THE EFFECTS OF MINDPLAY VIRTUAL READING COACH (MVRC) ON THE SPELLING GROWTH OF STUDENTS IN SECOND GRADE

by

Breanna L. Sherrow

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF DISABILITY AND PSYCHOEDUCATIONAL STUDIES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

WITH A MAJOR IN SPECIAL EDUCATION

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Breanna L. Sherrow, titled The Effects of MindPlay Virtual Reading Coach (MVRC) on the Spelling Growth of Students in Second Grade and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: April 22, 2015
Dr. Nancy Mather

_____ Date: April 22, 2015
Dr. Carl Liaupsin

_____ Date: April 22, 2015
Dr. John Umbreit

_____ Date: April 22, 2015
Dr. Janice Sammons

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: April 22, 2015
Dissertation Director: Dr. Nancy Mather

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

SIGNED: Breanna L. Sherrow

ACKNOWLEDGEMENTS

This dissertation and the process of completing my doctoral studies would not have been complete without the effort and support of many people. First and foremost, I'd like to thank my friends and family. I owe my deepest thanks to all of you, for your never ending support and for keeping me sane during my education endeavors.

I'd like to acknowledge the members of my dissertation committee for their feedback and continued support and for helping me through the growing pains of this entire process. My sincere gratitude to my dissertation committee chair, Dr. Nancy Mather, for your encouragement, the opportunities you provided along the way, and your countless hours of editing. Dr. Carl Liaupsin, I thank you for providing your support and mentorship which has been invaluable. Also, for allowing me the pleasure to teach alongside you. Dr. John Umbreit, thank you for serving as my minor advisor, answering my questions along the way, and pushing me to always thoroughly think something out before beginning. Dr. Janice Sammons, thank you for your friendship, advice, and for engraining in me that a good dissertation is a done dissertation.

I'd like to thank Dr. Mark Borgstrom, for his statistical knowledge and suggestions he provided along the way for multilevel modeling. I am eternally grateful and happy to say I am no longer intimidated by multilevel modeling.

I would like to thank Graduate Christian Fellowship (GCF) for grounding me in the Lord and introducing me to some very wonderful individuals who I am happy to now call my friends. I would mention you all by name but I'd be afraid I'd forget someone. Therefore, I'll make a blanket statement. Thank you to all my fellow GCF members for the never ending support and care you have all provided throughout graduate school which has been extremely important to my studies and spiritual foundations.

To my fellow doctoral students also in the Special Education doctoral program, it has been a great pleasure getting to know all of you and working with all of you. I wish you all the best as you continue with your doctoral studies. Special thanks to Alison Zagona and Jennifer May White. Alison Zagona, has been a dear friend and colleague. Thank you for going on this adventure with me, for supporting me, and keeping me focused on the end goal. I never would have finished this dissertation without you. Also, thank you for your feedback and editing at every stage. Thank you to Jennifer May White, for your friendship and for spending many hours scoring spelling tests for inter-observer agreement.

I would also like to show my gratitude to the professors in the Department of Disability and Psychoeducational Studies who have instilled confidence in me, provided insightful responses, and challenged me. In addition, I'd also like to thank the wonderful administrative support of Cecilia, Kevin, and Sandy.

A special thanks to all of the teachers and individuals who motivated me to always give 100% and nothing less. One thing I will also remember is to follow your dreams and your passions and stay true to yourself.

Additionally, I would like to thank MindPlay Educational Software for Reading Instruction and the Chief Financial Officer, Judith Bliss for making this dissertation possible by sharing an archival dataset with me.

DEDICATION

I lovingly dedicate this dissertation to my parents and in memory of my brother. Mom and Dad, thank you for your constant love and support throughout my life. You are both a large part of who I am today and how I got here. Thank you for your words of encouragement and for pushing me to persevere over many challenges. Ryan, you will always hold a special place in my heart and I know you have been supporting me from above. May you always be my guardian angel.

TABLE OF CONTENTS

LIST OF FIGURES 9

LIST OF TABLES 11

ABSTRACT 14

CHAPTER 1: INTRODUCTION 15
- Theoretical Perspectives and Spelling Assessment 16
- Computer-Delivered Interventions 18
- Purpose.....19
- Research Questions 19
- Defintion of Terms 20

CHAPTER 2: REVEW OF LITERATURE 22
- The Relationship Between Word Reading and Spelling 22
- Theoretical Approaches to Spelling Development 24
  - Stage Theories.....24
    - Ehri 24
    - Frith 26
    - Gentry 26
    - Henderson 27
  - Strategy Approach 28
- The Differences in Reading and Spelling Achievement of Males and Females, ELLs, and Students in SPED 29
- Spelling Assessment 32
- Curriculum-Based Measurement Scoring for Spelling 33
- Computer-Based Interventions for Promoting Reading and Spelling Growth 34
  - Leescircus 35
  - MultiFunk 37
  - ABRACADABRA 38
  - IntelliTools 41

CHAPTER 3: METHOD 43
- Research Questions 43
- Data Source 43
- Research Design 44
  - Setting and Participants.....44
  - Group Assignment 46
  - Description of Measures: Woodcock-Johnson IV Achievement 47
  - Curriculum-Based Measurement Spelling Scoring for the Qualitative Analysis 48
  - Dependent Variable 48
  - Independent Variable.....49
    - Intervention description 49
  - Student Predictors.....50

TABLE OF CONTENTS-Continued

Description of Classroom Instruction 50
Intervention and Comparison Procedures 51
Experimental Setting.....51
Teacher Training.....52
Procedures: Experimental Instruction 52
Procedures: Comparison Instruction.....52
Data Collection 53
Administration of Dependent Variable Measures.....53
Scoring Reliability (Inter-Observer Agreement) 54
Fidelity of the Intervention 54
Data Analysis 55
Why Use Multilevel Modeling?..... 55
Preparing the Data.....58
Data Analysis Plan 60

CHAPTER 4: RESULTS 65
Fidelity of Implementation 65
Product Usage Log 65
Planned Activities Check and Inter-Observer Agreement of Student Engagement 66
Inter-Observer Agreement of Scoring the Spelling Tests 67
Comparison 67
Experimental 68
Results of Multilevel Modeling of Repeated Measures Analyses 69
Visual Inspection of Data for Both Tests.....69
Multilevel Model for Repeated Measures Formulas 71
Multilevel Model Building Process 73
Null model for test three 74
Within-individual model for test three.....77
Between-individual model for test three.....79
Between-group model for test three.....82
Residual analysis for test three 87
Null model for test sixteen.....88
Within-individual model for test sixteen 91
Between-individual model for test sixteen 94
Between-group model for test sixteen 97
Residual analysis for test sixteen.....100
Qualitative Spelling Analysis for Students' Written Spellings 101
Scoring Evaluation for Test Three 102
Scoring Evaluation for Test Sixteen 102

CHAPTER 5: DISCUSSION 104
Summary of Findings for Each Research Question 104
Summary of Findings for Research Question One 104

TABLE OF CONTENTS-Continued

Summary of Findings for Research Question Two 105
Summary of Findings for Research Question Three 105
Limitations 106
Test Administration 106
Success for All 107
Planned Activity Checks and Inter-Observer Agreement (IOA) 107
Additional Data.....108
Sample Size.....108
Outcome Measures 108
Recommendations and Implications for Future Research 108
Future Research 110

APPENDIX A-FIGURES 112

APPENDIX B-PLANNED ACTIVITY CHECK FORM 133

APPENDIX C-ADDITIONAL FIGURES 134

REFERENCES 175

LIST OF FIGURES

Figure 1. The three levels of the multilevel model of repeated measures. 113

Figure 2. Test 3 spelling growth trajectories for all students with traditional scoring. 114

Figure 3. Test 3 spelling growth trajectories for all students with non-traditional scoring. 115

Figure 4. Test 16 spelling growth trajectories for all students with traditional scoring. 116

Figure 5. Test 16 spelling growth trajectories for all students with non-traditional scoring. 117

Figure 6. Mean of comparison and experimental groups for test 3 with traditional scoring. 118

Figure 7. Mean of comparison and experimental groups for test 3 with non-traditional scoring. 119

Figure 8. Mean of comparison and experimental groups for test 16 with traditional scoring. 120

Figure 9. Mean of comparison and experimental groups for test 16 with non-traditional scoring. 121

Figure 10. Subset of individual spelling growth trajectories for test 3 with traditional scoring. 122

Figure 11. Subset of individual spelling growth trajectories for test 3 with non-traditional scoring. 123

Figure 12. Subset of individual spelling growth trajectories for test 16 with traditional scoring. 124

Figure 13. Subset of individual spelling growth trajectories for test 16 with non-traditional scoring. 125

Figure 14. Mean of students by teacher for test 3 with traditional scoring. 126

LIST OF FIGURES-Continued

Figure 15. Mean of students by teacher for test 3 with non-traditional scoring. 127

Figure 16. Mean of students by teacher for test 16 with traditional scoring. 128

Figure 17. Mean of students by teacher for test 16 with non-traditional scoring. 129

Figure 18. Histogram of the between-group model residuals. 130

Figure 19. Histogram of the null model residuals. 131

Figure 20. Histogram of the between-group model residuals. 132

LIST OF TABLES

Table 1. Frequency Counts by Teacher 47

Table 2. Frequency Counts of Nominal Data for Gender 50

Table 3. Frequency Counts of Nominal Data for Special Education 50

Table 4. Frequency Counts of Nominal Data for English Language Learners 50

Table 5. Data Definition of SPSS Data set 59

Table 6. Results of the Planned Activities Check for Student Engagement Behavior 67

Table 7. IOA Scoring Results by Teacher for the Comparison Group 68

Table 8. IOA Scoring Results by Teacher for the Experimental Group 69

Table 9. Variability of the Student Subset (N=16) 70

Table 10. Information Criteria for the Null Model for Test Three 75

Table 11. Estimates of Fixed Effects for the Null Model for Test Three 75

Table 12. Estimates of Covariance Parameters for the Null Model for Test Three 76

Table 13. ICC Calculations for the Null Model for Test Three 77

Table 14. Information Criteria for the Within-Individual Model with Time and Treatment Parameters for Test Three 78

Table 15. Estimated of Fixed Effects for the Within-Individual Model with Time and Treatment Parameters for Test Three 79

Table 16. Estimates of Covariance Parameters for the Within-Individual Model with Time and Treatment Parameters for Test Three 79

Table 17. Information Criteria for the Between-Individual Model for Test Three 80

Table 18. Estimated of Fixed Effects for the Between-Individual Model for Test Three 81

Table 19. Estimates of Covariance Parameters for the Between-Individual Model for Test Three 82

LIST OF TABLES-Continued

Table 20. Information Criteria for the Between-Group Model for Test Three 83

Table 21. Estimated of Fixed Effects for the Between-Group Model for Test Three 83

Table 22. Estimates of Covariance Parameters for the Between-Group Model for Test Three 85

Table 23. Random Effects Covariance Structures (G) for the Between-Group Model for Test Three 87

Table 24. Random Effects Covariance Structures (G) for the Between-Group Model for Teacher, Test Three 87

Table 25. Hypothesis Test Summary for the Between-Group Residuals for Test Three 88

Table 26. Information Criteria for the Null Model for Test Sixteen 89

Table 27. Estimates of Fixed Effects for the Null Model for Test Sixteen 90

Table 28. Estimates of Covariance Parameters for the Null Model for Test Sixteen 91

Table 29. ICC Calculations for the Null Model for Test Sixteen 91

Table 30. Information Criteria for the Within-Individual Model with Time and Treatment Parameters for Test Sixteen 93

Table 31. Estimated of Fixed Effects for the Within-Individual Model with Time and Treatment Parameters for Test Sixteen 93

Table 32. Estimates of Covariance Parameters for the Within-Individual Model with Time and Treatment Parameters for Test Sixteen 94

Table 33. Information Criteria for the Between-Individual Model for Test Sixteen 95

Table 34. Estimates of Fixed Effects for the Between-Individual Model for Test Sixteen 95

Table 35. Estimates of Covariance Parameters for the Between-Individual Model for Test Sixteen 97

LIST OF TABLES-Continued

Table 36. Information Criteria for the Between-Individual Model for Test Sixteen 98

Table 37. Estimated of Fixed Effects for the Between-Groups Model for Test Sixteen 98

Table 38. Estimates of Covariance Parameters for the Between-Groups Model for Test Sixteen 99

Table 39. Hypothesis Test Summary for the Between-Group Residuals for Test Sixteen 101

Table 40. Hypothesis Test Summary for the Between-Group Residuals for Test Sixteen 102

Table 41. Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Three 135

Table 42. Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three 145

Table 43. Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen 155

Table 44. Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen 165

ABSTRACT

First, this study was conducted to determine the effects of MVRC on the spelling development of second-graders. Second, this study sought to determine if spelling trajectories vary by gender, English Language Learner (ELL) enrollment and/or Special Education (SPED) enrollment. Lastly, students' spelling tests were evaluated with two different spelling scoring methods: traditional standardized scoring, correct and incorrect, and Curriculum-Based Measurement-spelling (CBM), correct letter sequences, to determine which method was more sensitive to growth from pre-test to post-test. Students were pre-tested and post-tested with two measures from the Woodcock-Johnson IV Achievement, Test 3: Spelling and Test 16: Spelling of Sounds. Participants included 159 students, 83 students were enrolled in the experimental condition and 76 students were enrolled in the comparison condition. Using a multilevel model for repeated measures, the researcher estimated the between group-model analyses for Test 3: Spelling and Test 16: Spelling. Students who participated in the experimental condition, receiving MVRC, had significantly different spelling scores than their peers in the comparison group. For Test 3: Spelling, the experimental group increased on average by 1.786 words compared to the comparison group. For Test 16: Spelling of Sounds, the experimental group increased on average by 1.741 words compared to the comparison group. Student spelling trajectories did vary by gender, ELL enrollment, and SPED enrollment. However, these differences were not found to be significant. Neither traditional scoring nor CBM-spelling scoring was found to be the more sensitive scoring method for growth for both tests. Instead, CBM-spelling was more sensitive for Test 3: Spelling, while traditional scoring was more sensitive for Test 16: Spelling of Sounds.

CHAPTER 1: INTRODUCTION

The abilities to read and spell are integral to a student's achievement not only in school, but also in society (National Association for the Education of Young Children, 1998). Skills gained in reading transfer to spelling and skills gained in spelling transfer to reading (Conrad, 2008). Spelling has been important for reading because "Knowing the spelling of a word makes the representation of it sturdy and accessible for fluent reading" (Snow, Griffin, & Burns, 2005, p. 86). Learning to spell not only increases children's reading, but also their writing (Moats, 2005).

Students who struggle with reading, also find challenges with spelling. A significant amount of research has demonstrated that reading failure has serious consequences for individuals and society as a whole. In fact, a longitudinal study of about 4,000 students confirms that students who are not reading at a proficient level by the end of third-grade will be four times less likely to graduate from high school on time than proficient readers (Hernandez, 2012, p. 4). Also, important to note have been the apparent differences in spelling for males versus females, English Language Learners (ELL), and students with special education (SPED) needs. Often, females outperform males on measures of spelling and females can considerably spell more words than males (Allred, 1990; Berninger, Nielson, Abbott, Wijsman, & Raskind, 2008; Foorman & Petscher, 2010). Foorman and Petscher also reported that English Language Learners and students in Special Education perform poorly on spelling, performing approximately a half a standard deviation below their average peers.

Without additional instruction and intervention students who struggle with reading and spelling often will continue to struggle and fall behind their peers (Alexander & Entwisle, 1988; Foorman & Torgesen, 2001). Students at risk for reading and spelling

difficulties need more comprehensive, systematic, and explicit instruction than has normally been provided in the general education classroom (Foorman & Torgesen, 2001). In addition, early intervention has been critical for struggling readers who have not been improving with typical classroom instruction (Clay, 1993; Morris, Tyner, & Perney, 2000; Snow, Burns, & Griffin, 1998).

Increased demands on teachers from schools, districts, and policymakers to show student reading progress make it difficult to individualize instruction for all learners in one classroom. Thus, teachers often provide large group instruction that is not targeted to an individual student's needs. However, not all students will acquire skills through one single method or instructional approach (Foorman & Torgesen, 2001). Thus, teachers are faced with the challenge of trying to provide individualized literacy instruction to a diverse group of students. In addition, students with learning disabilities in reading also exhibit spelling difficulties. Spelling, more than any other ability, has been the largest discriminator between students with learning disabilities and other students who have low achievement (Fulk & Stormont-Spurgin, 1995). Students who struggle with spelling not only need to receive evidence-based and individualized instruction, but also must have their spelling abilities assessed in a way that accurately represents their abilities. Furthermore, individuals with dyslexia will always experience academic challenges because spelling problems persist throughout their lifespan (Bruck, 1993; Lefly & Pennington, 1991).

Theoretical Perspectives and Spelling Assessment

Just as with reading, students advance through different aspects of spelling development. Stage theories and the strategy approach are two theoretical perspectives

that explain how students' progress developmentally in spelling. Although stage theories and the strategy approach differ, each proposes that students develop increased spelling knowledge over time (Ehri, 1986, 1989; Gentry, 1982, 1984, 1987; Henderson, 1990; Reid, 1988; Weiner, 1994; Young, 2007). Even though the theories of spelling document that growth is developmental, teachers continue to provide students with a list of spelling words on Monday to memorize by the following spelling test on Friday. Instead of assessing spelling development, the emphasis has been on memorization skills. Traditional spelling tests scored as correct or incorrect do not document the student's developmental progress in spelling a word. Ehri (1989) explained, "…if spelling ability is measured only by the number of correct words, this development is totally missed" (p. 357).

Marking spelling words as either correctly spelled or incorrectly spelled only allows us to know whether a student has the ability to completely spell a word or that they are unable to spell the whole word. For example, because some students may only be able to detect the initial and final sounds of a word (often the easiest sounds to hear), a student asked to spell the target word "lion" might initially spell the word "ln." Although the student may not have spelled the entire word correctly, further analysis reveals that this student needs help with the middle vowel sounds of the word but does not need any help with the initial consonant and final consonant sounds. Another example of a common misspelling is when students fail to double a middle consonant in a word, such as "ladder" which is often spelled "lader."

Scoring spelling assessments in alternative ways allows a more accurate measure of what a student can and cannot do. Using the traditional method of scoring "correct" or

"incorrect" does not accurately indicate how well the student may have performed (Rao, Prakash, & Joshi, 2006). Typically, teachers assess spelling by the total number of words correct; however, additional procedures can be used to assess spelling, such as counting the number of correct letter sequences.

Computer-Delivered Interventions

School districts and teachers have been exploring the use of computer-based technologies that were designed to assist students with literacy development to provide individualized and supplemental instruction. Although a vast amount of research has been conducted regarding the application of computer-based reading interventions, specific research on the use of computer-based programs and the impact on spelling achievement has been extremely limited (Comaskey, Savage, & Abrami, 2009; Di Stasio, Savage, & Abrami, 2012; Fasting & Lyster, 2005; Howell, Erickson, Stanger, & Wheaton, 2000; Savage et al., 2010; Van Daal & Reitsma, 2000). Thus, more research is needed in this area to guide the selection of computer-based interventions that can potentially promote spelling skills acquisition and literacy improvement.

Presumably, an intervention that was designed to increase basic reading skills would also result in improvement of basic spelling skills, as Ehri (2000) states they are "two sides of a coin" (p. 19). One recently developed online program, MindPlay Virtual Reading Coach (MVRC),was designed to provide targeted, individualized reading instruction to students in Kindergarten to 12th grade. Because the program was relatively new (released in March 2013), little research exists that explores the efficacy of this program on literacy development, and no studies to date have examined the impact on spelling achievement. MVRC is a computer-based technology for learning that has been web-based and was designed for students and adults with varying proficiency levels of

reading. MVRC positively impact students' and teachers' experiences with learning by offering an internet-based reading intervention program to improve student's reading with one-to-one instruction. One-to-one individualized instruction and intervention is not always possible in a classroom with 25 students and one teacher. Even though a student or adult would sit in front of a computer which may seem to take away from one-to-one instruction, MVRC provides virtual instructors with pre-recorded videos from real-life reading specialists and speech pathologists.

Purpose

The purpose of this study was three-fold. The first purpose was to determine the effects of MVRC on the spelling development of second-grade students from two schools in one district in the Southwest. The second purpose was to determine if spelling achievement trajectories varied by gender, ELL classification, and/or SPED classification. The third purpose was to score the student's spelling tests with two different metrics of spelling: tests that were scored with traditional standardized scoring (correct and incorrect) and tests that were scored with Curriculum-Based Measurement (CBM) scoring for spelling (correct letter sequences) and to decide which was a more sensitive to growth from pre-test to post-test.

Research Questions

The following research questions were posed:

1. Does MVRC have a significant effect on students' spelling scores from the WJ IV ACH for Test 3: Spelling, and Test 16: Spelling of Sounds over time?
2. Do spelling achievement trajectories vary by gender, ELL classification, and/or SPED classification for the WJ IV ACH, Test 3: Spelling, and Test 16: Spelling of Sounds?

3. Which scoring measurement is more sensitive to student growth: traditional standardized scoring (correct or incorrect) or CBM spelling scoring (which credits each correct letter sequence) from the WJ IV ACH Test 3: Spelling, and Test 16: Spelling of Sounds?

Definitions of Terms

The following terms are defined as:

- **Computer-based intervention**: intervention provided on the computer program that supplements regular instruction
- **Correct letter sequences (CLS)**: procedure in which spelling is scored with curriculum-based measurement, points are awarded for each correct letter sequence in the word
- **Curriculum-based measurement**: method monitoring progress that assesses skills in reading, math, spelling, and written expression
- **Differentiated instruction**: philosophy or approach to teach students with different learning styles, giving each student an opportunity to explore the content in the same classroom in a way that fits them best
- **Grapheme**: printed or written symbol that represents a particular sound
- **Individualized instruction**: instruction based on the abilities specific to each learner
- **Nonsense words**: words that are not real and have no meaning but conform to English spelling patterns
- **Non-traditional scoring**: scoring of spelling based on correct letter sequences
- **Phoneme**: single speech sound

- **Phonemic awareness**: the use and manipulation of individual speech sounds
- **Phonics**: instructional reading method for teaching the relationships between sounds and letters and how to represent those sounds in spelling
- **Phonological awareness**: an umbrella term for many different types of sound manipulations, knowing that spoken language is composed of sounds , the ability to manipulate and integrate language sounds
- **Real words**: words that exist in the English language
- **Traditional scoring**: spelling scored as correct or incorrect based on the whole word

CHAPTER 2: REVIEW OF LITERATURE

The purpose of this chapter was to review the literature pertaining to: (a) the relationship between word reading and spelling; (b) theoretical approaches to spelling development; (c) the differences in reading and spelling achievement of males and females, ELLs, and students in SPED; (d) spelling assessment; (e) CBM scoring of spelling; and (f) computer-based interventions for promoting reading and spelling growth.

The Relationship Between Word Reading and Spelling

The exact nature of the relationship between reading and spelling is still not entirely clear; however, researchers have demonstrated that reading and spelling skills are interrelated. There is a strong relationship between early reading and spelling as well phonemic awareness and spelling, as confirmed by correlational studies (Ellis & Large, 1987; Juel, Griffith, & Gough, 1986; Snowling & Perin, 1983). Ehri (2000) suggests that reading and spelling are closely related because each relies on the knowledge of the alphabetic system, specific words, and the memory involved in spelling the specific words. For example, being able to recognize that there is a relationship between letters and sounds is the same foundational knowledge needed by children learning to read but also learning to spell (Ehri, 2000; Moats, 2005). Additionally, some longitudinal correlational studies indicate that spelling itself contributes to early literacy (Cataldo & Ellis, 1988; Mommers 1987; Morris & Perney, 1984).

