M., (1962). *Growth at adolescence*. 2nd ed., Oxford England: Blackwell Scientific Publications.
- Tashakkori, A., & Teddlie, C. (2003). *Mixed methodology: Combining qualitative and quantitative approaches*. (Eds.). Thousand Oaks, CA: Sage Publications.
- Taylor, N. E., Wade, M. B., & Yekovich, F. R. (1985). The effects of test manipulation and multiple reading strategies on the reading performance of good and poor readers. *Reading Research Quarterly*, 20, 506-574.
- Therrien, W.J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. *Remedial and Special Education*, 25 (4), 252-261.
- Thorndike, E.L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology*, 8, 276-282.
- Thorndike, E.L. (1921). *The teachers word book*. New York: Columbia University Teachers College.

- Topping, K.J. (2006). Building reading fluency: Cognitive, behavioral, and socioemotional factors and the role of peer-mediated learning. In S.J Samuels, A.E. Farstrup, (Eds.), *What research has to say about fluency instruction* (pp. 106-129). Delaware: International Reading Association.
- Tovani, C. (2000). *I read it but I don't get it*. Portland, ME: Steinhouse.
- Vacca, R. T. & Vacca, J. L. (2002). *Content area reading: Literacy and learning across the curriculum* (7th ed.). Boston MA: Allyn & Bacon.
- Van Til, W.G., & Lounsbury, J. (1961). *Modern education for the junior high school years*. New York: Boobs-Merril.
- Vygotsky, L. (1978). *Mind and society*. Cambridge, MA. Harvard University Press.
- Wade, S.E., & Moje, E.B. (2000). The role of text in classroom learning. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, R. Barr (Eds.). *Handbook of reading research* (Vol. 3, pp. 609-627). Mahwah New York: Erlbaum.
- Wagner, E. (1938). *Reading and ninth grade achievement*. (Teachers College Contributions to Education, No. 756). New York: Columbia University Teachers College.
- Webster, P. R. (2002). Computer-based technology and music teaching and learning. In R. Colwell& C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning*. (pp. 416-439).
- Wentzel, K.R. (1996). Social and academic motivation in middle school: Concurrent and Long term relations to academic effort. *Journal of Early Adolescences*. 16(4), 390-406.

- Whipple, G. M. (1920). *New materials of instruction*. (Nineteenth Yearbook of the National Society for the Study of Education Part I) Bloomington: Public School Publishing Co.
- Wiles, J., & Thomason, J. (1975). Middle school research 1968-1974: A review of substantial studies. *Educational Leadership*, 32, 412-423.
- Wiley, K.B. (1977). *The status of pre-college science, mathematics, and social science educational practices in the U.S schools: An overview and summaries of three studies*. Washington DC: U.S. Government Printing Office.
- Witty, P. A. (1948). Current role and effectiveness of reading among youth. In N. B. Henry (Ed.), *Reading in the high school and college* (Forty-Seventh Yearbook of the National Society for the Study of Education Part I) Bloomington: Public School Publishing Co.
- Worthy, J., Moorman, M. & Turner, M. (1999). What Johnny likes to read is hard to find in school. *Reading Research Quarterly*, 3(1), 12-27.
- Zimmerman B. J. (2000). Attaining self-regulation: A social –cognitive perspective. In M. Boekaerts, P.R., Pintrich, and M.H. Zeidner. *Handbook of self-regulation* (Vol. 41, pp. 64-70). San Diego CA: Academic Press.

APPENDICES

Appendix A: Qualitative Reading Inventory-4 (QRI-4)

Level: Six



Narrative

Concept Questions:
Who is Pele?

(3-2-1-0)
What is soccer?

(3-2-1-0)
What are professional athletes?

(3-2-1-0)
Why do some sports become popular?

(3-2-1-0)
Score: _____ /12 = _____ %
_____ FAM _____ UNFAM
Prediction:

"Pele"

Pele was born in the South American country of Brazil. He lived in a small village and his family was very poor. But Pele had a dream. He wanted to be-

come a professional soccer player. He could not afford a soccer ball so he fashioned one. He took an old sock, stuffed it with newspapers, and sewed it together with string. It was a poor substitute, but it was better than nothing. Pele and his friends formed their own team. They did not have enough money to purchase shoes, but that did not stop them. They played barefoot and became known as the "barefoot team."