Learning to spell not only increases children's reading but also their writing (Moats, 2005). Graham and Hebert (2011) found that spelling instruction improves word reading skills in first-grade to fifth-grade with a moderate effect size of 0.68. In their meta-analysis, they outlined how students' reading skills benefited from writing

instruction in many different ways (Graham & Hebert, 2011). The findings of their meta-analysis support the previous claim that learning to spell supports reading (Graham, 2000; Moats, 2005). Moreover, Graham and Hebert also found that teaching spelling and sentence construction skills improves reading fluency with a robust moderate effect size of 0.79. Providing instruction on how words are spelled gives students an organizational pattern of the connections between letters and sounds; this pattern helps students identify and remember word connections more easily (Ehri, 1987; Moats, 2005). Spelling is also important for reading because "knowing the spelling of a word makes the representation of it sturdy and accessible for fluent reading" (Snow et al., 2005, p. 86). Ellis and Cataldo (1990) found that information from spelling and reading is unidirectional, meaning that knowledge in spelling contributes to reading.

Conrad (2008) offers evidence for how reading and spelling are related. More specifically, she investigated the transfer from reading to spelling and spelling to reading which previous studies had not yet explored (p. 970). Conrad hypothesized that spelling and reading had reciprocal benefits but that practicing spelling would benefit reading more than practicing reading would benefit spelling (p. 870). Forty-one students in second-grade were randomly assigned to either repeated practice spelling or repeated practice reading the same 40 words. Then, to address word-specific transfer, students either read or spelled the words they had just practiced. Students in the repeated practice spelling group were asked to read the same 40 words, and students in the repeated practice reading group were asked to spell the same 40 words. Conrad found that transfer of skills for reading to spelling and spelling to reading does go both ways. There was complete transfer of skills from spelling to reading the words; however, there was a less

complete transfer from reading to spelling. Therefore, practicing reading a word does not mean that a student will be able to spell a word correctly. Perhaps this can be attributed to the fact that spelling requires additional knowledge because the English language has many letters (graphemes) which make up the 44 sounds (phonemes). Spelling is a more complicated task than reading because it requires retrieving the letters that make up the sounds instead of just being able to recognize the letters (Perfetti, 1997; Zutell, 1992). Reading requires only recognition of a word, whereas spelling requires a complete reproduction.

Theoretical Approaches to Spelling Development

Spelling development can best be understood through the lens of two theoretical perspectives: developmental stage theories and the strategy approach. The first of the two theoretical perspectives known as developmental stage theories (Ehri, 1986; Frith, 1980; Gentry, 1982; Henderson, 1985), includes four predominate theories of spelling development (Ehri, 1986; Frith, 1980; Gentry, 1982; Henderson, 1985). Stage theories are based on the underlying concept that children go through a number of stages as they develop, and each stage is a prerequisite for the next stage. Even though most theorists have assigned a typical age for children in each stage, the stages are based on the student's experiences, not necessarily their current age or grade level. This focus on the student's experiences allows for the student's abilities and needs to be taken into account instead of their age or grade level; therefore consideration of theories of spelling development can provide an argument for how to assess spelling.

Stage Theories

Ehri. Ehri (1986) originally began her theory of spelling development with three stages: *semiphonetic*, *phonetic*, and *morphemic*. However, she added an additional stage,

the *precommunicative stage* (Ehri, 1989). In the precommunicative stage children produce spellings with strings of letters, scribbling, and even numbers sometimes to represent sentences or words (Ehri, 1989, p. 358). In the second stage, semiphonetic, children are introduced to the sounds that letters make and children use this knowledge to spell (Ehri, 1989). Although, in the child's actual spelling, only one or two of the letters actually correspond correctly to the sounds the letters make. For example a student in the semiphonetic stage might spell the word "tickle" as "tl." In the third stage, phonetic, children produce spellings which demonstrate a more complete awareness of sounds and letters that match. The largest difference between the semiphonetic stage and the phonetic stage is the child's ability to be able to spell vowels. Also, children start to acquire an understanding of rules for spelling in the phonetic stage as well. Children in the phonetic stage often believe that one sound should represent one letter. For example, in this stage, "doctor" is spelled "doktdr" (Ehri, 1989). In the final stage, morphemic or transitional, children advance their spelling abilities by acquiring knowledge of both orthography and morphology. Orthography is the writing system of a language, including the spelling and specific spelling rules, punctuation, and capitalization rules. Morphology refers directly to how words are formed with morphemes, the smallest units of meaning. Children in the morphemic or transitional stage are also able to recognize spelling patterns that occur in words and understand that spellings for sounds can include more than one letter. For example, the word "runner" might be spelled "runnur." As a child progresses through each stage and gains more experiences they are able to "…spell more sounds in words" (Ehri, 1989, p. 358).

Frith. Frith (1980) also proposed a stage theory that was less detailed than Ehri's. Her theory of spelling development is of great importance because Frith's stages take into account the spelling errors students make in spelling nonsense words, in addition to spelling real words. Looking at the errors made by students in spelling real words and nonsense words, Frith (1980) proposed three stages of spelling development. In the first stage, students are able to guess the smallest units of sound. In the second stage, the smallest units of sound are converted into the correct letters. For example, a student would be able to recognize that the sound of /t/ in tee is represented by the letter t. In the third stage, the student must to be able to differentiate the correct grapheme out of all the graphemes available for selection. For example, the phoneme /k/ can be spelled in grapheme form <c>, <ck>, <k>, <qu>, <ch>. Or if a child hears the word "clock" he or she has to decide what letter or letters those sounds represent.

Gentry. Gentry (1982), presented five stages of spelling development which are similar but different in description from Ehri's stages. The five stages outlined in Gentry's (1982) theory of spelling development are *precommunicative, semiphonetic, phonetic, transitional,* and *correct*. In the first stage, the precommunicative stage, children use letters of the alphabet, however they do not have foundational knowledge of letter-sound correspondence. Examination of a child's writing in this stage would reveal a string of random letters combined together. The child also may not know all the letters of the alphabet and is not able to make a distinction between uppercase and lowercase letters. In the second stage, the semiphonetic stage, children start to understand letter-sound correspondences and try to use letters to match the sounds in a word. Consonants and vowels are often represented by one letter, for example using the letter "u" to

represent the word “you.” In the third stage, the phonetic stage, children spell by representing the sounds; however, children do not adhere to conventional English spelling rules. For example, the word “take” is spelled as “tak” because the “e” is silent and the child does not hear it. In the fourth stage, the traditional stage, a child takes into account how words look as well as how they sound. The child begins to take into consideration the visual representation and structure of words. A child in this stage also starts to realize that some words have silent sounds such as the word “have”, which in the previous stage would have been spelled “hav” because the child was unaware of silent sounds such as this “e.” Also, children may use the correct letters but put them in the wrong order; for instance, the word house is spelled “huose. Finally, in the fifth stage, the correct stage, the child has a grasp of basic spelling rules. Also, children are aware of alternative spellings, silent consonants, and irregular spellings. Even though a child may move to a new spelling stage, characteristics from a former stage may exist in a particular writing sample (Gentry, 1982).

Henderson. Henderson’s (1985) stages of spelling development, like Gentry (1982), also included five stages: *preliterate*, *letter-name spelling*, *within-word pattern*, *syllable juncture*, and *derivational principles*. In the preliterate stage, children are making their first attempts at writing; these marks include drawings, letter symbols, and scribbles. Children’s marks are all over the page, and they do not adhere to left to right writing. In the letter-name alphabet stage, children are beginning to use invented spelling and recognizing that sounds represent letters. As children write, they create their own spellings for words based on their current knowledge of phonology, which is known as “invented spelling” (Read, 1975). Contrary to popular belief, invented spelling does not

conflict with correct spelling rules and plays a vital part as children learn to write (Burns, Griffin, & Snow, 1999). Invented spelling is also known as phonetic spelling, or spelling things the way they sound. In the next stage, within word pattern, the child still uses invented spelling but spells long and short vowels properly. In this stage, sight words are often spelled correctly and used to help the child spell words that are less familiar. In the stage, syllable juncture, the child starts to apply and gain spelling rules, for example doubling consonant sounds. In the last stage, derivational principles, children understand morphology (the relationship between spelling and its meaning). The child also understands the root word and the meaning it carries and what happens in spelling when a prefix or suffix is added.

Strategy Approach

The second theoretical perspective that can be applied to spelling development is the strategy approach. The difference between the two perspectives, stage theories and the strategy approach, is that the strategy approach characterizes spelling development as continuous and ongoing. In contrast to stage theories, which describe spelling development as a more sequential process passing through different stages one at a time throughout learning to spell. Gentry (1982) did, however, mention that even though students may proceed to later stages of spelling development they can still include characteristics of earlier stages in their writings. Since the strategy approach is a more continual and evolving developmental process, children incorporate an array of spelling strategies from the very time they start learning to spell (Treiman, 1998). Although stage theories help explain spelling development, they "…do not fully capture the complexities of phonological, orthographic, and morphological representations as they relate to spelling" (Treiman & Bourassa, 2000, p. 2). Therefore, the strategy approach in

combination with stage theories can provide an overall picture of students' spelling development.

The Differences in Reading and Spelling Achievement of Males and Females, ELLs, and Students in SPED

The differences in academic performance between varying groups of students is often referred to as the achievement gap. In hopes of closing the achievement gap the federal government enacted the No Child Left Behind Act (NCLB) (Chamberlain, 2004). NCLB mandated that all children should be reading by third-grade by the 2013-2014 school year.

The majority of students in the United States still have not acquired the reading skills necessary to succeed in society; in fact, more than half of all fourth-grade students (66%) and 80% of all low-income fourth-grade students in the United States still cannot read proficiently at grade level, according to the *National Assessment of Educational Progress* (NAEP, 2013). In addition, there are discouraging statistics for achievement gaps regarding gender, ELLs, and SPED. Since the first administration of the NAEP *Reading*, for fourth-graders, female students have scored higher than male students in reading achievement. However, the score difference between the female average and the male average for reading was only 9 points at the 4^{th} grade level, in comparison, the 8^{th} grade scores differed by 10 points, indicating that the gap between females and males' reading scores widens over time (NAEP). Moreover, per the same assessment, students with disabilities had an average score of 189 points (considered below basic level) and students without disabilities had an average score of 227 (above basic level but below proficient level), a difference of 38 points (NAEP). Lastly, the largest achievement gap

was between Non-ELLs and ELLs. In fourth-grade, Non-ELLs on average scored 226 points while ELL students scored 187 (below basic level), a difference of 39 points (NAEP). Even though efforts have been made to increase the reading performance of all students in this country, the data provided from NAEP indicated that the achievement gap for some groups still continues and that the NCLB mandate was not successfully met by the 2013-2014 school year.

A search for current research on spelling achievement being the outcome elicits very few studies for gender differences. Typically, some researchers have simultaneously investigated the discrepancies of achievement scores within gender with regard to writing disabilities, which includes spelling, and with regard to reading disabilities (e.g., Berninger et al., 2008). Additionally, only one team of researchers has investigated writing disabilities separately from reading disabilities.

Allred (1990) conducted a comprehensive study in the United States that examined the gender differences of students in first-grade to sixth-grade in spelling achievement. Allred reported that girls were able to correctly spell considerably more words than boys on both a standardized proofreading test and a researcher-constructed spelling test with the same words (p. 190). The difference between girls and boys' spelling was significant as determinged by analyses of variance for both spelling tests at $p < .001$ with the girls scoring significantly higher than the boys. Due to the fact that this was the first comprehensive study to address gender differences in spelling achievement, Allred hypothesized that these gender differences existed because of cultural expectations, earlier maturation of girls, teacher expectations, and student sex-role expectations. As such, further research is needed to determine the existence of gender

differences in spelling abilitities despite the complexities associated with gender differences.

More recently, Foorman and Petscher (2010) investigated the usefulness of spelling assessment in grades three to twelve. Similar to Allred's (1990) findings, Foorman and Petscher reported that females outperform males on spelling. Additionally, at the student level from multi-level modeling analyses, Foorman and Petscher found that boys performed worse than girls at the beginning of the year on spelling except in 12th grade. However, ELLs and students in SPED performed poorly on spelling throughout the year. Both groups of students performed about half a standard deviation below average performing students (p. 14).

Berninger et al. (2008), investigated gender differences of writing disabilities and investigated reading disabilities at the same time. More specifically, these researcheers explored the gender differences in writing and reading disabilities of children with dyslexia, boys and girls, as well as their parents, both biological parents, and two additional nuclear or extended family members (grandparents, aunts, uncles, and cousins). The researchers found that gender differences were not present in reading, they were only present in writing disbailities but not in reading. Berninger and collegues reported that adult males scored significantly lower on average that adult females on two measures of spelling. In contrast, no significant differences were found between the boys and girls with dyslexia for spelling.

Katusic et al. (2005) is the only group of researchers to examine gender differences in reading disbailities as well as writing disabilities in separate studies (Katusic et al., 2001). More often, gender differences in writing and reading are explored

within the same study, which may lead to limitations in results or findings. Katusic et al. (2005) reported that men were more likely to have a writing disability than women. The ratio of the probability for men compared to women varied from 2.0 to 2.9 (Katusic et al., 2005). Katusic et al. (2001) also reported that men were more likely to have a reading disability than women.

Spelling Assessment

Unlike reading, no national tests exist to evaluate students' spelling abilities. Therefore, one could reason that there is less interest and attention placed on assessing students' spelling. However, the evidence that spelling and reading are mutually related to one another and even show bidirectional transfer between skills is unavoidable. Ignoring spelling and not helping students improve their spelling, "...can lead to decreased academic, vocational, and social success" (Masterson & Apel, 2010, p. 185). Assessing a student's spelling is the first place to begin when making informed decisions about creating interventions to enhance spelling performance.

The methods for spelling assessment are fairly standard in United States classrooms. A national survey conducted by Graham et al. (2008) found that 90% of first-grade to third-grade teachers reported using a "Friday test" for teaching and assessing spelling. Most of the teachers picked words from an array of sources without any consideration of the developmental characteristics of the students. Words were also not individualized at the student level as the entire class was given the same set of words to spell each week. These tests were scored using traditional scoring where the word is scored as either correct or incorrect. By assessing students spelling as either correct or incorrect the progess students makes from one time point to another is diffcuty to judge except that they are either able to spell a word in its entitreity or not at all. With a careful

analysis of spelling errors and use of an alternatives away to score spelling a more accurate representation of a student's current abilities is possible. Then successful instructional plans and interventions can be developed. In fact, Otaiba and Hosp (2010) described how it is important to have detailed spelling assessments to inform instruction.

Masterson and Apel (2010) conducted a case study of spelling assessment for one student using prescriptive spelling assessment, which identifies specific student deficits by evaluating their misspellings. Not only were words chosen for the student to spell based on an appropriate developmental level, but also the words were scored for each correct sound. This type of analysis allows the teacher to identify spelling patterns the student has yet to master and leads to the realization that alternative procedures for scoring spelling may be more diagnostically accurate. Detailed error analysis provides a wealth of data when done systematically and different types of error analysis can be completed depending on the purpose (Otaiba & Hosp, 2010).

Curriculum-Based Measurement Scoring for Spelling

Alternative methods for documenting students' abilities in spelling are needed because of the complexity of spelling. Also the rationale for using an alternative method to score spelling is based on the notion that spelling is developmental and progresses in stages/phases; therefore, an assessment which can document these different changes is necessary to document the student's growth. Ehri (1989) said it best, "…if spelling ability is measured only by the number of correct words, this development is totally missed" (p. 357). One of these alternative methods for scoring spelling is referred to as Curriculum-Based Measurement (CBM) scoring for spelling. CBM was developed over 30 years ago by Stanley Deno at the University of Minnesota and is supported by more than 30 years of school-based research (Deno, Mirkin, & Marston, 1982). CBM is a method of

monitoring student progress with direct assessment of academic skills such as reading, math, spelling, and writing expression. Students can be tested weekly or daily with CBM probes because administration takes usually one minute to five minutes at most.

Ritchey (2008) assessed kindergarten spelling responses for real words and nonsense words with CBM scoring. Scoring with CBM allowed Ritchey (2008) to document that students who could not spell an entire word correctly could at least spell some words partially. Abu-Hamour, Urso, and Mather (2013) conducted a case study investigating the value of normative-based assessments and CBM with one student, Miguel, after receiving a 24-week intensive intervention. The authors found that while each assessment provided beneficial data, each provided different kinds of information. The standardized assessment helped the researchers to identify Miguel's reading disability while the CBM probes provided Miguel's current level of performance as well as Miguel's ongoing performance throughout the intervention. Abu-Hamour et al. also noted that CBM is a more sensitive measure of growth.

Additionally, L.S. Fuchs, Fuchs, Hamlett, and Allinder (1991) conducted a study assessing the contribution of ongoing skills analysis to CBM. They found that teachers who used skill analysis were better able to plan for instruction and in turn increase students' achievement. Overall, it is evident that CBM is a beneficial alternative method for scoring spelling that provides additional information beyond examining correct or incorrect responses.

Computer-Based Interventions for Promoting Reading and Spelling Growth

The complexity of spelling and the interrelationship between writing and reading lay the foundation for alternative methods of instruction which incorporate the instruction

of reading, spelling, and writing. Although desirable, one-to-one instruction with students may not be possible due to the lack of trained teachers and lack of funding to acquire additional teachers. Computer-based reading interventions offer plausible solutions to increase reading, spelling, and writing achievement. Many variations of computer-based programs have been developed to increase performance in reading and spelling (Comaskey et al., 2009; Di Stasio et al., 2012; Fasting & Lyster, 2005; Howell et al., 2000; Savage et al., 2010; Van Daal & Reitsma, 2000).

Leescircus

Van Daal and Reitsma (2000) investigated the effects of *Leescircus* in two pilot studies. Leescircus is an instructional program for beginning readers with a variety of engaging activities, feedback, and supplemental instruction when needed. In the first study Van Daal and colleagues examined how kindergarten students could acquire reading and spelling skills with computer-assisted practice. The researchers selected 21 children from two kindergarten classrooms in the Netherlands for this first study, and they randomly assigned children to either the experimental or control group. Nine students were assigned to the computer-based reading program, the experimental group, and 13 students were assigned to the control group. At the end of the study, all students were assessed on three post-test measures: a standardized Dutch test for word recognition (Brus & Voeten, 1973), naming letters, and a standardized Dutch word reading test for the measurement of decoding skills (Van den Bos, Scheepstra, & Spelberg, 1993).

On average, all nine students in the experimental group spent three hours and thirteen minutes on the computer. However, the time spent on the computer-based reading program was highly variable. For example, one student spent five hours and 43

minutes on the computer while another student spent just 94 minutes on the computer. Results from pre- to post-test scores for the experimental group showed a significant increase in letter knowledge but not in concepts related to reading and writing. Additionally, at post-test, the experimental group performed better on real word reading and non-word reading than the control group. Overall, the experimental group showed improvements on letter knowledge and reading real words and nonwords than the control group after using the computer-based reading program, despite the variability in time spent on the computer. (Note: Though the study referenced 13 students in the control group and 9 students in the expeirmal group this number totals 22 students. However, the study only referenced 21 students total in the study.)

The second study conducted by Van Daal and Reitsma (2000) investigated spelling tasks that were computer-based from Leescircus and matched to the student's ability level to promote success. Fourteen students (three girls and 11 boys) identified as having low levels of motivation and being rated by their teachers for having poor spelling abilities participated in this study. All students in the study had learning disabilities and were from many different classes in a special education school. All 14 students worked on the computer-assisted spelling tasks provided by Leescircus and spent at least five minutes a day, at least three times a week, for approximately a half a year, on the computer. A Dutch standardized dictation test (Geelhoed & Reitsma, 1999) with blocks of 15 words read aloud in a sentence to clarify meaning, as well as an adapted version of the Coping Analysis Schedule for Educational Settings (CASES) (Spaulding & Papageorgiou, 1999) were used. This version of CASES was used to monitor learning behavior, i.e. on task and off-task behavior. In the spelling dictation test, before using the

computer-assisted program students, mean total number of words correctly spelled was 58.2; after using the computer-assisted program the student average was 74.3. The improvement after using the computer-assisted program was significant. Another encouraging result was an increase in positive learning behaviors during the use of the Leescircus program.

MultiFunk

Fasting and Lyster (2005) examined the effects of MultiFunk, a computer program to assist struggling readers with reading in Norwegian, on the reading and spelling proficiency of students. Firty-two children in grades five to seven who, per teacher report, were struggling in reading and spelling were randomy assigned to either an experiemental or control group, each with 26 students. The first language of all participants was Norwegian. Beyond teacher report the researchers examined reading and spelling skills assessments and found these selected students were indeed performing below their classmates. During the use of the program, 106 classmates who were considered normal readers and spellers were included for comparisons between changes in reading and spelling, MultiFunk. Fasting and Lyster employed an experimental *pre-test-intervention-post-test design*. One spelling test composed of 32 words and two reading tests (i.e., a single word decoding test and a sentence reading test) were given at the beginning of the study, as well as 11 weeks after the intervention for follow-up. Students enrolleld in the experimental group used the program for at least 20 minutes every day for seven weeks and on average, for a total of 26 out of 33 days because of various school and classroom interruptions. To make up for the lost time teachers allowed students to work more than 20 minutes on certain days. The expiermntal group

aprticpants increased their reading and spelling skills significantly in comparison to their classmates in the control group. Furthermore, boys in the experimental group demonstrated a significant gain in spelling. The results of this study justify the need to use computer-based technologies, such as MultiFunk, to combat the problem of low achievement in reading and spelling in particularly with regard to narrowing the achievement gap between boys and girls in spelling.

ABRACADABRA

Three studies have been published to date investigating the effects of *A Balanced Reading Approach for Canadians Designed to Achieve Best Results for All* (ABRACADABRA; also known as ABRA) on measures of reading and spelling (Comaskey et al., 2009; Di Stasio et al., 2012; Savage et al., 2010;). ABRA is a free web-based literacy intervention created by the Center for the Study of Learning and Performance out of Concordia University in Montreal, Canada.

Comaskey et al. (2009) investigated the effects of a web-based literacy intervention (ABRA) in conjunction with either a synthetic phonics or an analytic phonics approach on the early phonological and reading skills of 53 kindergarten children in a randomized control trial. Children were randomly assigned into either the synthetic phonics approach or the analytic phonics approach. All students were given pre- and post-tests consisting of the following eight measurements: letter-sound knowledge, Wide Range Achievement Test Word Recognition Subtest (WRAT) (Wilkinson, 1993), segmenting, blending, common unit articulation tasks (i.e., one required students to articulate a common rime between two words and another required students to identify common coda units in to words such as the “t” in pat and cat), nonsense word

recognition, and unique word recognition (i.e., students were expected to read a list of eight words containing no similar rimes between words). Students were also given the Peabody Picture Vocabulary Test III (PPVT) at pre-test only to assess general receptive verbal ability (Dunn & Dunn, 1996).

No significant effects were reported for letter-sound knowledge, WRAT reading. Significant results were reported for consonant-vowel (CV) and vowel-consonant (VC) blending for both groups but the synthetic phonics group showed more of an increase on blending than the analytic phonics group. However, CV and VC segmentation had no significant effects for either group. There was a significant effect for common unit rime articulation in favor of the synthetic phonics group which suggests that students who are taught explicitly about rimes will be able to better identify and articulate the specific rime. Additionally, no significant effects were reported for word recognition, both real word reading and nonsense word reading.

Savage and colleagues (2010) used a quasi-experimental design with four first-grade classrooms. Three classrooms of which received the ABRA intervention and one classroom served as a control. The three classrooms which received the ABRA intervention were classified as (a) *entry level* (teacher one), (b) *adoption level* (teacher two), and (c) *adaptation level* (teacher three). Teacher one in the entry level classroom was anxious about technology, had difficulty running the intervention, and gave no directions on how to use ABRA. This teacher also taught the control group but did not use any technology. Teacher two in the adoption level classroom typically had the whole-class read an ABRA story on a big screen and complete a letter-sound or alphabetic activity, and then had students work on their own or in pairs using ABRA. Observers

who watched teacher two's classroom did not observe the teacher connecting ABRA to any non-technology based learning themes. Teacher three in the adaptation level classroom integrated ABRA into the classroom by making it a center activity and connected ABRA activities into other non-technology literacy components.

Savage et al. (2010) pre-tested and post-tested all students in all four classrooms with an array of measures for reading, spelling, and arithmetic. They reported a significant main effect for word reading, word meaning, combined vocabulary task, and for elision task. Although there were no significant effects reported for the other measures, there was an increase from pre-test to post-test measures for all groups for all measures. Overall, the adaptiontion group performed better than all groups except for the adoption group on elision and listening comprehension but their pre-test score average also started higher than the adaption group at 102.19 vs. 98.56 at pre-test average for elision and 5.50 vs. 4.17 at pre-test average for listening comprehension.

Di Stasio et al. (2012) conducted a one year follow-up study in 2012 from the randomized control trial study of ABRA that investigated a synthetic phonics approach in contrast to an analytic phonics approach (Comaskey et al., 2009). Of the original 53 participants from Comaskey et al.'s (2009) study, 49 were available at follow-up. A significant main effect was reported for only reading comprehension for the analytic group. All other measurements at follow-up did not yield any significant effects in word decoding, spelling, phonological skills, blending, and WRAT Math.

Although the above mentioned studies for ABRA, provide evidence to support web-based reading intervention, critics feel these types of results can only be accomplished in a research setting. For example, Harper et al. (2012) discussed the

amount of support schools normally receive to implement an intervention outside of a research setting and argued that the previous studies of ABRACADABRA were conducted under optimal conditions where key variables could be controlled and research personnel or teachers were available for closely checking the fidelity of the intervention. Harper et al.'s study attempted to address how well a technology like ABRA works in a school setting without any researcher supports. Six schools implemented this computer-based program and teachers reported that they did not know how to connect ABRA to the elements of literacy that they were teaching within their classroom. Teachers also felt that one day of training, which is what teachers would normally receive in this area if a new intervention were implemented was inadequate and that the teachers in general did not feel appropriately prepared to teach reading in connection with this program. Comaskey et al., 2009, Di Stasio et al., 2012, and Savage et al., 2010, provided strong cases for using ABRA to increase reading and spelling skills but they also highlighted the importance of recognizing the obstacle of obtaining similar results outside of a research setting.

IntelliTools

Howell et al. (2000) investigated the effects of *IntelliTools Reading* on the reading performance of first-grade students identified with disabilities or at risk for reading failure. Intellitools is a computer-based prototype of an early reading program that requires software via a CD-ROM to run the program developed by IntelliTools Inc. In Howell et al.'s students were placed in two different groups: one group used IntelliTools Reading (experimental group), and one group served as the control group. All students were tested with a variety of instruments to address five areas of skills that

children who read at the beginning to middle first-grade level would be able to complete after using the program. The five subtests covered the following: onsets, rimes, phonemic awareness, spelling words (i.e, words scored in two ways, by total words written and by a three-point scale of developmental spelling), and word identification. The authors reported significant effects for onsets, rimes, phonemic awareness, and word identification. Spelling words scored by total words written and by the three-point scale of developmental scoring were not found to be significant. However, the experimental group's trend reflected a higher perfomance than the control group. Finally, Howell et al.'s results offer one more piece of evidence on why a computer-based technologies, such as IntelliTools, may increase students' reading and spelling skills and also the need to score spelling with an alternative method (i.e., CBM-spelling).

CHAPTER 3: METHOD

The purpose of this chapter is to discuss the methods used to answer the following research questions:

Research Questions

The following questions were addressed:

1. Does MVRC have a significant effect on students' spelling scores from the WJ IV ACH for Test 3: Spelling, and Test 16: Spelling of Sounds over time?
2. Do spelling achievement trajectories vary by gender, ELL classification, and/or SPED classification for the WJ IV ACH, Test 3: Spelling, and Test 16: Spelling of Sounds?
3. Which scoring measurement is more sensitive to student growth: traditional standardized scoring (correct or incorrect) or CBM spelling scoring (which credits each correct letter sequence) from the WJ IV ACH Test 3: Spelling, and Test 16: Spelling of Sounds?

Data Source

All data collection activities were supervised and coordinated by a third-year doctoral student in Special Education from the University of Arizona in the Department of Disability and Psychoeducational Studies. The data were originally collected for a larger study, and the author used a limited data set containing the following: (a) students' written responses from the *Woodcock-Johnson® IV Tests of Achievement* (WJ IV ACH) for the tests Test 3: Spelling (standard battery) and Test 16: Spelling of Sounds (extended battery) for pre-test and post-test, (b) students' gender (male or female), (c) ELL status (ELL, RFEP, or Non-ELL), (d) SPED status (SPED or Non-SPED), (e) whether the student was assigned to the experimental group or the comparison group, and (f) what teacher each participant had for 2013-2014 school year. The limited data set was de-

identified before the author received it and no names were provided to the author. A unique identification number was assigned to each student before the author received the limited data set. Additionally, each classroom was given a unique identification number (1-8). The information in the limited data set and accompanying documents supplied to the author cannot be combined to identify an individual or be readily ascertained by the author.

Research Design

The researcher compared the spelling performance of students in two different schools for eight different classrooms. Because the research questions examined the effect of the intervention on students' spelling and the growth of students' spelling, a quantitative approach was best (Creswell, 2009). More specifically, this study employed a quasi-experimental design with an additional qualitative component for analyzing the difference in the growth (i.e., percentage gain score) for the two different scoring techniques. Random assignment was not possible in this educational setting because it would have been very difficult to rearrange children into particular classes (Salkind, 2012). Therefore, each classroom served as a sample and the best design was a nonequivalent control group design, one of the most commonly used quasi-experimental designs. The one variable the researcher did not have control over in this design was is if the groups were equal. The issue of unequal groups was addressed by giving a pre-test to all of the students in the control and experimental groups.

Setting and Participants

This study was conducted during the 2013-2014 school year in a PK-12 public urban school district, Misty Mountain School District (pseudonym). Misty Mountain

School District, located in the southwest, serves a population of over 17,000 students with 22 schools (i.e., early childhood education centers, elementary schools, middle schools, high schools, and alternative education schools).

The specific ethnic make-up of the students in the district included Hispanic (81.6%), Native American (4.9%), African American (2.7%), Asian/Pacific Islander (0.6%), and White (12.1%). At the time of the study, about 77% of the district's students were eligible for free and reduced meals and about 6% were homeless. Nine percent of the students were classified as ELLS and 5.3% had been identified as recently reclassified ELL. A recently reclassified ELL (RFEP) was a student who had previously been considered an ELL but instead had achieved a satisfactory level of English-language proficiency per standardized testing. In addition, about 12% of the students' district-wide had an Individualized Education Program (IEP). School specific demographic data were not available to the researcher; this allowed schools to keep their anonymity in the district and ensured that the researcher would be unable identify the participants. Also specific ethnicity and socio-economic status were not provided at the student level to protect against confidentially and to allow the specific students to have anonymity. The only decriptive information available to the researcher at the student level included gender, ELL classification, and SPED classification.

Participants for this study were 209 second-grade students enrolled in Misty Mountain School District. At the beginning of the study, 102 students were assigned to the experimental group and 107 were assigned to the business-as-usual comparison group referred to as the comparison group. However, 31 participants, 13 from the experimental group and 18 from the comparison group, left the study prior to post-test collection.

These 31 students left the study because they either left the district or moved to a different school not participating in the study. Accounting for experimental mortality, 178 participants remained. Moreover, 19 additional students were either missing demographic information or did not have appropriate ceilings for the WJ IV ACH two subtests of spelling and were also excluded from this study. The purpose of basals (starting points) and ceilings (ending points) in the WJ IV ACH is to limit the span of items administrated but to be able to estimate the score the student would have obtained had they taken the entire test. Students were still getting the right answers but an adequate ceiling of six incorrect words had not been reached. Overall, the experimental condition contained 83 participants and the comparison condition contained 76 participants, a total of 159 participants.

Group Assignment

Random assignment was not possible at the student level due to ethical concerns of sorting students into particular groups and disrupting pre-existing routines of each already formed classroom. Therefore, students were assigned to groups according to the school and the classroom/teacher to which they were already assigned. Students were enrolled in two elementary schools within Misty Mountain School District; one elementary school served as the comparison school while the other served as an experimental school. In total, eight classrooms were a part of this study: four classrooms from the comparison school and four classrooms from the experimental school. Table 1 includes details of the frequency count of the students within each of the eight classrooms. Teachers one through four were within the comparison school/group and teachers four through eight were within the experimental school/group.

Table 1

Frequency Counts by Teacher

	Frequency of Students	Percent of Students (rounded)
Teacher 1	16	10.1
Teacher 2	17	10.8
Teacher 3	23	14.6
Teacher 4	20	12.5
Teacher 5	17	10.8
Teacher 6	18	11.4
Teacher 7	26	16.5
Teacher 8	22	13.9
Total	159	100.0

By examining each school's test scores for their third-grade students in reading, it was determined which school would serve as the experimental school and which would serve as the comparison school. Even though this study only included second-graders, the youngest grade tested the prior scholastic year, 2012-2013, with the state standardized test was third-grade. Therefore, there was no testing data available for second-graders. Group assignment was organized in this way to ensure that any demographic advantages would favor the comparison group. The school with a much lower pass rate (39%) on the reading section of the state standardized test for third-grade students was assigned to the experimental group, while the school with the higher pass rate (53%) was assigned to the comparison group.

Description of Measures: Woodcock-Johnson IV Achievement

The Woodcock-Johnson® IV Tests of Achievement (WJ IV ACH) is a standardized and reliable assessment tool for children and adults that contains various subtests designed to measure major areas of academic achievement for reading, writing and mathematics (Schrank, Mather, & McGrew, 2014). The WJ IV ACH Test 3: Spelling

(standard battery) and Test 16: Spelling of Sounds (extended battery), were used to measure the dependent variable, spelling scores (Schrank et al.). The WJ IV Test 3: Spelling is the spelling of real words, while Test 16: Spelling of Sounds is the spelling of nonsense words, words that are not real but follow regular spelling patterns (Schrank et al.).

Curriculum-Based Measurement Spelling Scoring for the Qualitative Analysis

Spelling CBM can be administered in a group or individually. When scoring CBM spelling probes, student responses are scored based on correct letter sequences (CLS). For each CLS that is correct in the word, the student receives a point. With CLS, a phantom letter exists at the beginning of the word (_cat) and at the end of the word (cat_). To find the correct letter sequences in a word, the total number of letters in a word are added, and then one more phantom letter is added. When scoring a student's spelling using CBM, carats represent correct letter sequences (^). For example, the word cat (^c^a^t^) has four CLS. If a student spells the word *cat* as *cath* the student would receive three CLS (^c^a^th).

Dependent Variable

All students were administered two spelling measures at pre- and post-test: WJ ACH IV Spelling and Spelling of Sounds. The dependent variable was spelling achievement. Students' spelling responses for real word spelling and non-word spelling for the experimental and comparison condition were scored in two different ways: tests that were scored with traditional scoring (correct and incorrect), represented as a point per correct word, and the same tests scored with Curriculum-Based measurement scoring (correct letter sequences), represented as a total number of correct letter sequences.

Independent Variable

The independent variable was the intervention, MVRC. Those students assigned to the experimental group received Mind Play Virtual Reading and Success for All language arts instruction, and those assigned to the comparison group received Success for All language arts instruction and two hours of regular reading instruction provided by their teacher.

Intervention description. MVRC is an internet-based reading intervention program recommended for students and adults with varying proficiency levels of reading. MVRC has been appropriate and beneficial for a wide range of individuals, including: English Language Learners, accelerated learners, teen and adult learners, and struggling readers. MVRC provides an individualized learning plan for each student based on comprehensive diagnostic reading assessments. During the use of MVRC, each student watches videos and receives one-to-one instruction from virtual instructors, real-life reading specialists, and speech pathologists. MVRC follows many of the principles of an Orton-Gillingham (OG) approach, an approach designed for people who have difficulty with writing, reading, and spelling. Ritchey and Goeke (2006) explain, "The OG approach is a systematic, sequential, multisensory, synthetic and phonics-based approach to teaching reading" (p. 171). MVRC also provides direct instruction for the five essential components of a reading program identified by the National Reading Panel (2000): phonemic awareness, phonics, vocabulary, fluency, and comprehension. Finally, MVRC includes meaning and grammar instruction for connecting the understanding between reading and writing.

Student Predictors

Additional variables included the student predictors of ELL classification (ELL or Non-ELL, and RFEP), gender (male or female), and SPED classification (SPED or Non-SPED). The frequency counts for all of student predictor data can be found in Tables 2, 3, and 4 provided below.

Table 2

Frequency Counts of Nominal Data for Gender

	Frequency	Percent of Students (rounded)
Male	89	55.9
Female	70	44.1
Total	159	100.0

Table 3

Frequency Counts of Nominal Data for Special Education

	Frequency	Percent of Students (rounded)
Non-SPED	139	87.4
SPED	20	12.7
Total	159	100.0

Table 4

Frequency Counts of Nominal Data for English Language Learners

	Frequency	Percent of Students (rounded)
Non-ELL	82	51.9
ELL	63	39.9
RFEP	14	8.8
Total	159	100.0

Description of Classroom Instruction

All participants in both the experimental and comparison condition received Success for All (SFA) language arts instruction in the regular classroom setting. SFA was

developed by Robert Slavin and colleagues, and it was founded on the belief that all students can succeed and need to succeed in the early grades (Slavin, Madden, Dolan, & Wasik, 1996). SFA is an intervention that schools can purchase; it provides an all-inclusive package with training, materials, ongoing professional development, and a plan for delivery and continuing the model (Borman et al., 2007). Schools that decide to adopt SFA implement the program for the entire school from Pre-Kindergarten to 5th grade. The main goal of SFA is to organize school resources to ensure that each student will have basic language arts skills by third-grade and will continue to build these competiencies throughout elementary school (Borman et al.). SFA activities include 90-minute blocks of daily reading which are highly intensive and also include guided practice, phonics instruction, and ongoing formative evaluation (Slavin & Madden, 2001).

Intervention and Comparison Procedures

Experimental Setting

Prior to the beginning of the study, MindPlay Educational Software for Reading personnel visited each elementary school and computer lab and confirmed the appropriateness of each school's computer hardware, broadband capabilities, and internet connectivity. Students in the experimental group received MVRC online reading instruction for the same period of time in the computer lab of their respective school. Students not in the experimental group were not permitted in the computer lab when students were working on MVRC. The environment in the computer lab was made as distraction-free as possible. Each student was provided with an individual computer, monitor, and headphones.

Teacher Training

Prior to the implementation of MVRC, teachers received training from MindPlay personnel. Each teacher completed a one-hour webinar and a subsequent in person four-hour training. Little teacher participation was needed during the intervention because MVRC provides adaptive instruction to students in response to their ongoing performance as they use the program. Because little teacher participation was needed, the trainings mainly focused on the structure of the program and the underlying foundations by which the program was designed, the description of the process of the lessons, the details of how to facilitate and support the use of the intervention, and the directions for accessing and understanding the data generated by the software in the student and classroom reports. In addition, teachers were provided with a guide for MVRC.

Procedures: Experimental Instruction

Classrooms assigned to the experimental group received MVRC online reading instruction in addition to SFA reading and language arts instruction. Students used MVRC for 30 minutes per day, Monday through Thursday, for a total of two hours per week. The students' use of MVRC occurred during mid-September to mid-April with the exceptions of holidays, school in-service days, school functions, district testing days, and mandatory state testing days. Classroom teachers brought their classes to the computer lab at their assigned lab time and assisted with student log-in as needed.

Procedures: Comparison Instruction

Students assigned to the comparison group received SFA but did not receive the MVRC intervention. Instead of receiving MVRC, students in the comparison condition received an additional two hours of supplemental reading instruction from their

classroom teacher. During the additional two hours of reading instruction, the teachers provided materials and employed instructional practices consistent with regular classroom routines.

Data Collection

Administration of Dependent Variable Measures

Pre-test dependent variable data collection was performed in September and post-test dependent variable data collection was performed in April. The two spelling measures from the WJ IV ACH were administered to participating students in whole-class groups in an effort to save time, and they were administered beginning with the first item of each test and ending with items that were thought to be beyond the performance of any participant. Therefore, the basal (starting point) for all participants was satisfied because the lowest basal was number one. However, not all participates reached an ending point (ceiling) in their performance because some students could spell beyond the ending word administrated to the group. Students who did not reach an appropriate ending point (ceiling) were not included in the data analysis and were excluded from this study because including their scores would not be an accurate representation of their true achievement in spelling.

Scoring Reliability (Inter-Observer Agreement)

At least 25% of the data was scored for inter-observer agreement (IOA) between the researcher and an additional observer trained for traditional spelling scoring (correct vs. incorrect) and CBM-spelling scoring (correct letter sequences). The additional observer was trained until 100% of agreement was achieved on practice data for both traditional spelling scoring and CBM-spelling scoring at least three consecutive times for

both Test 3 and Test 16 from the WJ ACH IV. The inter-obsever agreement will be reported in the Chapter Four.

Fidelity of the Intervention

Two measures of fidelity were selected and included with the archival data set, usage logs and a planned activity check created to capture student engagement with observations. The MVRC program generated usage logs, which recorded the duration and frequency for students as they used the program throughout the year. The specific data collected in the usage logs included the following: (a) time spent on the program in minutes, (b) the type and number of activities the student had completed, and (c) each student's progress toward specific achievement targets. It is important to note that only active use of MVRC was reported in the usage logs because MVRC will automatically log users out of the program after three minutes of inactivity.

The researchers at MindPlay Education Software for Reading Instruction in conjuction with the third-year doctoral student in special education who supervised and coordinated all data collection activities created a planned activities check for the larger study from which this limited data set was originally derived. The planned activity check used time sampling for the behavioral observations which is a useful for school environments and required little observe time yet can yield rather valuable information. With time sampling as the time-based method of measurement, the observer records whether or not the behavior was occurring at the end of the interval (Zirpoli & Melloy, 2001).

The planned activities check was used to document the percentage of students in the experimental group who displayed engagement behavior at each observed interval.

This was created with the intent of measuring student engagement behavior twice a week during the larger study from which this limited data originated. Student engagement was defined as screen activity of the student (i.e., looking at the screen), keyboard use, physical position (i.e., head up and eyes facing the screen). When administered, two observers scanned each classroom at the same time for off-task behavior, at the end of every 2-minute interval for a total of 20 minutes (10 intervals total). Non-engagement behavior (or off-task behavior) was defined as the following: (a) student with headphones either on only one ear or sitting on the desk, (b) student engaged in activities with another person (outside of instructional interaction), and/or (c) student with a software program open other than MVRC.

Then the total number of off-task students displaying non-engagement behavior was subtracted from the total number of students present in that classroom at that time. This allowed the observers a quick, yet efficient way to document the number of participants who were on-task, as well as those who were not on-task. Given that on-task or off-task behavior (or engagement and non-engagement) is a dichotomous variable, knowing one informs the status of the other. A copy of the planned activity check form can be found in Appendix B. Results from the product usage log and the planned activity check will be provided in Chapter Four.

Data Analysis

Why Use Multilevel Modeling?

Nested data structures often exist in educational and social settings because individuals are among a variety of different groups (Kreft & Leeuw, 1998). Multilevel modeling (MLM) is most appropriate for analysis of data that are complexly nested (Garson, 2013), and multilevel modeling can account for the complexity among the

nested factors. MLM can not only account for the variations between different participants nested in the classrooms, but also the relationships between the outcome variables and the predictors at the different levels. In education, one example of MLM might include students nested in specific classrooms and schools nested in different school districts. Conducting research at only the student level while ignoring the other levels (i.e., classroom and school) can pose threats to the soundness of the research design by leading to incorrect conclusions (Bell, Ene, Smiley, & Schoeneberger, 2013). For example, a traditional analysis may only allow for an aggregate group level analysis, in which the data are combined and the individuals are ignored. In addition, not accounting for nested data may contribute to a Type I error, or a false positive, a result that indicates a certain condition was significant when in reality it was not (Beretvas, 2009). Furthermore, the statistical power to detect a possible relationship is reduced by ignoring the number of individuals in the study when averaging estimates over groups (Tabachnick, 2008). Also the meaning of a variable at Level 1 (for example student SES) may not be the same as the meaning at Level 2 (school level SES); therefore, multilevel models have been created to account for the hierarchical nesting of data (Heck & Thomas, 2000; Hox, 2010; Klein & Kozlowski, 2000; Raudenbush & Bryk, 2002; Snijders & Bosker, 1999).

Data in educational settings often violate the basic assumption of independence and therefore, traditional statistical techniques cannot be employed. However, multilevel modeling (MLM) is a method of analysis which allows for data to be nested; as such, the data is not independent of one another. Ignoring the complexity of the data in an organizational setting by considering individuals as if they were independent from their

groupings introduces a source of bias into the analysis (Heck, Thomas, & Tabata, 2014). This is due to the fact that individuals who come from the same organization, such as a classroom or school, will be more homogenous than if individuals were randomly sampled from the larger population. The individuals within these organizations share certain characteristics, and the observations of these individuals cannot be regarded as fully independent. However, most analytic techniques have a basic assumption for analysis, that observations of these individuals are independent. To account for the dependence among observations within groups a more complex error structure needs to be added to the model (Heck et al.). MLM is capable of assessing the effects of each level (*i.e.*, student-level, and classroom level) as well as the between level effects at the same time (Marsh, Kong, & Hau, 2000). Additionally, within MLM, unequal sample sizes are expected (Tabachnick & Fidell, 2007). With a traditional statistical analysis, sample size must meet a specific criteria. However, with MLM and in a real classroom, the same number of students will not exist in each classroom. It is very likely that some teachers may teach a class consisting of 20 students while another teacher teaches a class consisting of 25 students. MLM is advantageous because the researcher can include data of each of these classes regardless of the class size as they do not have to be equal (Bryk & Raudenbush, 2002).

MLM was the most appropriate statistical method for this study for a number of reasons. First, given the nested structure of the data MLM was selected because time can be nested within students who can then be nested within different teacher's classrooms. Second, because students were not randomly assigned to their specific classrooms, interdependence of the data cannot be met. Third, the first two research questions can

only be answered by including a multilevel model for repeated measures with three levels (Level 1: Time, Level 2: Student, Level 3: Teacher) because the same tests were conducted at pre-test and post-test and because of the nesting of students among time, and students in teacher's classrooms. Lastly, all eight classrooms varied in size, and with MLM, this data can still be included. MLM takes into account the nested nature of the data and allows for unequal sample sizes which are often common among classrooms (Heck et al., 2014).

Preparing the Data

Because IBM SPSS uses single-equation representation approaches for estimating the multilevel model, only one file in SPSS is needed that combines data from each level. Additionally, with time-varying models and multivariate models, the data need to be organized vertically instead of horizontally. Additionally, the study variables needed to be restructured and created to account for the nesting within cases. More specifically, time was nested within students and indicator variables for students (treatment, female, ELL, and SPED) and then students were nested within the different teachers. The definitions of the SPSS data are contained in Table 5 which includes the variable name, level, description, values, and measurement type. The variable *schoolcode* is included in the table as well as the data set, but it was not included in the multilevel models of repeated measures because there were only two schools at this level (Level: 4 had it been included) and school was only important for identifying the experimental and comparison groups. As such a student level variable was created, treatment to be able to distinguish between the comparison and experimental group. Additionally, the student level

predictors were dummy coded as seen in Table 5 for the column values for treatment, female, ELL, and SPED.