Pele and his friends saved their money, and eventually the team was able to get a regular ball and shoes. Pele discovered that the ball could be better controlled when he wore shoes. Pele and his team practiced continuously. They soon began playing older and more established teams from the big cities. The team began to win most of its games. Pele was the star of the team. People thought this was amazing because he was only eleven years old!

Pele's skill at soccer came to the attention of influential people, and when he was fifteen, he was signed by the Santos team. Pele led the Santos team to many championships. He also led the Brazilian national team to three world championships. Pele also holds many records and has scored over twelve hundred goals in his career as a professional player.

Pele decided to retire in 1974. Then he changed his mind and came to the United States, where he joined the New York Cosmos. Soccer had not been very popular in the United States up to this point,

Appendix A: (Continued)

Level: Six

but Pele's presence had a dramatic effect. Crowds at games doubled and tripled as people came to see the famous and exciting Pele. Games began to be shown on television. Soccer gained in popularity and many children in the United States began to play soccer. Soccer is now one of the most popular sports in the United States, due in part to the dream of a young boy in Brazil. (351 words)

From *Holt Social Studies*, edited by JoAnn Cangemi, copyright © 1983 by Holt, Rinehart, and Winston, Inc. Reprinted by permission of Harcourt, Inc.

Number of Total Miscues (Total Accuracy): _____		
Number of Meaning-Change Miscues (Total Acceptability): _____		
Total Accuracy		Total Acceptability
0-8 miscues	_____ Independent	0-8 miscues
9-36 miscues	_____ Instructional	9-19 miscues
37+ miscues	_____ Frustration	20+ miscues
Rate: $351 \times 60 = 21,060 /$ _____ seconds = _____ WPM		
_____ WPM - _____ errors = _____ CWPM		

Retelling Scoring Sheet for "Pele"

Setting/Background

- _____ Pele was born
- _____ in Brazil.
- _____ Pele's family was poor.

Goal

- _____ Pele had a dream.
- _____ He wanted

- _____ to become a soccer player
- _____ a professional player.

Events

- _____ He could not afford a ball.
- _____ He fashioned a ball.
- _____ He took a sock
- _____ and stuffed it
- _____ with newspapers
- _____ and sewed it together
- _____ with string.
- _____ Pele
- _____ and his friends formed a team
- _____ their own team.
- _____ They did not have enough money
- _____ to purchase shoes.
- _____ They played barefoot
- _____ and became known
- _____ as the "barefoot team."
- _____ Pele
- _____ and his friends saved their money,
- _____ and eventually
- _____ the team was able
- _____ to get a ball
- _____ a regular ball
- _____ and shoes.
- _____ Pele discovered
- _____ that the ball could be controlled
- _____ better
- _____ when he wore shoes.
- _____ They began
- _____ to play teams
- _____ from big cities
- _____ and to win games
- _____ most of their games.
- _____ Pele was the star.
- _____ He was only eleven.
- _____ He was signed
- _____ by the Santos team.

Resolution

- _____ Pele led the team
- _____ to championships.
- _____ He led the team
- _____ the Brazilian team

Appendix A: (Continued)

Level: Six

- ___ to championships
- ___ world championships.
- ___ Pele held many records.
- ___ Pele decided
- ___ to retire
- ___ in 1974.
- ___ Then he changed his mind
- ___ and came to the United States,
- ___ where he joined the Cosmos
- ___ the New York Cosmos.
- ___ Pele's presence had an effect.
- ___ Crowds doubled
- ___ and tripled
- ___ and people came
- ___ to see Pele.
- ___ Soccer gained
- ___ in popularity.
- ___ Soccer is now one
- ___ of the most popular sports
- ___ in the United States

66 Ideas

Number of ideas recalled _____

Other ideas recalled, including inferences:

Questions for "Pele"

1. What was Pele's main goal?
Implicit: to become a professional soccer player
2. What did Pele use to make a soccer ball?
Explicit: an old sock, string, and newspaper
3. What was Pele's team called when they had no shoes?
Explicit: the barefoot team