Table 5

Data Definition of SPSS Data set

Variable	Level	Description	Values	Measurement
schoolcode	School	School identifier (2 schools)	0=Comparison, 1=Experimental	Scale
Student_id	Individual	Student identifier (1-159)	Integer	Ordinal
Teacher_id	Classroom	Classroom identifier (1-8)	Integer	Ordinal
Female	Individual	Demographic predictor variable representing students gender	0=Male, 1 =Female	Scale
ELL	Individual	Demographic predictor variable representing student ELL status	0=Non-ELL, 1=ELL, 2=RFEP	Scale
SPED	Individual	Demographic predictor variable representing students Special Education status	0=Non-SPED, 1=SPED	Scale

(continued)

Table 5 (continued)

Data Definition of SPSS Data set

Variable	Level	Description	Values	Measurement
Index1time	Within Individual	Identifier variable resulting from representing the occasions at which the tests were given (pre and post)	1=pre, 2=post	Scale
treatment	Individual	Represents which students were in which treatment conditions	0=comparison, 1=experimental	Scale
spelling_3_trad	Within Individual	Dependent variable measuring achievement on the outcome for test 3 spelling	0 to 28	Scale
spelling_16_trad	Within Individual	Dependent variable measuring achievement on the outcome for test 3 spelling	0 to 17	Scale

Note. Within Individual=Level 1; Individual=Level 2; Classroom=Level 3; School=Level 4; SPED=Special Education; ELL=English Language Learners; trad= scoring as correct and incorrect; non-traditional scoring= coring using correct letter sequences.

Data Analysis Plan

Multilevel modeling of repeated measures was used to estimate the relationship of the predictors at each level of the outcome variable (spelling achievement) and find the

best model fit for the data. There are two options for estimation procedures, Restricted Maximum Likelihood (REML) or Maximum Likelihood Estimate (ML). ML produces more accurate estimates of fixed regressions parameters whereas REML produces more accurate estimates of random variances (Twisk, 2006). The decision of which estimation procedure to employ depended upon whether the study hypotheses focused on the fixed regression parameters or on estimating variances of the random effects. REML was selected as the estimation procedure because REML has been the best to use because of the smaller group sample size which "may result in a downward bias in the estimation of variance components and standard errors" (Heck et al., 2014, p. 29).

Because the data for this study were nested at levels of time, student, and teacher, a three-level multilevel model of repeated measures was constructed (Garson, 2013). Data existed at the following three levels: time (level 1), student or individual (level 2), and teacher or classroom (level 3), and are detailed in Figure 1. Multilevel modeling was used to investigate between-group variability as well as within-group variability (Garson, 2013; Heck et al., 2014). Between-group variability refers to the variability between the treatment conditions (experimental and comparison) and within-group variability refers to the variability of students between time 1 and time 2 from pre-test to post-test. Also specifically at the student level, demographic predictor variables can be added to the multilevel model of repeated measures. For this study gender, ELL classification, SPED classification, and the intervention (MVRC) on students' spelling achievement (score) for Test 3: Spelling and Test 16: Spelling of Sounds (from the WJ IV) were all student level predictors in the model.

IBM SPSS Statistics (version 22) was selected as the statistical program for running the statistical analyses because IBM SPSS MIXED is a flexible modeling routine which can be used to estimate a variety of multilevel designs with continuous and categorical outcomes (Heck et al., 2014). Also all analyses can be conducted from one data set because the statistical package, uses single-equation representation (with algebraic substitution) (Heck et al., 2014).

The following steps were taken to analyze the data and to answer the following research questions:

Research Question 1

1. Does MVRC have a significant effect on students' spelling scores from the WJ IV ACH for Test 3: Spelling, and Test 16: Spelling of Sounds over time?

Hypothesis 1

MindPlay Virtual Reading will have a significant effect on student's spelling scores for both Test 3: Spelling and Test 16: Spelling of Students from the WJ ACH IV.

Research Question 2

2. Do spelling achievement trajectories vary by gender, ELL classification, and/or SPED classification for the WJ IV ACH, Test 3: Spelling, and Test 16: Spelling of Sounds?

Hypothesis 2

Spelling achievement for all students will vary by gender. Existing literature supports the notion that females will perform better on spelling than males. Additionally, it is likely that students in SPED and ELL will make less progress in spelling from pre-test to post-test than Non-ELL or Non-SPED students.

Research Question 3

3. Which scoring measurement is more sensitive to student growth: traditional standardized scoring (correct or incorrect) or CBM spelling scoring (which credits each correct letter sequence) from the WJ IV ACH Test 3: Spelling, and Test 16: Spelling of Sounds?

Hypothesis 3

Spelling growth will be best documented by Curriculum-Based Measurement.

In order to answer the research questions one and two, a series of multilevel models of repeated measures were built using either the outcome variable of spelling achievement for Test: 3 Spelling or Test 16: Spelling of Sounds with traditional scoring. Alpha was set at .05 to designate significant parameters. After the final model was constructed, the residuals were plotted to check to make sure that the assumption of MLM had been met, that the residuals are randomly distributed residuals. The formulas of the multilevel models of repeated measures and the results as well as the residual plots are presented in Chapter Four.

To answer research question three, students' written spellings that were produced for the WJ IV, Test 3: Spelling and Test 16: Spelling of Sounds at both pre-test and post-test were treated as qualitative data. Forty-three students' written spellings were the same ones randomly pulled for inter-rater reliability of scoring which represented approximately 27.04% of the total participants of the study. For each student the number of correct was put over the number of points possible, for example 10/20 which would be 50% correct. For traditional scoring (correct or incorrect) students written spellings are scored for either the whole word correct (correct=1) or as incorrect (incorrect=0). For

non-traditional scoring (correct letter sequences) students earned a point for each correct letter sequence in the word. For CLS a phantom letter exists at the beginning of the word and at the end of the word (_dog) and at the end of the word (dog_). To calculate the correct letter sequences possible in a word, the total number of letters are added plus one additional phantom letter is added. Carrots represent correct represent correct letter sequences (^). For example, the word cat (^d^o^g^) has four CLS possible. If a student spells the word *dog* as *dogh* the student would receive three CLS (^d^o^gh). This would be recorded as 3/4. Results of the qualitative analysis of these selected 43 students' spellings are reported in Chapter Four.

CHAPTER 4: RESULTS

The purpose of this study was three-fold. The first purpose was to determine the effects of MVRC on the spelling development of second-grade students from two schools in one district in the Southwest. The second purpose was to determine if spelling achievement trajectories varied by gender, ELL classification, and/or SPED classification. The third purpose was to score the student's spelling tests with two different metrics of spelling: tests that were scored with traditional standardized scoring (correct and incorrect) and tests that were scored with Curriculum-Based Measurement (CBM) scoring for spelling (correct letter sequences) and to decide which was a more sensitive to growth from pre-test to post-test. This chapter will present the results in the following order: (a) the fidelity of implementation (including the product usage log and planned activities check), (b) the inter-observer agreement of scoring the spelling tests, (c) multilevel modeling of repeated measures analyses, (d) the qualitative spelling analysis for students' written spellings, and (e) a summary of findings for each research question.

Fidelity of Implementation

Product Usage Log

The product usage logs generated by the MindPlay Virtual Reading software for all participants in the experimental classrooms recorded that the participants actively used the intervention 90% of the time or more (a mean average of 44 hours of active use). This percentage included the remaining 159 participants that were retained for the research study.

Planned Activities Check and Inter-Observer Agreement of Student Engagement

Throughout the year, participants in the experimental classrooms were observed 45 times (approximately twice a week) during the 2013-2014 school year by two observers while they used the MVRC intervention in the computer lab. One observer served as the primary observer and the second observer served as the observer for inter-observer agreement. The planned activities check was used to record behavior observations for student engagement on MVRC using time sampling as the measurement method. Data recorded using the planned activities check was converted into a percentage of students displaying engagement with the following formula for each two minute interval:

$$\frac{number\ of\ students\ displaying\ enaggement\ or\ on\ task\ behavior}{total\ number\ of\ students\ in\ the\ computer\ lab\ for\ the\ observation} X100$$

Then the percentage from each interval (all ten of them) were averaged together, then finally aggregated together for each month. Inter-observer agreement was calculated for agreement among the two observers by dividing the number of agreement between the observers per interval divided by the total number of agreements plus disagreements. Next, the result was multiplied by 10. The amount of time students were found to be displaying engagement behavior is displayed in Table 6. Overall, participants in the experimental group were displaying engagement behavior for at least 90% or greater during the observed intervals while students used MVRC. However, inter-observer agreement (IOA) was not as high at 87.7% overall. In the behavior field, 80% or higher is often considered a high reliability for IOA. Although, there is concern for the 77% IOA for the month of November, but since the author received the data as archival data, this low IOA recorded could not be investigated further.

Table 6

Results of the Planned Activities Check for Student Engagement Behavior

Month (2013-2014)	Percentage of Engagement Behavior Per Primary Observer	Inter-Observer Agreement
October 2013	91.5%	84.5%
November 2013	91%	77%
December 2013	90.5%	86%
January 2014	93%	81.5%
February 2014	91%	85.5%
March 2014	92%	87.5%

Inter-Observer Agreement of Scoring the Spelling Tests

Comparison

For teacher one, four students out of 16 were scored by a second rater (25% of the students' data in this class). For teacher two, five of 17 students were scored by a second rater (29.4% of the students' data in this class). For teacher three, six out of 23 students were scored by a second rater (26% of the students' data in this class). For teacher four, five out of 20 students were scored by a second rater (25% of the students' data in this class). The IOA for both tests at pre-test and post-test were aggregated by the scoring metric for each teacher, summarized in Table 7. IOA was at 99% or greater each different scoring procedure and for each teacher. The disagreements that occurred were merely because of a judgment call in evaluating and scoring a student's handwriting. No disagreements were found because of incorrect scoring. Disagreements occurred over whether a letter was identified as an "n" or an "m," an "h" or an "n", and an "f "or a "t."

Table 7

IOA Scoring Results by Teacher for the Comparison Group

Teacher	Traditional Scoring	Non-Traditional Scoring
1	99.02%	99.16%
2	100%	99.14%
3	99.74%	99.27%
4	100%	100%

Experimental

For teacher five, five students out of 17 were scored by a second rater (29.4% of the students' data in this class). For teacher six, five of 18 students were scored by a second rater (27.8% of the students' data in this class). For teacher seven, seven of 26 students were scored by a second rater (26.9% of the students' data in this class). For teacher eight, six out of 22 students were scored by a second rater (27.3% of the students' data in this class). The IOA for both tests at pre-test and post-test were aggregated by the scoring metric by teacher, summarized in Table 8. IOA was at 99% or higher for each teacher and each different scoring procedure. The disagreements that occurred were merely because of a judgment call in scoring a student's handwriting. No disagreements were found because of incorrect scoring. Disagreements occurred over whether a letter was identified as an "h" or an "n", an "r" or an "n", an "e" or a "c", and a "g" or a "q".

Table 8

IOA Scoring Results by Teacher for the Experimental Group

Teacher	Traditional Scoring	Non-Traditional Scoring
5	100%	100%
6	99.98%	99.92%
7	100%	99.92%
8	99.67%	99.84%

Results of Multilevel Modeling of Repeated Measures Analyses

Visual Inspection of Data for Both Tests

Visual inspection of the data can provide preliminary clues about the shape of the change in spelling scores taking place among the individuals over time. Figures 2, 3, 4, and 5 provide a plot of the linear growth trajectories of all 159 students in the data set from pre-test to post-test for traditional scoring and non-traditional scoring. For a closer visual inspection of the data set, the first two students from each classroom were selected for a total of 16 students. Readers will note that the intercepts (*i.e.*, individual's status at Time 1) appear to vary considerably and the steepness of the growth over time varies within this subset of 16 students. The variability for the subset of 16 students is summarized in Table 9.

Table 9

Variability of the Student Subset (N=16)

Test and Scoring Metric	Range	Minimum	Maximum	Mean	Std. Deviation
Test 3 spelling with traditional scoring	20.00	1.00	21.00	9.6250	4.94975
Test 3 with non-traditional scoring	145.00	17.00	162.00	74.6563	40.65560
Test 16 spelling of sounds with traditional scoring	14.00	1.00	15.00	8.6875	2.65716
Test 16 spelling of sounds with non-traditional scoring	110.00	22.00	132.00	81.1875	23.79406

The figures of the individual's scores over time show that some individuals have increase in their knowledge of spelling from pre-test to post-test in both groups (experimental and comparison), while others have decreases in their knowledge of spelling from pre-test to post-test. However, for purposes of answering the research question, the focus is directed toward whether the experimental group increases more in spelling ability than the comparison group. This question is addressed in the multilevel model of repeated measures model analyses further. Additionally, the plot of the experimental versus comparison groups shows that the experimental group is increasing at a greater rate than the control group over time. A limitation of visually inspecting the data shown in Figures 2, 3, 4, and 5 is that these figures do not take into account the nesting that is present among the different time points or the individuals level predictors (such as demographic variables). Instead, this is aggregate data of the experimental group and the comparison group as displayed in Figures 6, 7, 8, and 9. Moreover, Figures 10,

11, 12, and 13 were constructed at classroom level for Test 3: Spelling and Test 16: Spelling of Sounds (WJ ACH IV) by traditional and non-traditional scoring to represent the growth of each teacher's students averages, each line in each figure represents a different teacher's classrooms. Evident in Figures 14, 15, 16, and 17 is the fact that each classroom has a different slope that describes the direction and the steepness of the line.

Multilevel Model for Repeated Measures Formulas

Multilevel models can be longitudinal in nature and account for the growth present in an outcome variable (dependent variable) with a handful of specific categorical or continuous observations nested within subjects (Hox, 1995). Another hierarchical level occurs in longitudinal students when the same subjects are measured at more than one time. Measurement repetitions make up the Level 1 and subjects make up the Level 2 of the multilevel model for repeated measures. A multilevel model for repeated measures not only contains students' original scores or status at pre-test but also the rate at which students change or grow. Additionally, the teacher or school can be added as a third level. A straightforward model for Level 1 would be:

$$Y_{ti} = \pi_{oi} + \pi_{1i} time_{ti} + \pi_{2i} time_{ti} + \varepsilon_{ti.} \quad (4.1)$$

Where Y_{ti} is the observed status at time t for the individual i which is a function of the systematic growth trajectory plus random error. Where π_{oi} is average achievement across the two occasions and ε_{ti} represents the errors in predicting the average spelling achievement of students.

Since the period of time is shorter (about one year) and there are only two time points (pre-test and post-test), one might assume that a linear model will fit reasonably

well. The figure suggests that the comparison and experimental groups are close before the intervention begins at time one when the pre-test was given.

At Level 2 it is assumed that the intercept varies between students for time one:

$$\pi_{0i} = \beta_{00} + B_{01}treatment_i + B_{11}treatment_i + u_{0i}. \quad (4.2)$$

At Level 2 it is assumed that the intercept varies between students for time two (varying slope):

$$\pi_{1i} = \beta_{00} + \beta_{10} + B_{o1}treatment_i + B_{11}treatment_i + u_{1i.} \quad (4.3)$$

Where β_{00} and β_{10} are the intercepts which describes the average initial means between students Where $B_{01}treatment_i$ and $B_{11}treatment_i$ are the different occasions at each time the measurement was given by each of the treatment groups (comparison and experimental).

Where u_{0t} and u_{1i} are the Level 2 random component associated with describing the differences in average achievement (an outcome measure such as math achievement) between students.

From substituting the Level 2 equations into the Level 1 equation, the combined model is:

$$Y_{ti} = \beta_{00} + \beta_{01}treatment_i + \beta_{011}treatment_i + \beta_{10}time_{ti} + \beta_{11}time_{ti} * treatment_i + u_{1i}time_{ti} + u_{0i} + \varepsilon_{ti.} \quad (4.4)$$

The key parameter in the above model is $\beta_{11}time_{ti} * treatment_i$ which is used to test whether growth trajectories for individuals differ in comparison and experimental groups. Next teacher level was added to the polynomial growth model to predict spelling achievement. The significance of the parameter estimates determined if teacher level was

significantly related to the spelling achievement. Below is the equation of teacher added at level 3 to model the initial status and the rate of growth accounting by teacher:

$$Y_{ti} = \beta_{00} + \beta_{01}treatment_i + \beta_{011}treatment_i + \beta_{10}time_{ti} + +\beta_{11}time_{tij} * treatment_{ij} + \beta_{02}gender_{ij} + \beta_{03}ell_{ij} + \beta_{04}sped_{ij} + u_{1i}time_{tij} + u_{0ij} + \varepsilon_{tij.} \quad (4.5)$$

Where j represents teacher.

At level 1 and 2 an unstructured covariance matrix was specified:

$$\begin{bmatrix} \sigma_{1,1}^2 & \sigma_{1,2} \\ \sigma_{2,1} & \sigma_{2,2}^2 \end{bmatrix} \quad (4.7)$$

Teacher was added at a random effect for Level 3 to account for the variability of students being a part of different classrooms during the study. At Level 3 it is assumed that the intercept varies between teacher.

Multilevel Model Building Process

This section describes the multilevel model building process completed by the author during this study. First, the process began with a null model which was used as a baseline for the beginning relative fit and for comparing later models as they were built. The null model also allowed for the inter-class correlation (ICC) to be determined, which is a vital component in the analysis which would support the need for a Multilevel Model. Next, a within subjects model was developed to represent the student's growth trajectories. For this model time and treatment were are added and the results indicate the relationship between Level 1 and the outcome. Next, a between-individual model was constructed with the student predictors at added (female, SPED, and ELL). Finally, a between group model was built, which included the teacher and student at random effects with the student predictors added (Female, ELL, and SPED). Within each step of building the model, the significant variables were identified from the output tables generated in

IBM SPSS MIXED. Additionally, the author also compares the relative fit statistics of each model in relation to the null using a likelihood ratio test that is often used to compare the goodness of fit of two models, usually the null model to another alternative model. The statistical test indicates how many times more likely the data fits one model over the other. Moreover, a model with smaller -2 log likelihood (LL) and a smaller number of parameters are usually the best fit for the model (degrees freedom).

Null model for test three. The null model was generated for a three-level multilevel model of repeated measures. Students and teachers were specified as the subjects with each student's scores from pretest to post-test for Test 3: Spelling specified as the dependent variable.

This basic model is necessary for comparisons between more complicated models as they are built. The equations for each level of this model are specified below:

Level 1: $Y_{ti} = \pi_{oi} + \pi_{1i} time_{ti} + \pi_{2i} time_{ti} + \varepsilon_{ti.}$

Level 2: $\pi_{tij} = \beta_{t0j} + r_{tij}.$

Level 3: $\beta_{t0j} = \gamma_{t00} + u_{toj}.$

The ICC for each level can be found with the variance, σ^2, for each level using the following equations:

Level 1: $ICC = \frac{\sigma^2_{level\ 1}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

Level 2: $ICC = \frac{\sigma^2_{level\ 2}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

Level 3: $ICC = \frac{\sigma^2_{level\ 3}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

After running the null model in IBM SPSS MIXED with a variety of difference covariance types (i.e., Diagonal, AR heterogeneous, AR (1), and Unstructured) it was

determined that the covariance type diagonal had the lowest -2LL which meant it was a better fit. In addition, the covariance type for repeated effects for the null model was specified as scale identity. The null model estimated five parameters (the intercept for the fixed effect, two random effects and two repeated effects). The -2 log likelihood produced for the null model was 1873.923 provided in Table 10. The resulting estimates of fixed effects are presented in Table 11 and includes the intercept of 7.517.

Table 10

Information Criteria[a] for the Null Model

-2 Restricted Log Likelihood	1873.923
Akaike's Information Criterion (AIC)	1881.923
Hurvich and Tsai's Criterion (AICC)	1882.052
Bozdogan's Criterion (CAIC)	1900.959
Schwarz's Bayesian Criterion (BIC)	1896.959

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 11

Estimates of Fixed Effects for the Null Model

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	7.516768	.544177	6.500	13.813	.000	6.209668	8.823869

a. Dependent Variable: spelling_3_trad.

In Table 12, the estimates of covariance parameters, it is noted that the variance component within individuals is 2.002 for time one and 30.610 for time two, between individuals is 13.991, and between teachers is 1.652. From these numbers, we can gather that there is a greater amount of variability due the repeated measures (time) and between individuals. Next, the ICC was calculated and checked at each level, Table 13 displays

those results. With the results from Table 12, it was found that 4.15% of the variability in students' spelling achievement for Test 3: Spelling could be attributed to time 1, 63.43% of the of the variability in students' spelling achievement for Test 3: Spelling could be attributed to time 1, 29% of the variability could be attributed to differences between students, and 3.42% of the variability could be attributed to differences between the teacher groups. These results support the author's decision for using three-level model, specifically a three-level multilevel model of repeated measures. It also emphasizes the importance of accounting for all three levels of the hierarchal structure of the data (time, student, and teacher).

Table 12

Estimates of Covariance Parameters[a] for the Null Model

						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	Var: [Index1time=1]	2.002110E-9[b]	.000000	.	.	.	.
	Var: [Index1time=2]	30.610063	3.433054	8.916	.000	24.569606	38.135570
Intercept [subject = student_id]	Variance	13.991388[b]	.000000	.	.	.	.
Intercept [subject = teacher_id]	Variance	1.651924	1.310522	1.261	.207	.348904	7.821215

[a]Dependent Variable: spelling_3_trad.
[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Table 13

ICC Calculations for the Null Model

Level	Calculation	Result
Level 1 (time 1)	2.002110/48.255485	4.15%
Level 1 (time 2)	30.610063/48.255485	63.43%
Level 2	13.991388/48.255485	29%
Level 3	1.651924/48.255485	3.42%

Within-individual model for test three. The within-individual model was used to build the individual student growth trajectories for spelling achievement for Test 3: Spelling. This model was created by adding time into the null model as well as the variable treatment (used to identify who was in the comparison and who was in the experimental group). Table 14 contains a lower -2LL of 1677.847 therefore we can say that this model is a better fit than the null model which generated a -2LL of 1873.923. Table 15 displays the results of the fixed effects for the within-model with the two parameters, time and treatment added. For the *F*-statistics one parameter (treatment) was not significant (.301, $p>.05$) and three parameters were significant. The three significant parameters were intercept ($t=3.914$, $p<.05$), Index1time ($t=10.494$, $p<.05$), and the interaction effect of Index1time and treatment ($t=3.646$, $p<.05$). Students in the comparison group started with a mean score of 3.314 words at pre-test per Table 16 in the fixed effect estimates. At post-test the comparison group increased by 3.645 points on average. Over time, it is observed that students who were in the experimental group and received MindPlay Virtual Reading increased their scores at a greater rate 1.752854 points, than their peers the comparison group.

Table 16 contains the variance component table for the random effects. At Level 1, the variance summarizes the variability of student spelling achievement for test three

around his her own growth trajectory for time one and time two (Singer & Willet, 2003). At Level 2, the variance summarizes the variability of student spelling achievement between students. At Level 3, the variance summarizes the variability of student spelling achievement between teacher groups.

Shown in Table 16, the Wald Z test is a used to test whether each coefficient in the model is statistically significant. A Wald Z test calculates a *z statistic.* The null of the Wald Z test is that the population parameter for variance equals zero (Singer & Willet, 2003). None of the Wald Z tests could be estimated because they have already been estimated in the fixed effects of this model.

Table 14

Information Criteria[a] for the Within-Individual Model with Time and Treatment Parameters

-2 Restricted Log Likelihood	1677.847
Akaike's Information Criterion (AIC)	1685.847
Hurvich and Tsai's Criterion (AICC)	1685.977
Bozdogan's Criterion (CAIC)	1704.845
Schwarz's Bayesian Criterion (BIC)	1700.845

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 15

Estimated of Fixed Effects for the Within-Individual Model with Time and Treatment Parameters

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	3.313930	.846736	8.152	3.914	.004	1.367689	5.260171
Index1time	3.644737	.347309	157.002	10.494	.000	2.958736	4.330737
treatment	-.651065	1.187044	7.869	-.548	.599	-3.396349	2.094220
Index1time * treatment	1.752854	.480701	157.002	3.646	.000	.803378	2.702330

[a]Dependent Variable: spelling_3_trad.

Table 16

Estimates of Covariance Parameters[a] for the Within-Individual Model with Time and Treatment Parameters

						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	Var: [Index1time=1]	2.002110E-9b	.000000	.	.	.	.
	Var: [Index1time=2]	30.610063	3.433054	8.916	.000	24.569606	38.135570
Intercept [subject = student_id]	Variance	13.991388b	.000000	.	.	.	.
Intercept [subject = teacher_id]	Variance	1.651924	1.310522	1.261	.207	.348904	7.821215

[a]Dependent Variable: spelling_3_trad.
[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Between-individual model for test three. Additional student parameters (female, ELL, and SPED) were added at Level 2. The significance tests and relative fit statistics were used to determine if the added parameters explained the variation in student spelling achievement for test three. This model should be a better fit than the previous model because additional student level predictors are being added to the model. In Table 17 the information criteria for the within-individual model containing student demographics, time, and treatment parameters. Indeed, this model has a lower -2LL of 1615.207. In Table 18, the significant parameters are bolded (time, female, ell_status, and the interaction of time and treatment. This model estimated four parameters as significant: time, female, and ELL. Table 19 contains the estimates of covariance parameters for the between-individual model.

Table 17

Information Criteria[a] for the Between-Individual Model

-2 Restricted Log Likelihood	1615.207
Akaike's Information Criterion (AIC)	1625.207
Hurvich and Tsai's Criterion (AICC)	1625.409
Bozdogan's Criterion (CAIC)	1648.759
Schwarz's Bayesian Criterion (BIC)	1643.759

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 18

Estimated of Fixed Effects for the Between-Individual Model

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	.944423	1.296563	16.921	.728	.476	-1.792061	3.680907
Index1time	3.556136	.617205	151	5.762	**.000**	2.336664	4.775608
treatment	1.574317	1.776614	15.000	.886	.390	-2.212442	5.361075
female	2.517503	1.056982	147.190	2.382	**.019**	.428682	4.606323
sped_status	-1.602584	1.959030	146.968	-.818	.415	-5.474091	2.268923
ell_status	1.946253	.850119	159.152	2.289	**.023**	.267283	3.625223
Index1time * treatment	1.786009	.820109	151.000	2.178	**.031**	.165638	3.406380
Index1time * female	.152550	.695700	151.000	.219	.827	-1.222014	1.527114
Index1time * sped_status	-1.715185	1.290472	151.000	-1.329	.186	-4.264899	.834528
Index1time * ell_status	.216399	.538464	151.000	.402	.688	-.847498	1.280296
treatment * female	-2.634812	1.462985	145.896	-1.801	.074	-5.526192	.256569
treatment * sped_status	1.025137	2.411128	150.394	.425	.671	-3.738923	5.789196
treatment * ell_status	-1.227965	1.222245	163.743	-1.005	.317	-3.641358	1.185427
Index1time * treatment * female	-.379183	.966617	151.000	-.392	.695	-2.289024	1.530657
Index1time * treatment * sped_status	.419164	1.571769	151.000	.267	.790	-2.686335	3.524662
Index1time * treatment * ell_status	.612748	.759853	151.000	.806	.421	-.888569	2.114065

[a]Dependent Variable: spelling_3_trad.

Table 19

Estimates of Covariance Parameters[a] *for the Between-Individual Model*

						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	UN (1,1)	5.437823	1.301935	4.177	.000	3.401172	8.694037
	UN (2,1)	5.306396	1.530707	3.467	.001	2.306264	8.306527
	UN (2,2)	14.248481	2.333709	6.106	.000	10.336017	19.641918
Intercept [subject = student_id]	Variance	5.633103[b]	.000000	.	.	.	.
Intercept [subject = teacher_id]	Variance	3.202180	2.287056	1.400	.161	.789772	12.983435

[a]Dependent Variable: spelling_3_trad.

[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Between-group model for test three. In the final model female, SPED, and ELL were added at the student level and teacher level of random effects. The covariance type at repeated was specified as unstructured. The covariance type at random effects student and teacher was also specific as unstructured. In Table 20, the final model provided the best fit with a 1610.32 -2LL, the lowest estimated for all models to this point. In Table 21 the significant results are bolded, time and the interaction of time and treatment. Because the variable treatment is significant, this illustrates that the treatment growth for the comparison group was significantly different from the treatment group at the beginning of study (3.556, p=.000). Over the time period of this study, the treatment group demonstrated considerably higher growth compared to the comparison group (1.786,p<.05). Table 22 provides the estimates of covariance parameters. Tables 23 and 24 provide the random effects for the model.

Table 20

Information Criteria[a] for the Between-Group Model

-2 Restricted Log Likelihood	1610.132
Akaike's Information Criterion (AIC)	1656.132
Hurvich and Tsai's Criterion (AICC)	1660.103
Bozdogan's Criterion (CAIC)	1764.472
Schwarz's Bayesian Criterion (BIC)	1741.472

The information criteria are displayed in smaller-is-better form

Table 21

Estimated of Fixed Effects for the Between-Group Model

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	1.859466	1.492444	.154	1.246	.756	-575048.076325	575051.795258
Index1time	3.556136	.631686	112.249	5.630	**.000**	2.304561	4.807711
treatment	1.332590	2.014948	.139	.661	.836	-263311.160469	263313.825649
female	2.559024	1.391500	3.988	1.839	.140	-1.308904	6.426951
sped_status	-.999759	2.090191	.380	-.478	.788	-1924.993573	1922.994054
ell_status	.699415	1.494379	.785	.468	.737	-36.266722	37.665552
Index1time * treatment	1.786009	.839352	112.249	2.128	**.036**	.122981	3.449037
Index1time * female	.152550	.712024	112.249	.214	.831	-1.258200	1.563300
Index1time * sped_status	-1.715185	1.320751	112.249	-1.299	.197	-4.332021	.901651
Index1time * ell_status	.216399	.551098	112.249	.393	.695	-.875505	1.308303
treatment * female	-2.841327	1.933362	4.146	-1.470	.213	-8.135675	2.453022
treatment * sped_status	.561984	2.641653	1.065	.213	.865	-28.519256	29.643223

(continued)

Table 21 (continued)

Estimated of Fixed Effects for the Between-Group Model

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
treatment * ell_status	-.315225	2.089558	1.168	-.151	.902	-19.341951	18.711500
Index1time * treatment * female	-.379183	.989297	112.249	-.383	.702	-2.339301	1.580935
Index1time * treatment * sped_status	.419164	1.608648	112.249	.261	.795	-2.768089	3.606416
Index1time * treatment * ell_status	.612748	.777682	112.249	.788	.432	-.928092	2.153588

[a]Dependent Variable: spelling_3_trad.

Table 22

Estimates of Covariance Parameters[a] for the Between-Group Model

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Repeated Measures	UN (1,1)	4.481783	2.703544	1.658	.097	1.373981	14.619106
	UN (2,1)	2.835001	2.332939	1.215	.224	-1.737476	7.407477
	UN (2,2)	10.692518	3.630384	2.945	.003	5.496361	20.801026
Intercept + female + sped_status + ell_status [subject = student_id]	UN (1,1)	4.448892[b]	.000000	.	.	.	.
	UN (2,1)	-.694589	1.801396	-.386	.700	-4.225260	2.836081
	UN (2,2)	3.332245[b]	.000000	.	.	.	.
	UN (3,1)	-2.800000	1.992548	-1.405	.160	-6.705323	1.105323
	UN (3,2)	.468921	14.256375	.033	.974	-27.473061	28.410903
	UN (3,3)	3.065032[b]	.000000	.	.	.	.
	UN (4,1)	.670706	11.231313	.060	.952	-21.342264	22.683676
	UN (4,2)	-.347765	8.457664	-.041	.967	-16.924482	16.228952
	UN (4,3)	-1.532109	52.031827	-.029	.977	-103.512615	100.448397
	UN (4,4)	1.127286	25.213723	.045	.964	1.031398E-19	12320895200716945000.000000

(continued)

Table 22 (continued)

Estimates of Covariance Parameters[a] for the Between-Group Model

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept + female + sped_status + ell_status [subject = teacher_id]	UN (1,1)	3.946512	30.766544	.128	.898	9.127374E-7	17064004.823143
	UN (2,1)	-1.095501	9.594015	-.114	.909	-19.899424	17.708422
	UN (2,2)	2.461475	4.529173	.543	.587	.066829	90.662273
Intercept + female + sped_status + ell_status [subject = student_id]	UN (3,1)	-1.892300[b]	.000000	.	.	.	.
	UN (3,2)	2.174005[b]	.000000	.	.	.	.
	UN (3,3)	2.277314[b]	.000000	.	.	.	.
	UN (4,1)	-1.458366[b]	.000000	.	.	.	.
	UN (4,2)	.215587[b]	.000000	.	.	.	.
	UN (4,3)	-.075713[b]	.000000	.	.	.	.
	UN (4,4)	4.238393	9.626482	.440	.660	.049420	363.495689

[a]Dependent Variable: spelling_3_trad.

[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Table 23

Random Effects Covariance Structures (G)for the Between-Group Model[a] for Student[b]

	Intercept | student_id	female | student_id	sped_status | student_id	ell_status | student_id
Intercept | student_id	4.448892	-.694589	-2.800000	.670706
female | student_id	-.694589	3.332245	.468921	-.347765
sped_status | student_id	-2.800000	.468921	3.065032	-1.532109
ell_status | student_id	.670706	-.347765	-1.532109	1.127286

[a]Dependent Variable: spelling_3_trad.
[b] Unstructured

Table 24

Random Effects Covariance Structures (G)for the Between-Group Model[a] for Teache[r]

	Intercept | teacher_id	female | teacher_id	sped_status | teacher_id	ell_status | teacher_id
Intercept | teacher_id	3.946512	-1.095501	-1.892300	-1.458366
female | teacher_id	-1.095501	2.461475	2.174005	.215587
sped_status | teacher_id	-1.892300	2.174005	2.277314	-.075713
ell_status | teacher_id	-1.458366	.215587	-.075713	4.238393

[a]Dependent Variable: spelling_3_trad.

Residual analysis for test three. The residuals of the all models were tested with the Kolmogorov-Smirnov (K-S) test for normality in SPSS which tests whether the results of the residuals were normally distributed or non-normally normality. Only the last model the between-subjects group had a normal distribution. These results are displayed in Table 25. Therefore, this is the only model that should be used in reporting results because it is the only model not to violate the assumption of MLM, that the residuals are normally distributed. The visual inspection of the histogram in Figure 18, as

well as the results from the K-S test confirmed that the residuals of the between-group model are normally distributed.

Table 25

Hypothesis Test Summary for the Between-Group Residuals

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of the Between-Group Residuals is normal with the mean -0.000 and standard deviation 1.08	One Sample Kolmogorov-Smirnov Test	.086[1a]	Retain the null hypothesis.

[1]Lilliefors Corrected
[a]The significance level is .05.

Null model for test sixteen. The null model was generated for a three-level multilevel model of repeated measures. Students and teachers were specified as the subjects with each student's scores from pretest to post-test for Test 3: Spelling, specified as the dependent variable.

This basic model is necessary for comparisons between more complicated models as they are built. The equations for each level of this model are specified below:

Level 1: $Y_{ti} = \pi_{oi} + \pi_{1i} time_{ti} + \pi_{2i} time_{ti} + \varepsilon_{ti.}$

Level 2: $\pi_{tij} = \beta_{t0j} + r_{tij}.$

Level 3: $\beta_{t0j} = \gamma_{t00} + u_{toj}.$

The ICC for each level can be found with the variance, σ^2, for each level using the following equations:

Level 1: $ICC = \frac{\sigma^2_{level\ 1}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

Level 2: $ICC = \frac{\sigma^2_{level\ 2}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

Level 3: $ICC = \frac{\sigma^2_{level\ 3}}{\sigma^2_{level\ 1} + \sigma^2_{level\ 2} + \sigma^2_{level\ 3}}$

After running the null model in IBM SPSS MIXED with a variety of difference covariance types (i.e., Diagonal, AR heterogeneous, AR (1), and Unstructured) it was determined that the covariance type diagonal had the lowest -2LL which meant it was a better fit. In addition, the covariance type for repeated effects for the null model was specified as scale identity. The null model estimated five parameters (the intercept for the fixed effect, two random effects and two repeated effects). The -2 log likelihood produced for the null model was 1602.828 provided in Table 26. The resulting estimates of fixed effects are presented in Table 27 and includes the intercept of 7.956.

Table 26

Information Criteria[a] for the Null Model

-2 Restricted Log Likelihood	1602.828
Akaike's Information Criterion (AIC)	1610.828
Hurvich and Tsai's Criterion (AICC)	1610.956
Bozdogan's Criterion (CAIC)	1629.863
Schwarz's Bayesian Criterion (BIC)	1625.863

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 27

Estimates of Fixed Effects for the Null Model

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	7.955666	.367101	6.885	21.672	.000	7.084650	8.826682

a. Dependent Variable: spelling_3_trad.

In Table 28, the estimates of covariance parameters, it is noted that the variance component within individuals is 6.580 for time one and 6.891 for time two, between individuals is 2.353, and between teachers is .784791. From these numbers, we can gather that there is a greater amount of variability due the repeated measures (time). Next, the ICC was calculated and checked at each level, Table 13 displays those results. With the results reported in Table 29, it was found that 39.62% of the variability in students' spelling achievement for Test 3: Spelling could be attributed to time 1, 41.59% of the of the variability in students' spelling achievement for Test 3: Spelling could be attributed to time 1, 14.16% of the variability could be attributed to differences between students, and 4.73% of the variability could be attributed to differences between the teacher groups. These results support the author's decision for using three-level model, specifically a three-level multilevel model of repeated measures. It also emphasizes the importance of accounting for all three levels of the hierarchal structure of the data (time, student, and teacher).

Table 28

Estimates of Covariance Parameters[a] for the Null Model

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Repeated Measures	Var: [Index1time =1]	6.580585	1.221048	5.389	.000	4.574266	9.466897
	Var: [Index1time=2]	6.891113	1.242409	5.547	.000	4.839773	9.811914
Intercept [subject = student_id]	Variance	2.352519	.759780	3.096	.002	1.249181	4.430380
Intercept [subject = teacher_id]	Variance	.784791	.581080	1.351	.177	.183866	3.349700

[a]Dependent Variable: spelling_3_trad.
[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed

Table 29

ICC Calculations for the Null Model

Level	Calculation	Result
Level 1 (time 1)	6.580585/16.608696	0.3962132247
Level 1 (time 2)	6.891113/16.608696	0.4149099363
Level 2	2.352519/16.608696	0.1416438112
Level 3	.784791/16.608696	0.0472518131

Within-individual model for test sixteen. The within-individual model was used to build the individual student growth trajectories for spelling achievement for Test 16: Spelling of Sounds. This model was created by adding time into the null model as well as the variable treatment (used to identify who was in the comparison and who was in the

experimental group). Table 30 contains a lower -2LL of 1499.612indicating a better fit than the null model which generated a -2LL of 1602.828.

Table 31 displays the results of the fixed effects for the within-model with the two parameters, time and treatment added. For the *F*-statistics one parameter (treatment) was not significant (-1.492, $p>.05$) and three parameters were significant. The three significant parameters were intercept ($t=8.231$, $p<.05$), Index1time ($t=4.509$, $p<.05$), and the interaction effect of Index1time and treatment ($t=4.228$, $p<.05$). Students in the comparison group started with a mean score of 5.270 words at pre-test per Table 31 in the fixed effect estimates. At post-test the comparison group increased by 1.395 points on average. Over time, it is observed that students who were in the experimental group and received MindPlay Virtual Reading increased their scores at a greater rate 1.810 points, than their peers the comparison group.

Table 32 contains the variance component table for the random effects. At Level 1, the variance summarizes the variability of student spelling achievement for test three around his her own growth trajectory for time one and time two (Singer & Willet, 2003). At Level 2, the variance summarizes the variability of student spelling achievement between students. At Level 3, the variance summarizes the variability of student spelling achievement between teacher groups.

Table 30

Information Criteria[a] for the Within-Individual Model with Time and Treatment Parameters

-2 Restricted Log Likelihood	1499.612
Akaike's Information Criterion (AIC)	1509.612
Hurvich and Tsai's Criterion (AICC)	1509.807
Bozdogan's Criterion (CAIC)	1533.359
Schwarz's Bayesian Criterion (BIC)	1528.359

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 31

Estimated of Fixed Effects for the Within-Individual Model with Time and Treatment Parameters

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	5.270532	.640325	25.290	8.231	.000	3.952524	6.588540
Index1time	1.394737	.309325	157	4.509	.000	.783762	2.005712
treatment	-1.492094	.891942	23.928	-1.673	.107	-3.333265	.349078
Index1time * treatment	1.810082	.428129	157	4.228	.000	.964447	2.655718

[a]Dependent Variable: spelling_3_trad.

Table 32

Estimates of Covariance Parameters[a] for the Within-Individual Model with Time and Treatment Parameters

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Repeated Measures	UN (1,1)	4.997736	.880953	5.673	.000	3.537787	7.060166
	UN (2,1)	1.277708	.690330	1.851	.064	-.075313	2.630730
	UN (2,2)	4.829502	.856214	5.641	.000	3.411884	6.836132
Intercept [subject = student_id]	Variance	2.626999b	.000000	.	.	.	.

[a]Dependent Variable: spelling_3_trad.
[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Between-individual model for test sixteen. Additional student parameters (female, ELL, and SPED) were added at Level 2. The significance tests and relative fit statistics were used to determine if the added parameters explained the variation in student spelling achievement for test three. This model should be a better fit than the previous model because additional student level predictors are being added to the model. In Table 33, the information criteria for the within-individual model containing student demographics, time, and treatment parameters. Indeed, this model has a lower -2LL of 1433.870. In Table 34, the significant parameters are bolded (intercept, Index1time, sped_status, and ELL status).

Table 33

Information Criteria[a] for the Between-Individual Model

-2 Restricted Log Likelihood	1433.870
Akaike's Information Criterion (AIC)	1443.870
Hurvich and Tsai's Criterion (AICC)	1444.073
Bozdogan's Criterion (CAIC)	1467.423
Schwarz's Bayesian Criterion (BIC)	1462.423

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 34

Estimates of Fixed Effects for the Between-Individual Model[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	3.665213	1.027544	73.015	3.567	**.001**	1.617329	5.713098
Index1time	1.706712	.562122	151.000	3.036	**.003**	.596073	2.817352
treatment	-.340411	1.381001	64.153	-.246	.806	-3.099150	2.418328
female	1.764794	1.053292	149.460	1.676	.096	-.316472	3.846061
sped_status	-4.308199	1.953206	149.317	-2.206	**.029**	-8.167693	-.448704
ell_status	1.626547	.825875	156.063	1.969	.051	-.004788	3.257881
Index1time * treatment	1.515597	.746918	151.000	2.029	**.044**	.039837	2.991356
Index1time * female	-.283640	.633612	151.000	-.448	.655	-1.535530	.968250
Index1time * sped_status	.192832	1.175303	151.000	.164	.870	-2.129331	2.514994

(continued)

Table 34 (continued)

Estimates of Fixed Effects for the Between-Individual Model[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Index1time * ell_status	-.276662	.490408	151.000	-.564	.573	-1.245610	.692287
treatment * female	-1.094940	1.461594	148.787	-.749	.455	-3.983102	1.793222
treatment * sped_status	2.282344	2.387367	151.151	.956	.341	-2.434574	6.999263
treatment * ell_status	-.456985	1.173311	159.251	-.389	.697	-2.774243	1.860272
Index1time * treatment * female	.279482	.880350	151.000	.317	.751	-1.459913	2.018878
Index1time * treatment * sped_status	-.036253	1.431495	151.000	-.025	.980	-2.864599	2.792093
Index1time * treatment * ell_status	.182115	.692039	151.000	.263	.793	-1.185216	1.549446

[a]Dependent Variable: spelling_3_trad.

Table 35

Estimates of Covariance Parameters[a] for the Between-Individual Model

						95% Confidence Interval	
	Parameter	Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	UN (1,1)	5.185459	.690459	7.510	.000	3.994360	6.731738
	UN (2,1)	1.553610	.526024	2.953	.003	.522622	2.584598
	UN (2,2)	5.447995	.715898	7.610	.000	4.210988	7.048382
Intercept [subject = student_id]	Variance	.644071b	.000000	.	.	.	.
Intercept [subject = teacher_id]	Variance	.729465	.578040	1.262	.207	.154347	3.447549

[a]Dependent Variable: spelling_3_trad.
[b]This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Between-group model for test sixteen. In the final model female, SPED, and ELL were added at the student level and teacher level of random effects. The covariance type at repeated was specified as unstructured. The covariance type at random effects student and teacher was also specific as unstructured. The model had a higher -2LL of 1454.583, seen in Table 36. This might be accounted for the fact that the residuals were no normally distributed in the between-individual model. Other covariance types were investigated and lead to lesser -2LL's however, none of the models were normally distributed which is an important assumption of the residuals in MLM. Therefore, this current model with the higher -2LL was retained.

In Table 37, the significant results are bolded intercept, time and the interaction of time and treatment. Because the variable treatment is not significant, this means that the

spelling growth for the comparison group was not significantly different from the treatment group at the beginning of the study (-.578606, $p>.05$). Over the time period of this study, the treatment group demonstrated considerably higher growth compared to the comparison group (1.741,$p<.05$).

Table 36

Information Criteria[a] for the Between-Group Model

-2 Restricted Log Likelihood	1454.583
Akaike's Information Criterion (AIC)	1500.583
Hurvich and Tsai's Criterion (AICC)	1504.512
Bozdogan's Criterion (CAIC)	1609.150
Schwarz's Bayesian Criterion (BIC)	1586.150

The information criteria are displayed in smaller-is-better form.
[a]Dependent Variable: spelling_3_trad.

Table 37

Estimated of Fixed Effects for the Between-Groups Model[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	4.070164	.984517	454.336	4.134	**.000**	2.135391	6.004936
Index1time	1.582881	.371524	301.253	4.261	**.000**	.851770	2.313991
treatment	-.578606	1.242716	631.313	-.466	.642	-3.018963	1.861750
female	1.650069	.965181	520.524	1.710	.088	-.246060	3.546198
sped_status	-4.225430	1.579072	.491	-2.676	.393	-290.619439	282.168580
ell_status	1.130897	.941985	55.696	1.201	.235	-.756352	3.018146
Index1time * treatment	1.740509	.364781	301.253	4.771	**.000**	1.022667	2.458351
Index1time * female	-.143776	.358179	301.253	-.401	.688	-.848625	.561073

(continued)

Table 37 (continued)

Estimated of Fixed Effects for the Between-Groups Model[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Index1time * sped_status	.155759	.545767	301.253	.285	.776	-.918241	1.229758
Index1time * ell_status	-.188914	.281771	301.253	-.670	.503	-.743401	.365574
treatment * female	-1.063524	1.113214	397.873	-.955	.340	-3.252040	1.124993
treatment * sped_status	2.182435	1.786339	.199	1.222	.718	-168711.885200	168716.250070
treatment * ell_status	-.010480	1.221445	21.264	-.009	.993	-2.548691	2.527730
Index1time * sped_status	.155759	.545767	301.253	.285	.776	-.918241	1.229758

a. Dependent Variable: spelling_16_trad

Table 38

Estimates of Covariance Parameters[a] *for the Between-Groups Model*[a]

						95% Confidence Interval	
	Parameter	Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	UN (1,1)	2.120849	.902523	2.350	.019	.921049	4.883563
	UN (2,1)	-.388988	.940820	-.413	.679	-2.232960	1.454985
	UN (2,2)	2.099505	1.181949	1.776	.076	.696498	6.328691

(continued)

Table 38 (continued)

Estimates of Covariance Parameters[a] for the Between-Groups Model[a]

						95% Confidence Interval	
	Parameter	Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Intercept + female + sped_status + ell_status [subject = student_id]	UN (1,1)	1.986523[b]	.000000	.	.	.	.
	UN (2,1)	-.320195[b]	.000000	.	.	.	.
	UN (2,2)	1.986523[b]	.000000	.	.	.	.
	UN (3,1)	.564678	1.537485	.367	.713	-2.448738	3.578094
	UN (3,2)	-.904985[b]	.000000	.	.	.	.
	UN (3,3)	1.959125[b]	.000000	.	.	.	.
	UN (4,1)	.585047	3.509973	.167	.868	-6.294373	7.464467
	UN (4,2)	-.441067[b]	.000000	.	.	.	.
	UN (4,3)	-1.010931[b]	.000000	.	.	.	.
	UN (4,4)	1.437340	6.501592	.221	.825	.000203	10182.241573
Intercept + female + sped_status + ell_status [subject = teacher_id]	UN (1,1)	1.669791[b]	.000000	.	.	.	.
	UN (2,1)	-.007084[b]	.000000	.	.	.	.
	UN (2,2)	1.435421[b]	.000000	.	.	.	.
	UN (3,1)	-.405352[b]	.000000	.	.	.	.
	UN (3,2)	.313707[b]	.000000	.	.	.	.
	UN (3,3)	2.033792	13.187888	.154	.877	6.148789E-6	672703.348474
	UN (4,1)	.014292[b]	.000000	.	.	.	.
	UN (4,2)	-.022956[b]	.000000	.	.	.	.
	UN (4,3)	.388371[b]	.000000	.	.	.	.
	UN (4,4)	1.458509[b]	.000000	.	.	.	.

a. Dependent Variable: spelling_16_trad.

b. This covariance parameter is redundant. The test statistic and confidence interval cannot be computed.