4. Why did the purchase of shoes affect the number of games won by Pele's team?
Implicit: they played better because they could control the ball more effectively
5. Why was it amazing that Pele became a star at the age of eleven?
Implicit: he was very young to be playing so well against older and more established teams from the big cities
6. How old was Pele when he was signed by a professional soccer team?
Explicit: fifteen
7. What American team did Pele join?
Explicit: New York Cosmos
8. How did Pele's presence help to make soccer popular in the United States?
Implicit: people came to see Pele and grew to like the game itself

Without Look-Backs

Number Correct Explicit: _____

Number Correct Implicit: _____

Total: _____

___ Independent: 8 correct

___ Instructional: 6-7 correct

___ Frustration: 0-5 correct

With Look-Backs

Number Correct Explicit: _____

Number Correct Implicit: _____

Total: _____

___ Independent: 8 correct

___ Instructional: 6-7 correct

___ Frustration: 0-5 correct

Appendix B: Qualitative Matrix

An Example of the Qualitative Classroom Observation Notes Transferred to the Categorical Matrix for the Peer Interactions

<p>Information Exchange Peer discussion/talk direct quotes from conversations during the literacy task</p>	<p>Peer 1 “ How did you get the song to slow down” Peer 2 “ Click on this button”</p>
<p>Modeling Peer Observation/ through descriptions of interactions during the literacy task</p>	<p>He looked around the classroom started to smile and went back to playing the drums</p>
<p>Peer Pressure Social reinforcement/ descriptions through looks / comments/ laughs during the literacy task</p>	<p>T hit the drum wrong, M laughed and then the class laughed T turned red and put his head down,</p>

Appendix C: Construct Key
Peer Interactions: During the Literacy Task of Rereading Through Singing

Construct Heading	Construct	Definition
Group Characteristics	Extrinsic Motivation	Students engaging in a task because of a reward or punishment.
Group Characteristics	Intrinsic Motivation	Students engaging in a task for their own personal learning.
Group Characteristics	Peer Observations	Peers' observing each other that influences behavior changes.
Group Characteristics	Peer Hierarchy	Social system in the classroom that positions some members of the peer group above other.
Group Characteristics	Peer Support	Peers providing or showing support for one another
Group Characteristics	Dominant and Vulnerable Peers	Peer positions that place dominant peers over the more vulnerable peers.
Group Characteristics	Students' Perceptions of Alternative Text	How the students understand and perceive the alternative text.
Group Characteristics	Disequilibrium	A cognitive state that occurs as new and different information occurs in the learning.
Group Characteristics	Fake Rereading/Singing	The students appear to be singing the songs however, they are not.
Group Characteristics	Peer Leaders	Students in the group take on the role of leadership over the other students.
Instructional Procedures	Alternative Approaches to Singing	Teaching approaches different to practice singing that used the drums.
Instructional Procedures	Safe Risk-Free Environment	A setting where the students feel comfortable enough to take a risk.

About the Author

Marie C. Biggs is currently an assistant professor at St. Petersburg College of Education. Her background is in the area of reading research with an emphasis on working with struggling “at risk” readers. Over the last twenty years she has held positions as a reading specialist for K-12 students and faculty positions as an instructor in Literacy Education and Assessment at Wheelock College in Boston and St. Petersburg College in Tampa, and the University of South Florida where she taught undergraduates and graduate preservice teachers. Also, in her fieldwork as a professional development coordinator and consultant through Tufts University in Boston she provided outreach services and taught graduate courses in reading to in-service teachers in various school systems throughout New England.

Currently, her research interests are focused on the use of Alternative Texts for Struggling Adolescent Readers grades 4-12. In 2004, she conducted a pilot study using an interactive singing software program, *Carry-a-Tune*, with middle school struggling readers. This was followed by becoming the lead field researcher in a seven-site study of 4th through 12th grade struggling readers in the Tampa area, who were using the *Carry-a-Tune* program. In addition, she was the coordinator for a multi-district study with 1,300 ELL across three districts in Florida. Her implications and findings from the research have been presented throughout the United States at conferences that focus in the field of Literacy, Music, Technology, and Educational Leadership.