Residual analysis for test sixteen. The residuals of the all models were tested with the Kolmogorov-Smirnov (K-S) test for normality in SPSS which tests whether the results of the residuals were normally distributed or non-normally normality. Only the null model and last model the between- group model had a normal distribution. These results are displayed in Table 39 and 40. The between-group model and the null model should be used in reporting results because they are the only model not to violate the assumption of MLM, that the residuals are normally distributed. The visual inspection of the histograms present in Figure 19 and 20, as well as the results from the K-S test confirmed that the residuals of the null model and the between-group model are normally distributed.

Table 39

Hypothesis Test Summary for the Between-Group Residuals

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of the Null Model Residuals is normal with the mean 0.027 and standard deviation 2.306	One Sample Kolmogorov-Smirnov Test	.200[1a]	Retain the null hypothesis.

[1]Lilliefors Corrected
[a]The significance level is .05.

Table 40

Hypothesis Test Summary for the Between-Group Residuals

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of the Between-Group Residuals is normal with the mean 0.000 and standard deviation 1.38	One Sample Kolmogorov-Smirnov Test	.099[1a]	Retain the null hypothesis.

[1]Lilliefors Corrected
[a]The significance level is .05.

Qualitative Spelling Analysis for Students' Written Spellings

Qualitative data included all 159 students' written spellings that students produced for the WJ ACH IV, Test 3: Spelling and Test 16: Spelling of Sounds at both pre-test and post-test. In order to compare the two different scoring methods side-by-side growth was converted into a percentage gain score.

Scoring Evaluation for Test Three

Table 41 and 42 (located in the Appendix C) provides the results of scoring all 159 student's written spellings from pre-test to post-test with two different scoring methods: traditional scoring (correct or incorrect) and non-traditional scoring (correct letter sequences). In Table 36 (located in the appendix) the results of the traditional and non-traditional scoring percent change from pre-test to post-test for Test 3: Spelling is provided. For traditional scoring, 77 of 159 of (48.43%) students showed a larger percentage again score from pre-test to post-test with traditional scoring. For non-traditional scoring, 81 out of 159 of (50.94%) students showed a larger percentage increase from pre-test to post-test with non-traditional scoring for Test 3. Then one

additional student showed a gain percentage score of zero for both traditional and non-traditional scoring because the student scored exactly the same at pre-test as they did at post-test. Regarding Test 3: Spelling for documenting the larger percentage of growth, non-traditional scoring worked well overall for the majority of the 159 students (50.94%). However, non-traditional scoring is not the most sensitive scoring measurement for all of the particpants as noted by the 48.43% of students that showed more growth with traditional scoring.

Scoring Evaluation for Test Sixteen

Table 43 and 44 (located in the Appendix C) provides the results of scoring all 159 student's written spellings from pre-test to post-test with two different scoring methods: traditional scoring (correct or incorrect) and non-traditional scoring (correct letter sequences). In summary, 94 out of 159 (59.12%) of students showed a larger percentage gain score from pre-test to post-test with traditional scoring for Test 16. For non-traditional scoring 65 of 159 (40.88%) of students showed a larger percentage increase from pre-test to post-test with non-traditional scoring. Thus, non-traditional scoring did not work as well at documenting growth for Test 16. Even though non-traditional scoring was a more sensitive scoring measurement for 40.88% of the 159 students, this was not true for all students.

CHAPTER 5: DISCUSSION

The purpose of this study was three-fold. The first purpose was to determine the effects of MVRC on the spelling development of second-grade students from two schools in one district in the Southwest. The second purpose was to determine if spelling achievement trajectories varied by gender, ELL classification, and/or SPED classification. The third purpose was to score the student's spelling tests with two different metrics of spelling: tests that were scored with traditional standardized scoring (correct and incorrect) and tests that were scored with Curriculum-Based Measurement (CBM) scoring for spelling (correct letter sequences) and to decide which was a more sensitive to growth from pre-test to post-test. This chapter will present the summary of findings for each research question, limitations, recommendations and implications for future research.

Summary of Findings for Each Research Question

Summary of Findings for Research Question One

Research question one was "Does MVRC have a significant effect on students' spelling scores from the WJ IV ACH for Test 3: Spelling, and Test 16: Spelling of Sounds over time?" Using multilevel models for repeated measures the researcher estimated the results for Test 3: Spelling with the between-group model found that the time by treatment was significant (p=.036, p<.05). On average, students who participated in MVRC increased their spelling score on Test 3: Spelling by 1.786 words. Using multilevel models for repeated measures the researcher estimated the results of Test 16: Spelling of Sounds with the between-group model found that the time by treatment was

significant (p=.000, p<.05). Students who participated in MVRC increased their spelling score on Test 16: Spelling of Sounds by an average of 1.741 words.

Summary of Findings for Research Question Two

Research question two was "Do spelling achievement trajectories vary by gender, ELL classification, and/or SPED classification for the WJ IV ACH, Test 3: Spelling, and Test 16: Spelling of Sounds?" Using multilevel models for repeated measures the researcher estimated the how the spelling trajectories varied by gender, ELL classification, or SPED classification. However, none of these student level predictors were significant for average spelling achievement (p<.05). Females demonstrated more of an increase than males over time. Students in special education demonstrated less growth over time than students not in special education. Students who were classified as ELLs scored more words than ELL students.

The between-groups model for Test 16: Spelling of Sounds estimated how the spelling trajectories varied by gender, ELL classification, or SPED classification, however, none of these student level predictors were significant of average spelling achievement (p<.05). Females demonstrated less of an increase than males over time. Students in the special education demonstrated more growth over time than students not in special education. Students who were classified as ELLs scored less words than ELL students.

Summary of Findings for Research Question Three

Research question three was " Which scoring measurement is more sensitive to student growth: traditional standardized scoring (correct or incorrect) or CBM spelling scoring (which credits each correct letter sequence) from the WJ IV ACH Test 3:

Spelling, and Test 16: Spelling of Sounds?" Qualitative analysis of the 159 students' written spellings from Test 3: Spelling and Test 16: Spelling of Sounds conducted by the author did not reveal a clear answer. For the 159 students, it seems that non-traditional scoring, scoring by correct letter sequences, was more representative of growth for Test 3: Spelling because 50.94% of 159 students had a larger percentage increase from pre-test to post-test. However, for Test 16: Spelling of Sounds, the opposite was found. Traditional scoring was more representative of student's growth pre-test to post-test because 59.12% of 159 students had a larger percentage increase. These results do not allow for any clear distinction of which scoring method is better than the other. Many times, students had an equal number of the amount of total words they had spelled correctly (i.e., 7 out of 10), but when their responses were evaluated with CBM-spelling one could clearly distinguish differences among students in spelling growth.

Limitations

Although the current study contributes to the existing knowledge of computer-based reading interventions and the comparisons of standardized scoring and CBM-spelling scoring, several limitations should be considered. The limitations to be discussed are as follows: test administration, success for all, planned activity checks and inter-observer agreement, additional data, sample size, and outcome measures.

Test Administration

First, the measures of spelling were only conducted at pre-test and post-test. This limited the multilevel model's ability to provide how spelling development differed in growth during the middle of the school year and did not allow for the use of a parameter of quadratic time (0, 1, 4). If a study like this were to be conducted again, the author

would recommend a pre-test, a mid-test, and a post-test. Additionally, a number of students' tests could not be used in this study because adequate ceiling rules were not observed for the WJ IV ACH tests. In other words, several of the students may have continued to get correct responses if more items had been given. This problem could have been addressed by administering additional items to all students to ensure everyone had reached a ceiling.

Success for All

Both the experimental and comparison groups received Success for All language arts instruction. Success for All is already an intervention that includes 90-minute blocks of daily reading which are highly intensive and also include guided practice, phonics instruction, and ongoing formative evaluation (Slavin & Madden, 2001). Therefore, the experimental group should be seen as Success for All plus MindPlay Virtual Reading, and the comparison group was Success for All plus supplemental class instruction.

Planned Activity Checks and Inter-Observer Agreement (IOA)

The results of lower levels of inter-observer agreement for the planned activities check for student engagement behavior make the author question the validity of the planned activities check. The additional observers needed to be trained to record engagement behavior in the same way as the primary observer. Since this study was completed with an archival data set, IOA could not be further investigated for student engagement.

Additional Data

As this study was completed with a limited data set, school level predictors of SES and free and reduced lunch as well as other characteristics were not available to the author. These factors could have allowed the author to explore school at Level 4 for the multilevel model. Moreover, only two schools and four teachers from each school were in this study. Also, no predictors were provided at the specific teacher level such as teacher effectiveness of implementing SFA which might have been important for consideration in this study.

Sample Size

A larger sample size would be necessary to generalize results to other similar school environments. Also the smaller sample size may have contributed to not finding significant effects for the student demographic variables.

Outcome Measures

The current outcome measures of spelling from the WJ IV ACH (Test 3: Spelling and Test 16: Spelling of Sounds) only included narrowly focused tasks of spelling real words and non-words. Additionally, testing which included the application of spelling skills in written expression would have further added to this study.

Recommendations and Implications for Future Research

This study adds to the limited body of research that supports the use of computer-based instruction and online instruction for enhancing student spelling performance. With increased pressures to ensure that all students are making adequate yearly progress and that special education students are getting the intervention and support they need, teachers do not have the adequate amount of time available to make certain that all of

things will happen or to be able to provide one-to-one individualized instruction to the students that really need it. A range of computer-based and online base programs have been created to increase performance in spelling as discussed in chapter two (e.g. Comaskey et al., 2009; Daal & Reitsma, 2000; Di Stasio et al., 2012; Fasting & Lyster, 2005; Howell et al., 2000; Savage et al., 2010). However, compared to the literature for computer-based and online based programs for the outcomes of reading achievement the literature of these similar programs for spelling is limited. As such computer-based instruction and online instruction can serve as a possible solution to increasing students' skills in many areas, particularly in spelling, because not all teachers have the time to actually teach spelling skills directly.

Although both groups made gains in spelling, the experimental group outperformed the control group. One novel aspect of this study was the use of standardized scoring and CBM scoring to evaluate growth. Although no differences existed between these two different scoring approaches, one indicated more growth for Test 3: Spelling, while the other showed indicated more growth for Test 16: Spelling of Sounds. CBM scoring was particularly useful for examining the spelling growth of individual students and for accurately assessing the spelling of a student's ability at pre-test. Because of this, CBM spelling seemed to provide a more accurate representation of a student's spelling ability in post-test, thus showing a lesser percentage gain increase than standardized scoring. Taking all of this into consideration, teachers could use CBM spelling to score students' spelling if they want a more accurate presentation of a student's abilities at the beginning. A teacher for example, may examine a student's spelling of the word "apple" spelled as "aple". With traditional scoring the teacher would

mark this student as correct. With CBM-spelling scoring, the word apple has five correct letter sequences and the student is earning four of those five correct letter sequences. Upon further examination with CBM-spelling, the teacher would notice that the student was only missing one correct letter sequence in their spelling of apple because they did not double the consonant "t" (^a^pl^e^). This information can be very valuable in planning for instruction and for working with students who make spelling errors.

Overall, scoring spelling traditionally still provides a quick assessment of how the students can or cannot spell an entire word while CBM spelling gives more information beyond scoring as correct or in correct. Because CBM spelling scoring takes so much time, it is not recommended for use in scoring spelling of a large number of students, like this study. CBM-spelling scoring would be recommended for scoring a standardized test such as the WJ IV ACH to find out what particular spelling errors the student is making.

Future Research

Future research could include having MindPlay Virtual Reading as the main intervention for students without Success for All. Due to the constraints, during the school year most schools are already using an additional intervention or structured language arts program, this would be recommended to be employed in the summer with students.

Also, a similar study could be conducted which included older students or a few grade levels of students within one study (e.g., 2nd grade, 3rd grade, 4th grade). Additionally, the same study could be conducted with three testing points throughout the year and also have students spend a different amount of time on MindPlay Virtual Reading. For example, one treatment group could spend 15 minutes a day on MVRC,

while another group could spend 30 minutes a day, and another group could spend one hour a day on MVRC. Then the results of the different treatment groups could be compared. Additionaly, a study on spelling achievement comparing special education and general education students would be of further interest. Finally, researchers should investigate the effects of spelling achievement among special education students of different classifications (e.g., Specific Learning Disabilities, Autism, Speech/Language Impairment).

APPENDIX A-FIGURES

Figure 1. The three levels of the multilevel model of repeated measures.

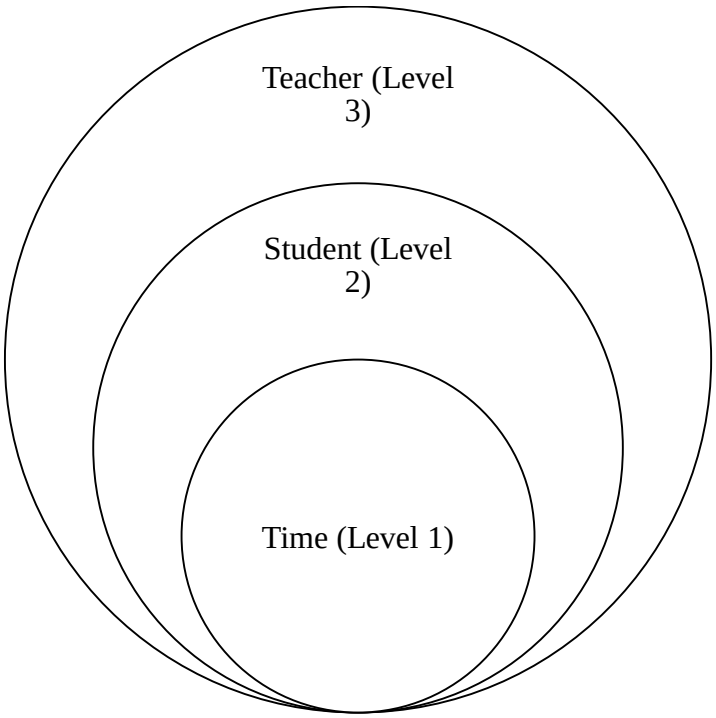


Figure 2. Test 3 spelling growth trajectories for all students with traditional scoring.

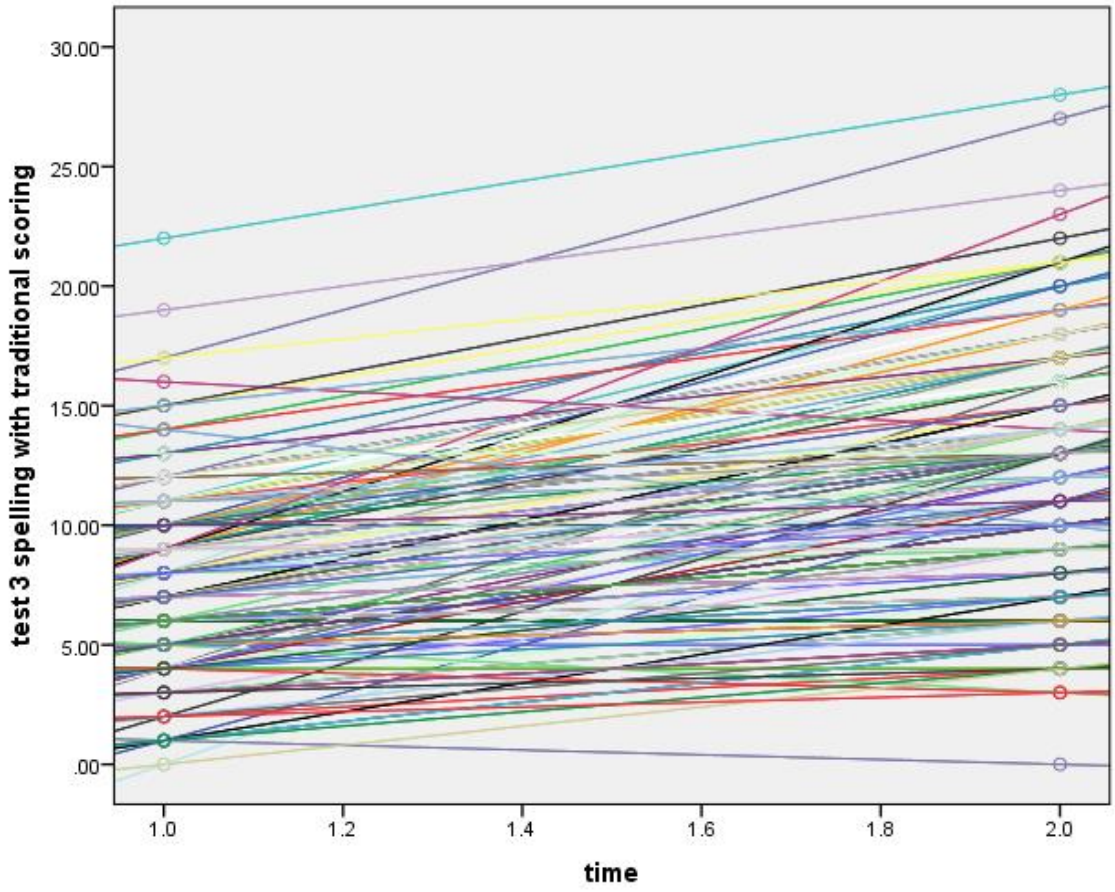


Figure 3. Test 3 spelling growth trajectories for all students with non-traditional scoring.

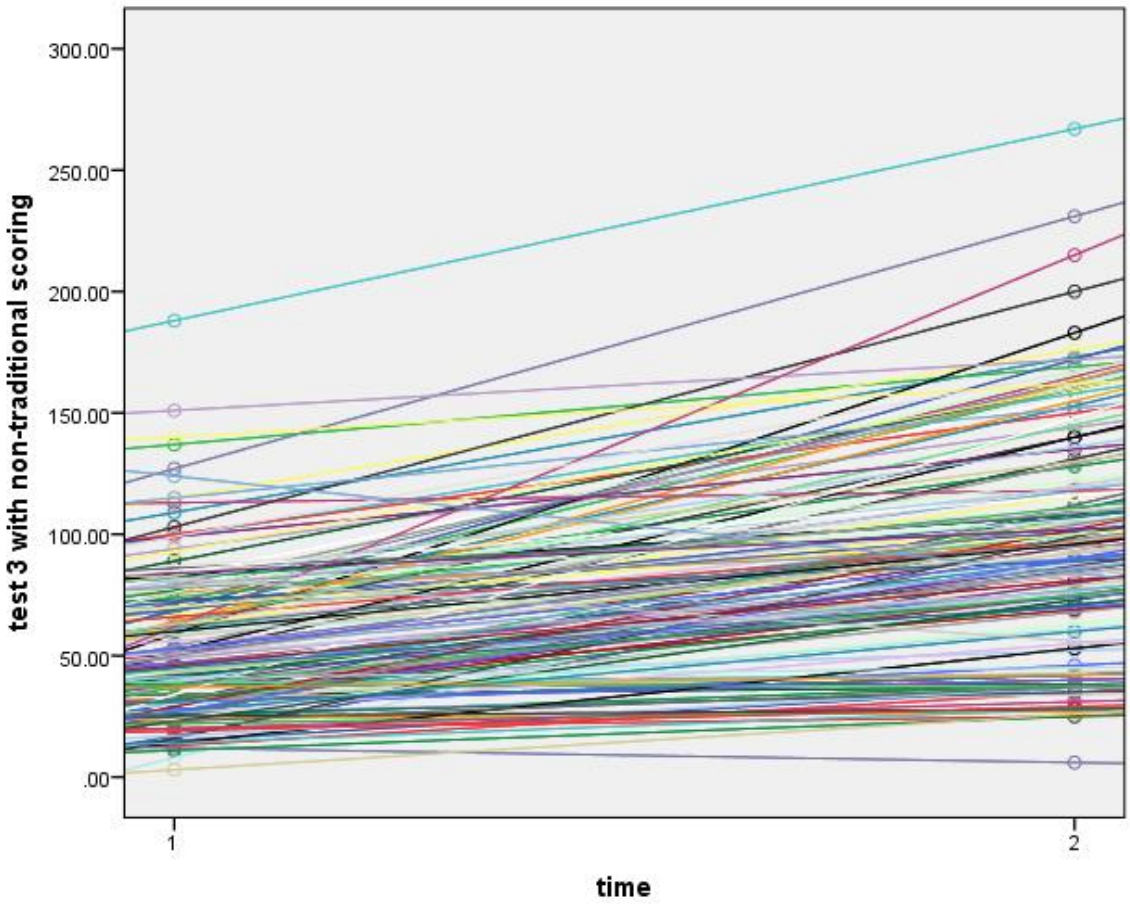


Figure 4. Test 16 spelling growth trajectories for all students with traditional scoring.

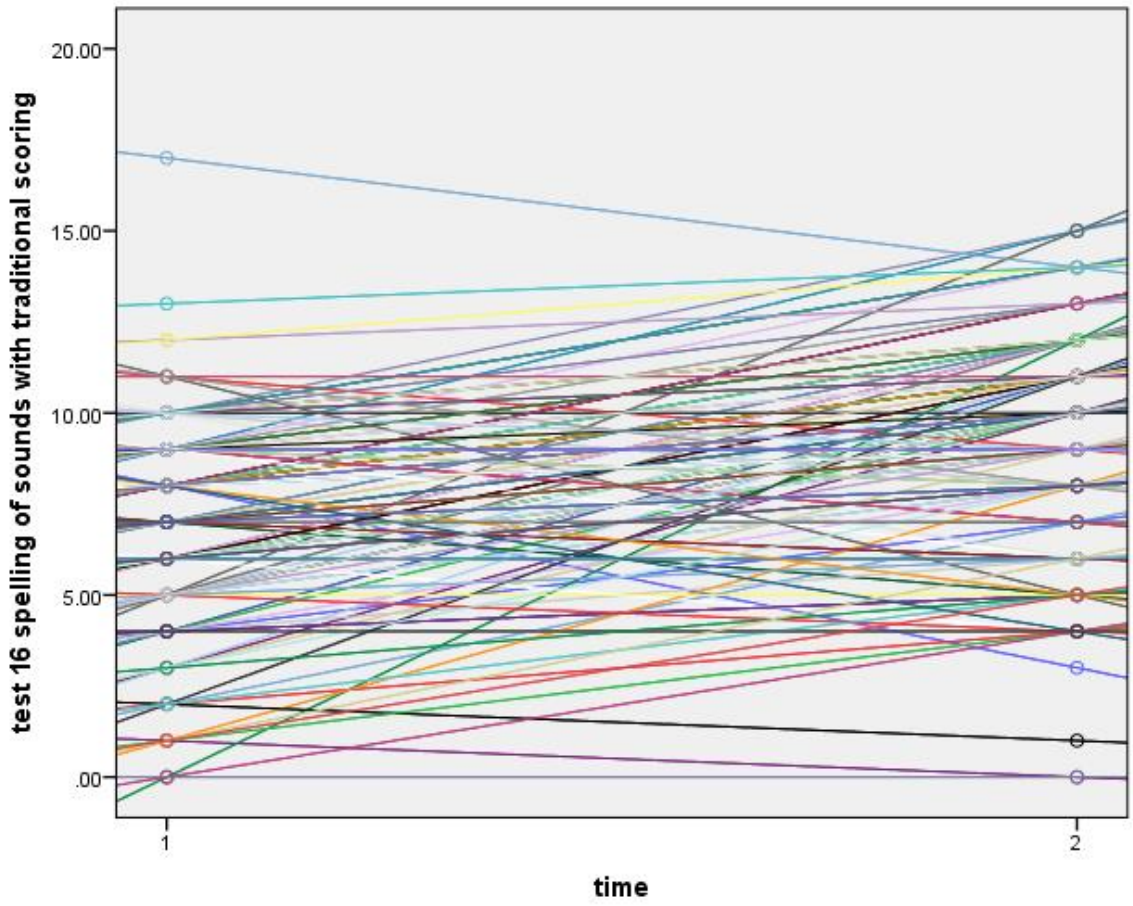


Figure 5. Test 16 spelling growth trajectories for all students with non-traditional scoring.

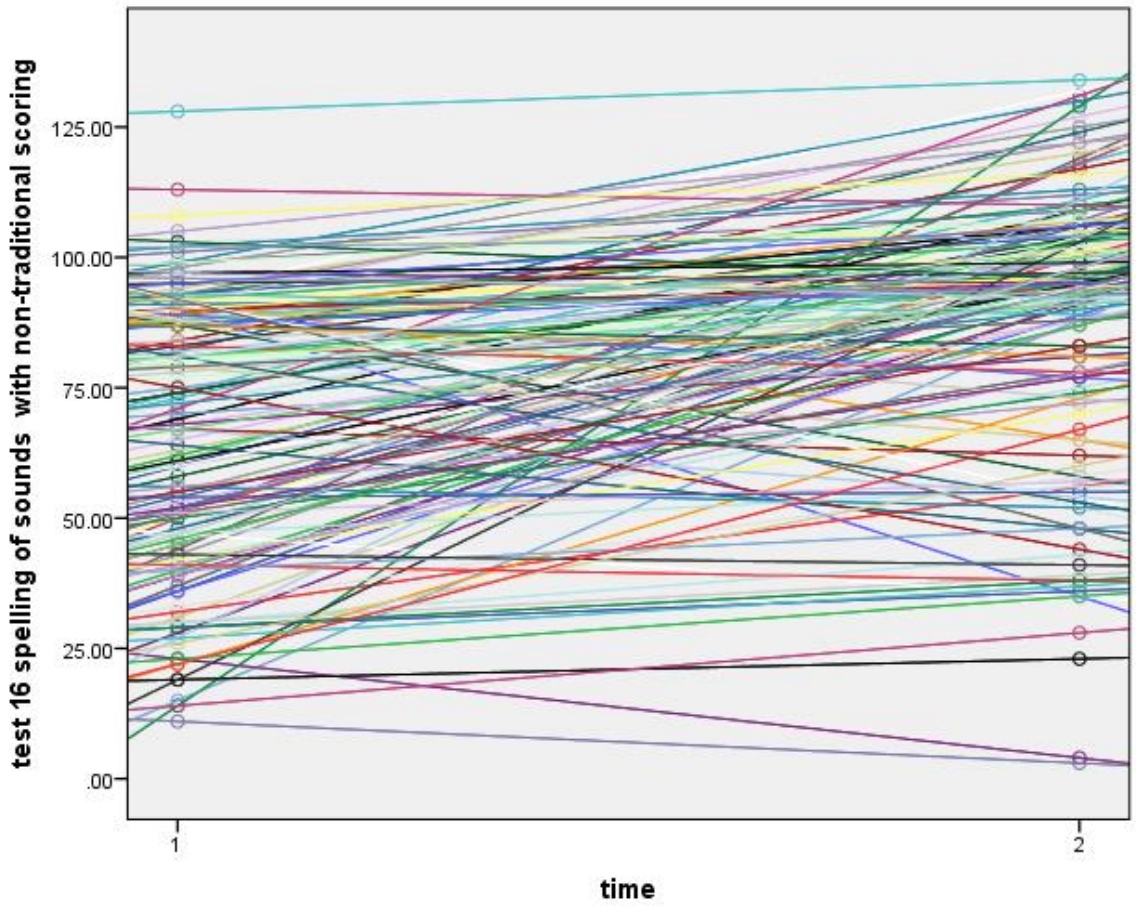


Figure 6. Mean of comparison and experimental groups for test 3 with traditional scoring.

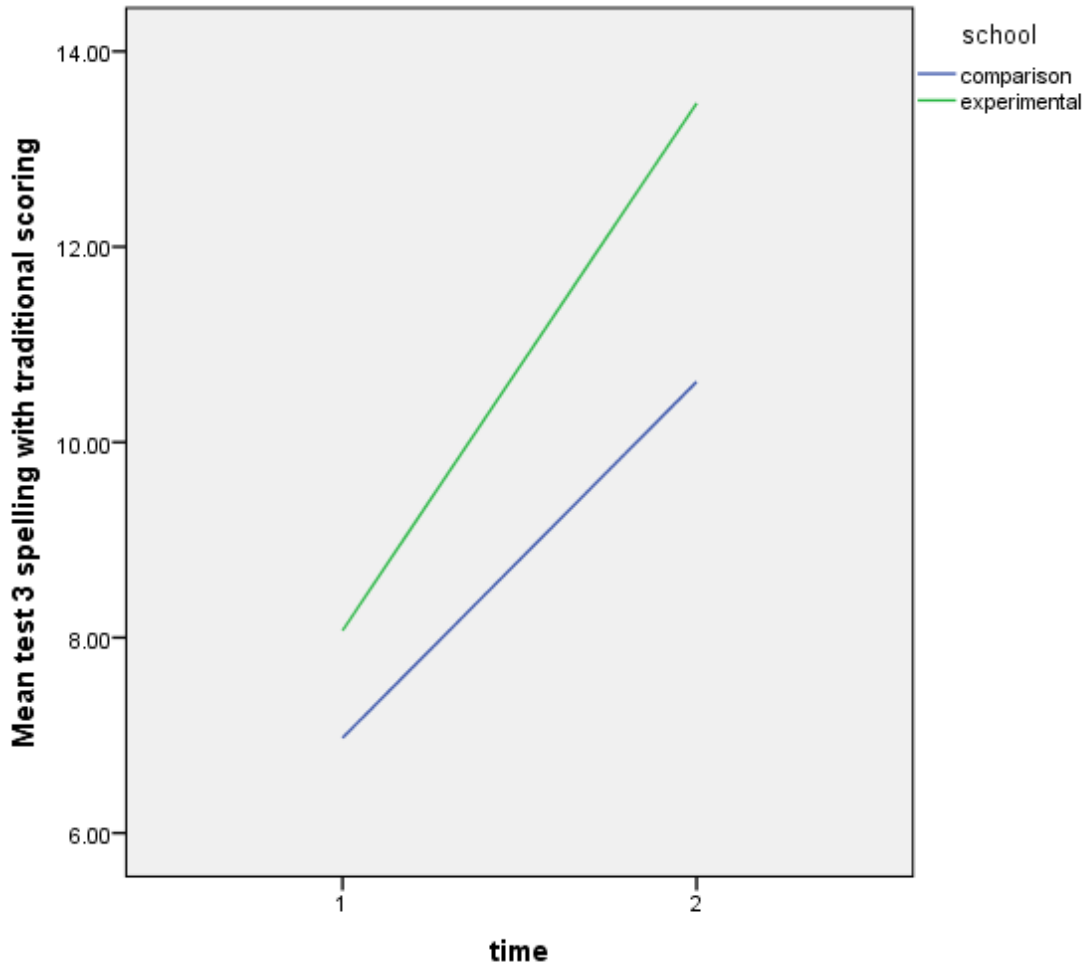


Figure 7. Mean of comparison and experimental groups for test 3 with non-traditional scoring.

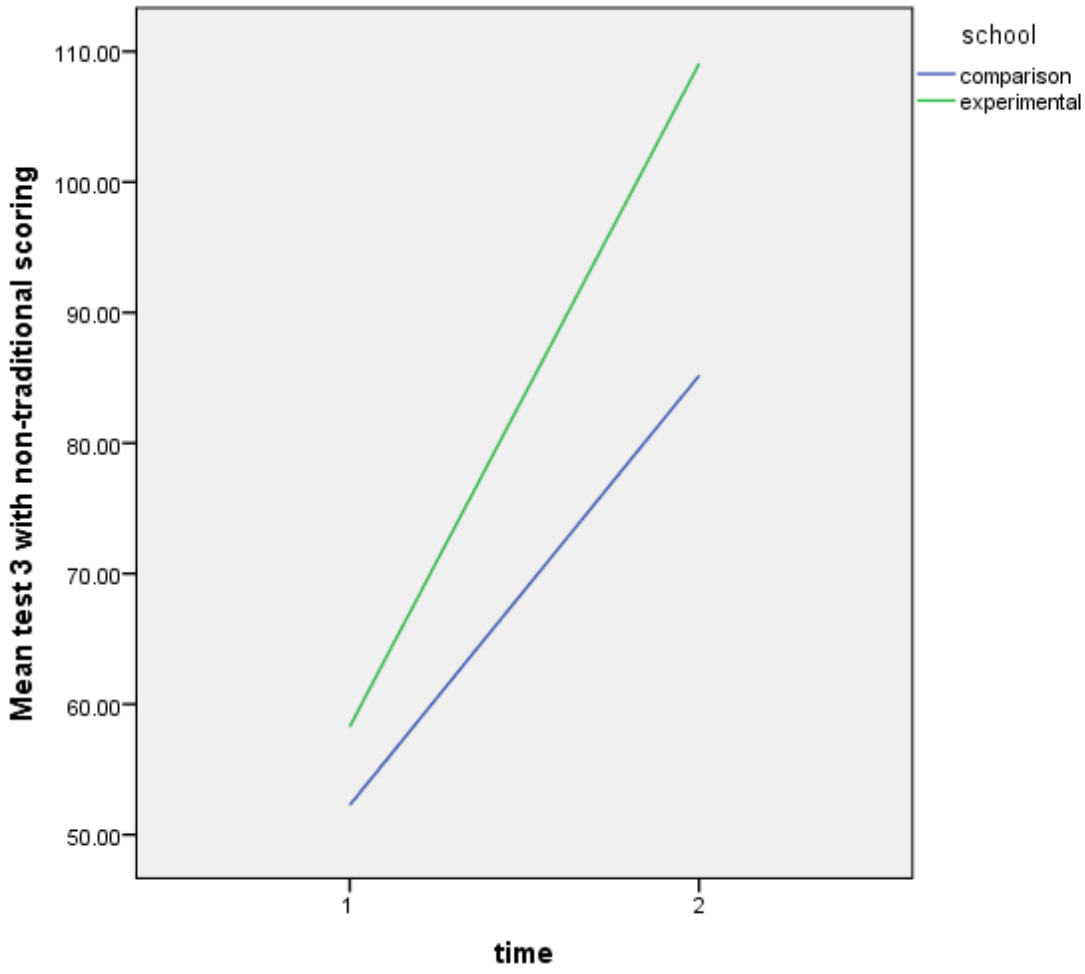


Figure 8. Mean of comparison and experimental groups for test 16 with traditional scoring.

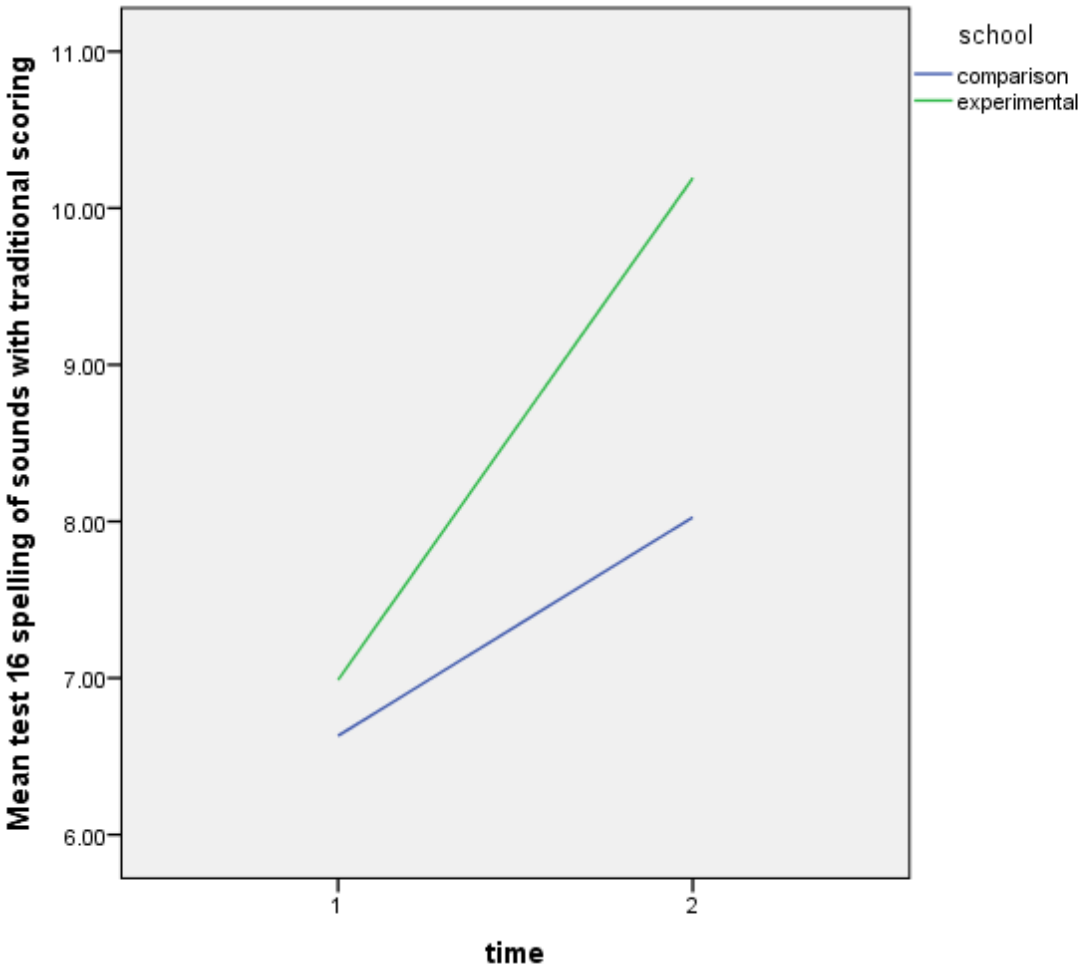


Figure 9. Mean of comparison and experimental groups for test 16 with non-traditional scoring.

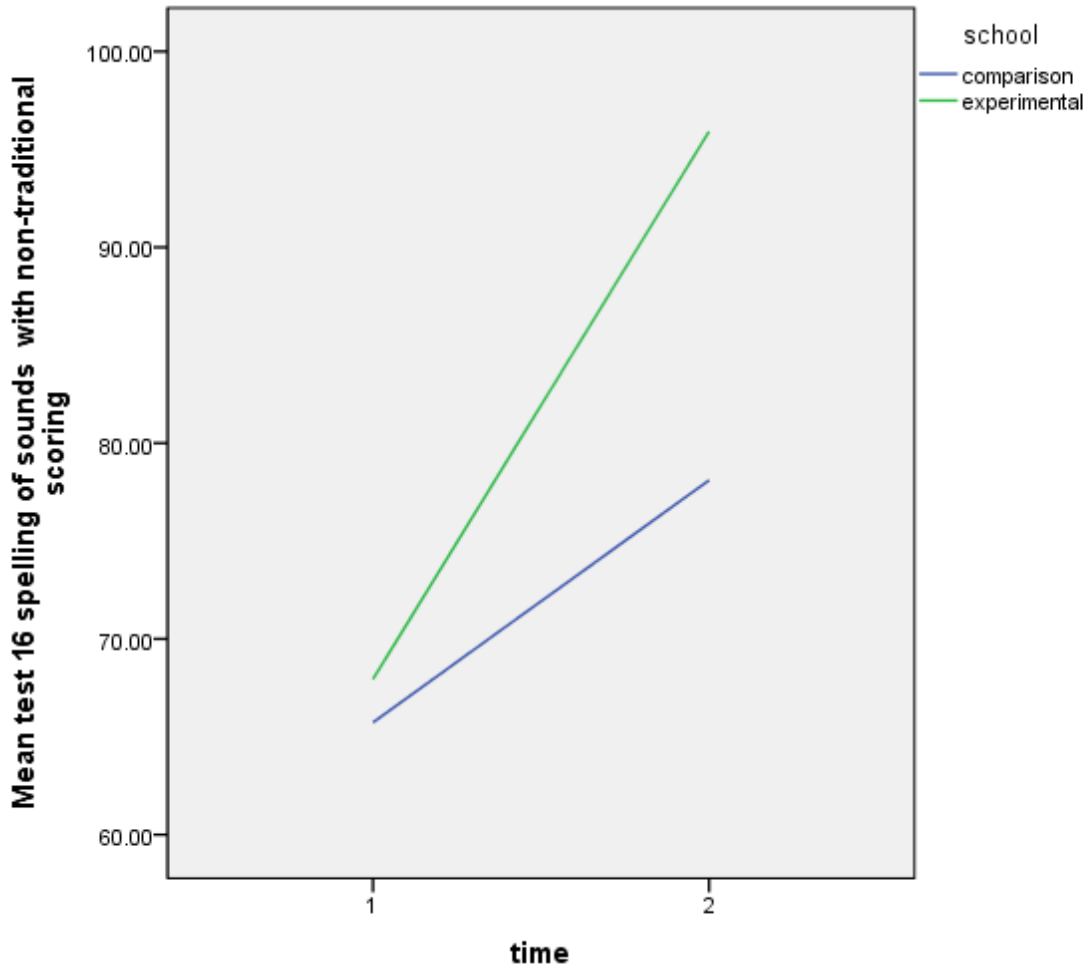


Figure 10. Subset of individual spelling growth trajectories for test 3 with traditional scoring.

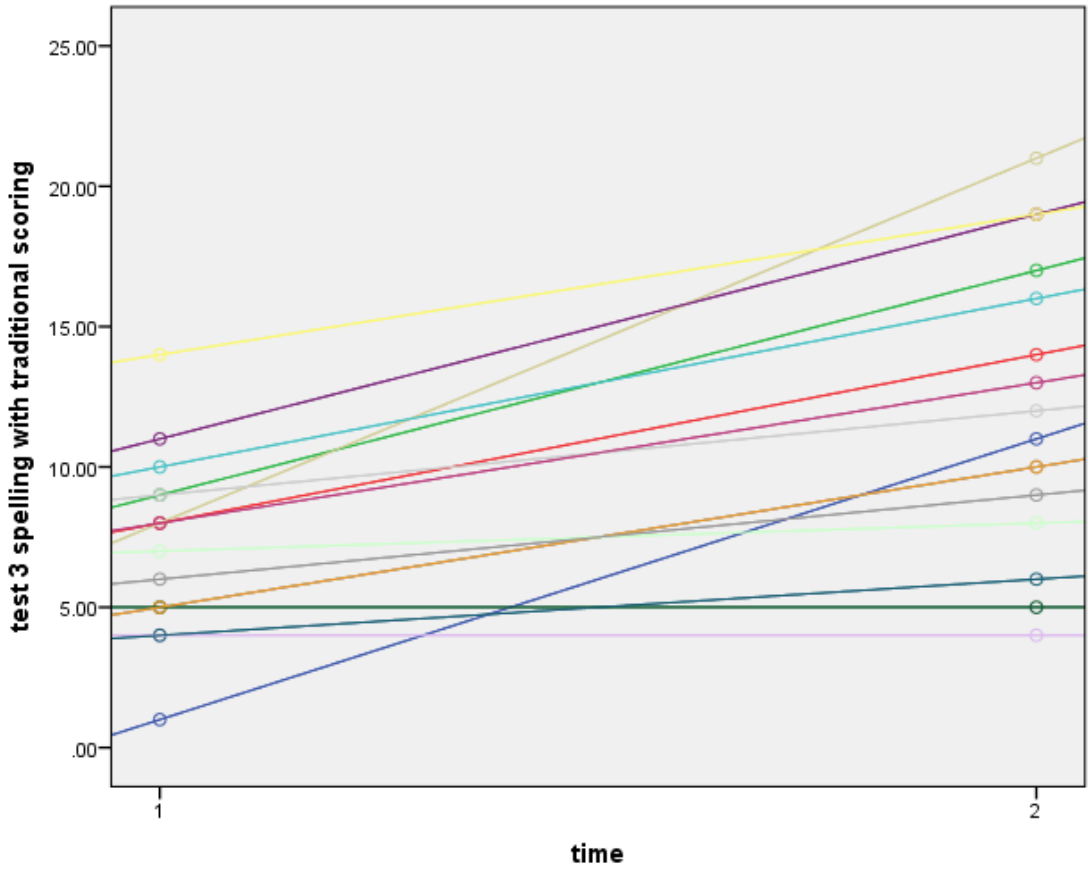


Figure 11. Subset of individual spelling growth trajectories for test 3 with non-traditional scoring.

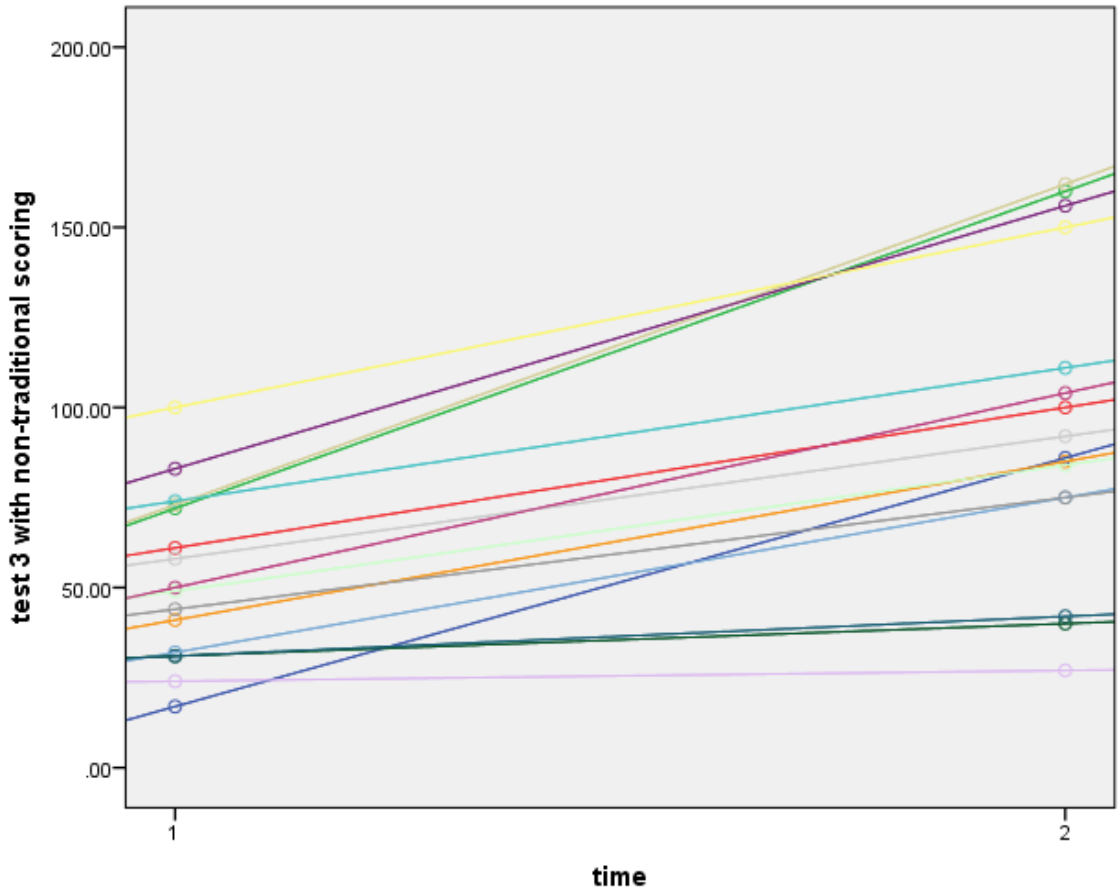


Figure 12. Subset of individual spelling growth trajectories for test 16 with traditional scoring.

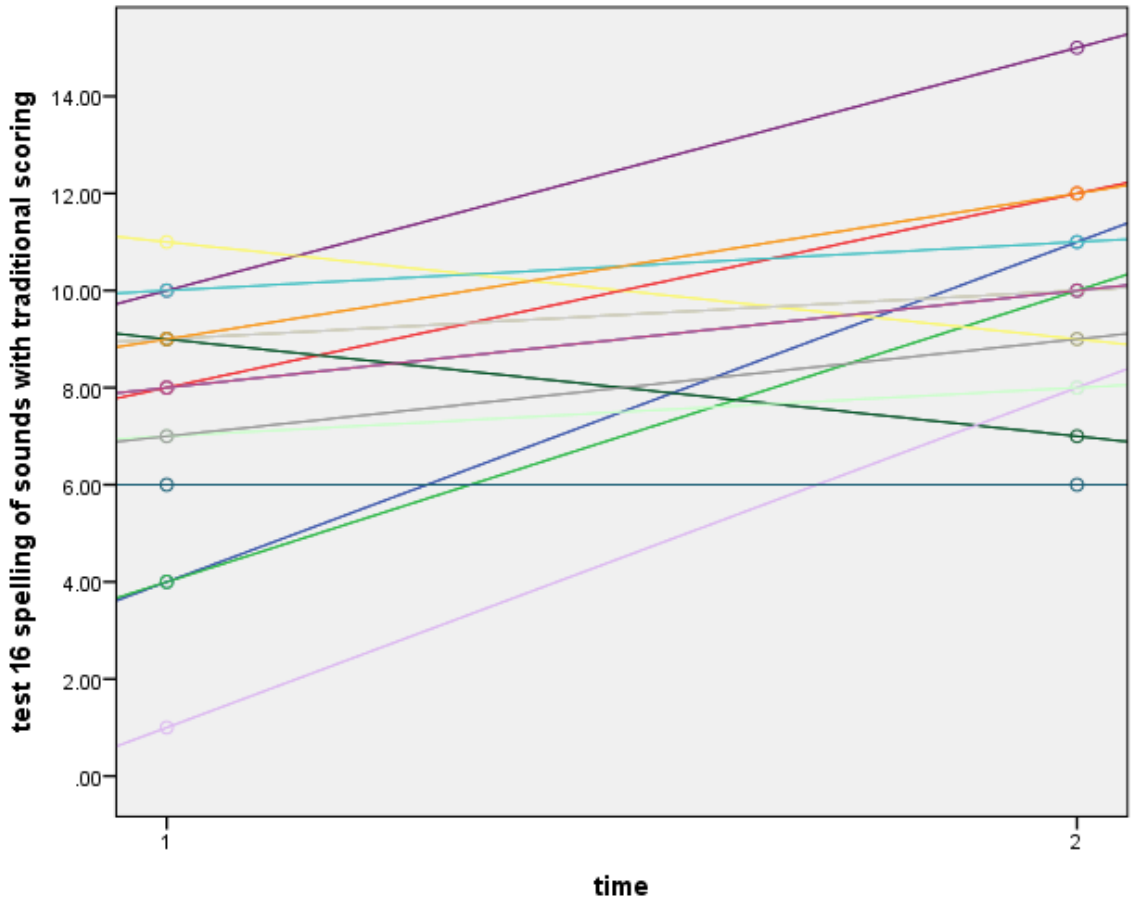


Figure 13. Subset of individual spelling growth trajectories for test 16 with non-traditional scoring.

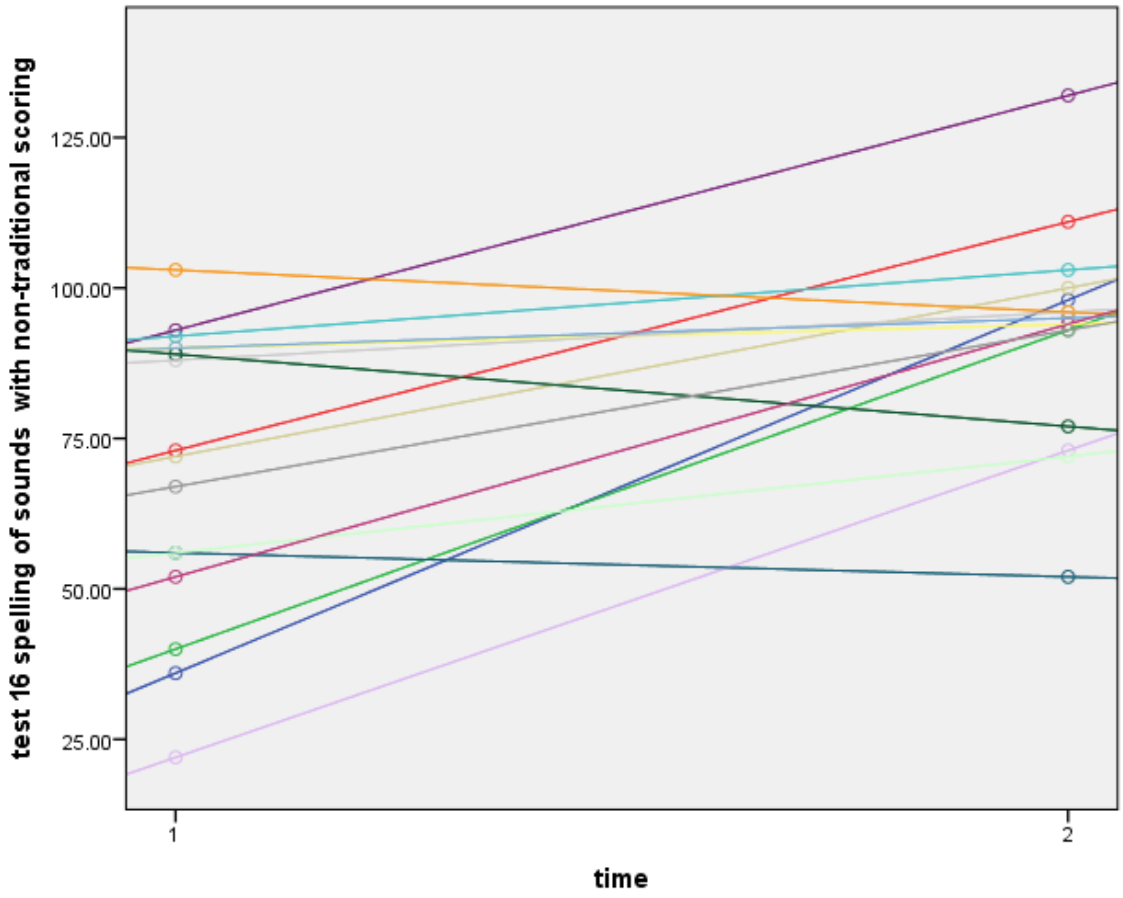


Figure 14. Mean of students by teacher for test 3 with traditional scoring.

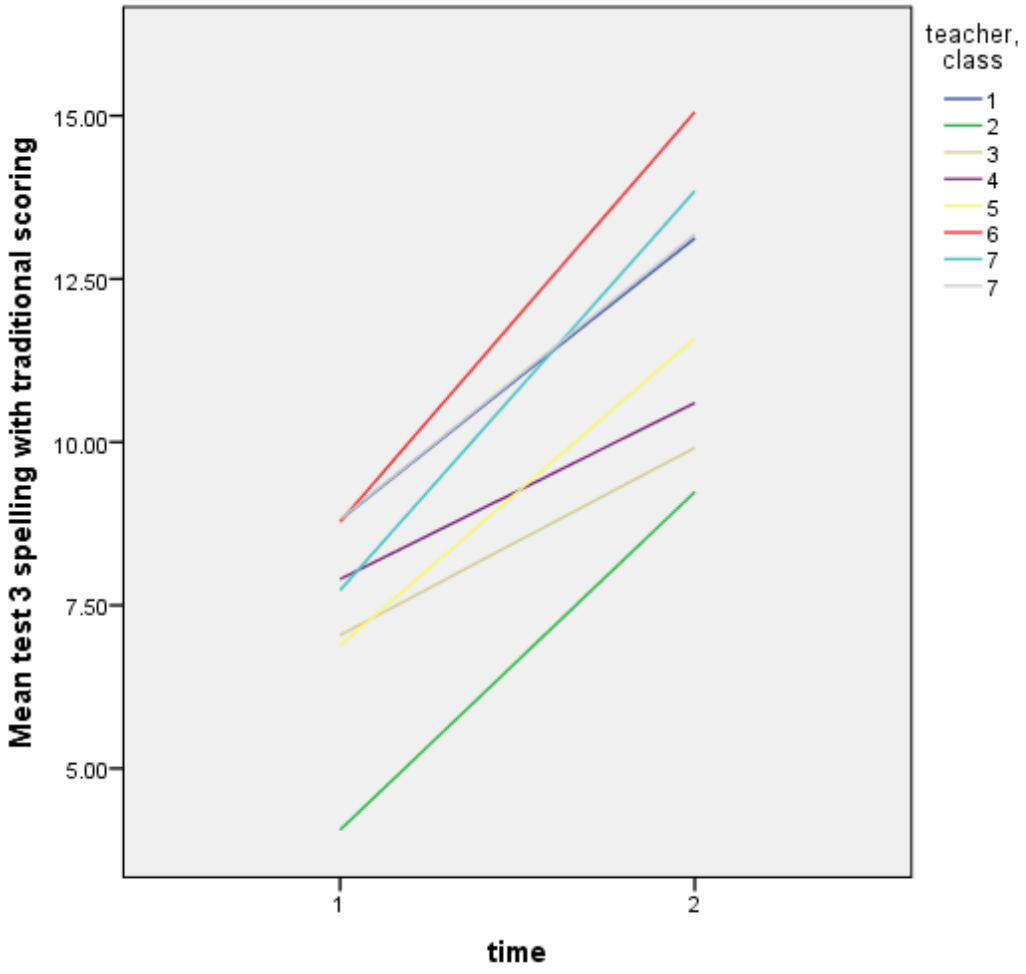


Figure 15. Mean of students by teacher for test 3 with non-traditional scoring.

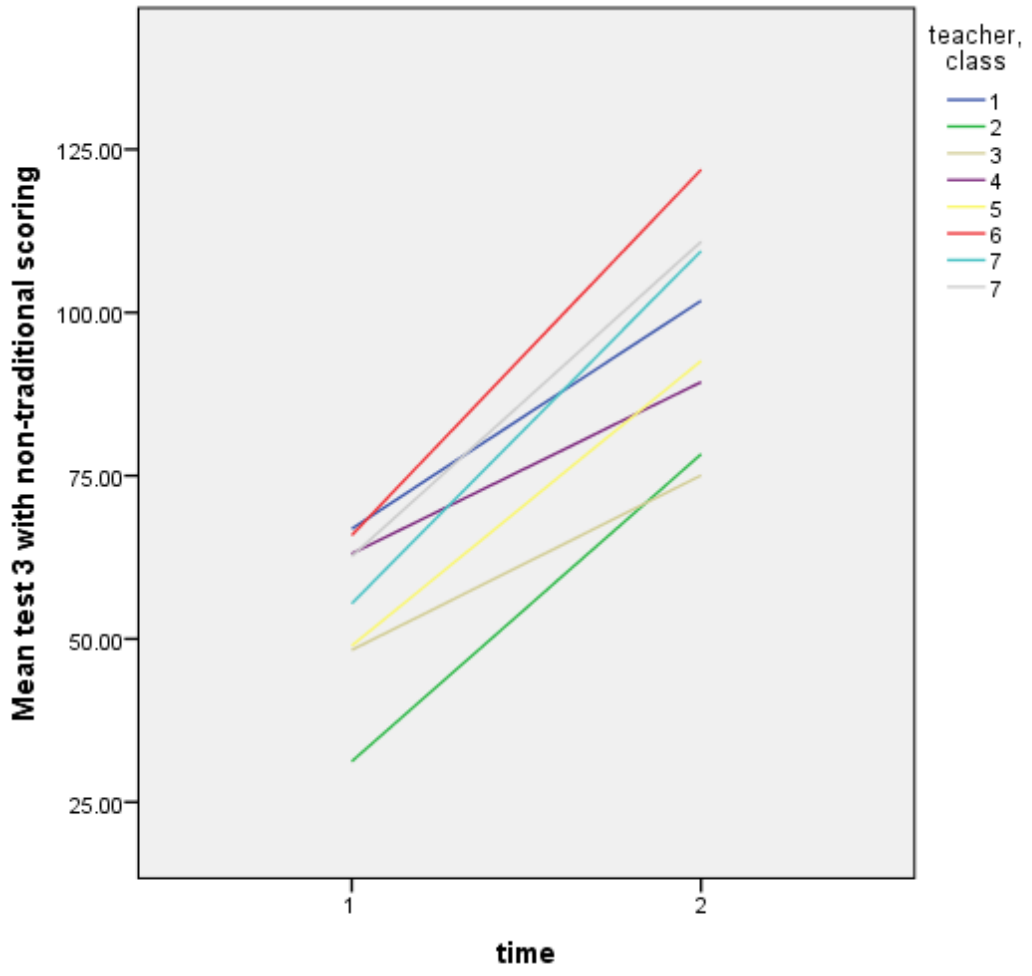


Figure 16. Mean of students by teacher for test 16 with traditional scoring.

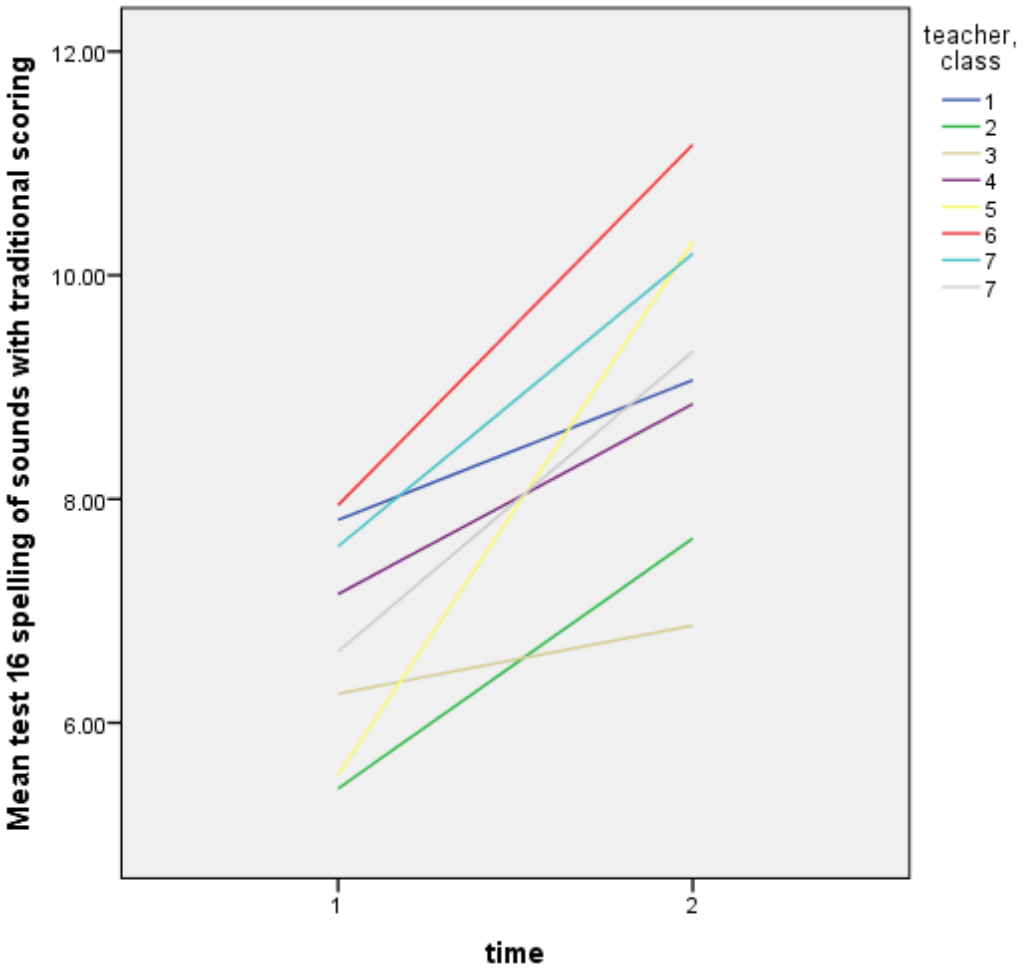


Figure 17. Mean of students by teacher for test 16 with non-traditional scoring.

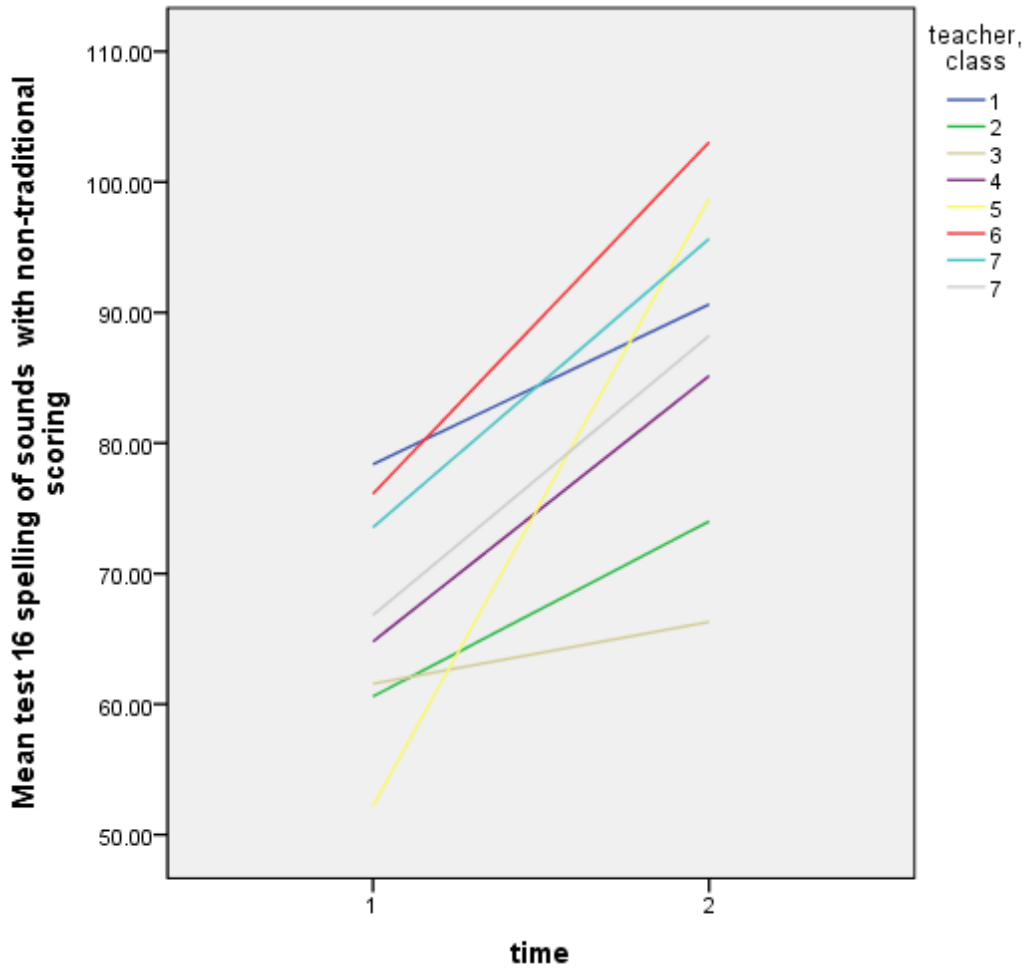


Figure 18. Histogram of the between-group model residuals.

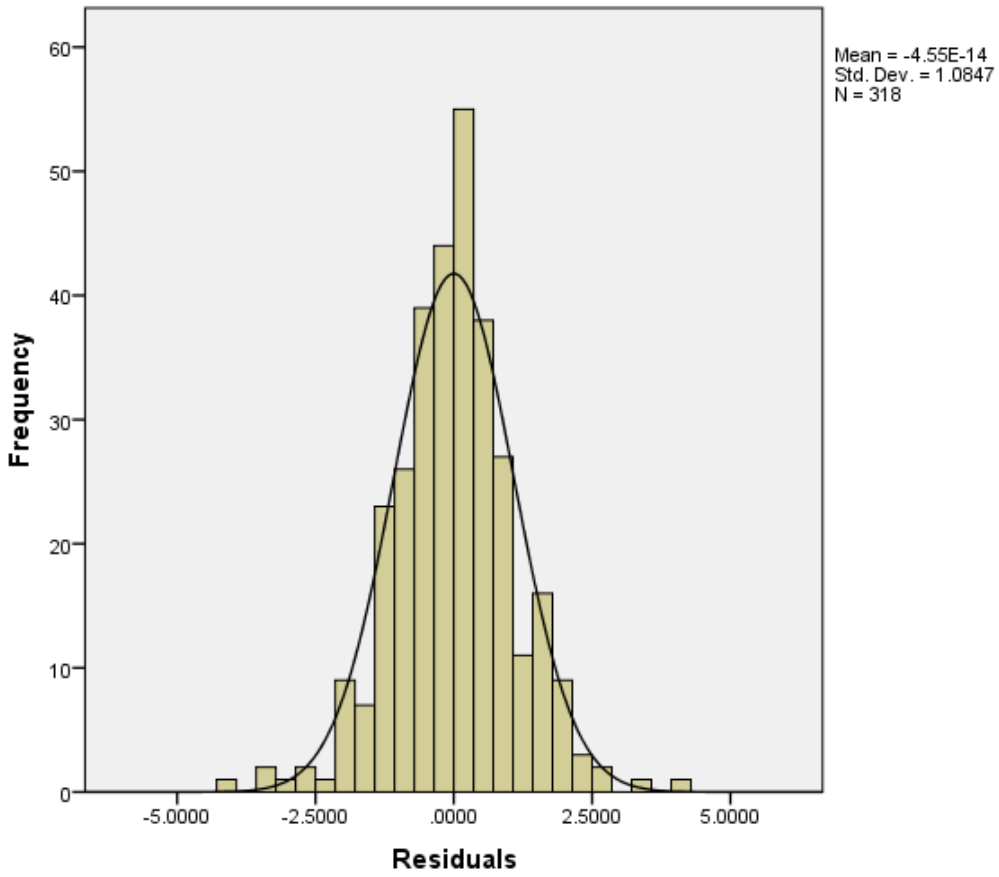


Figure 19. Histogram of the null model residuals.

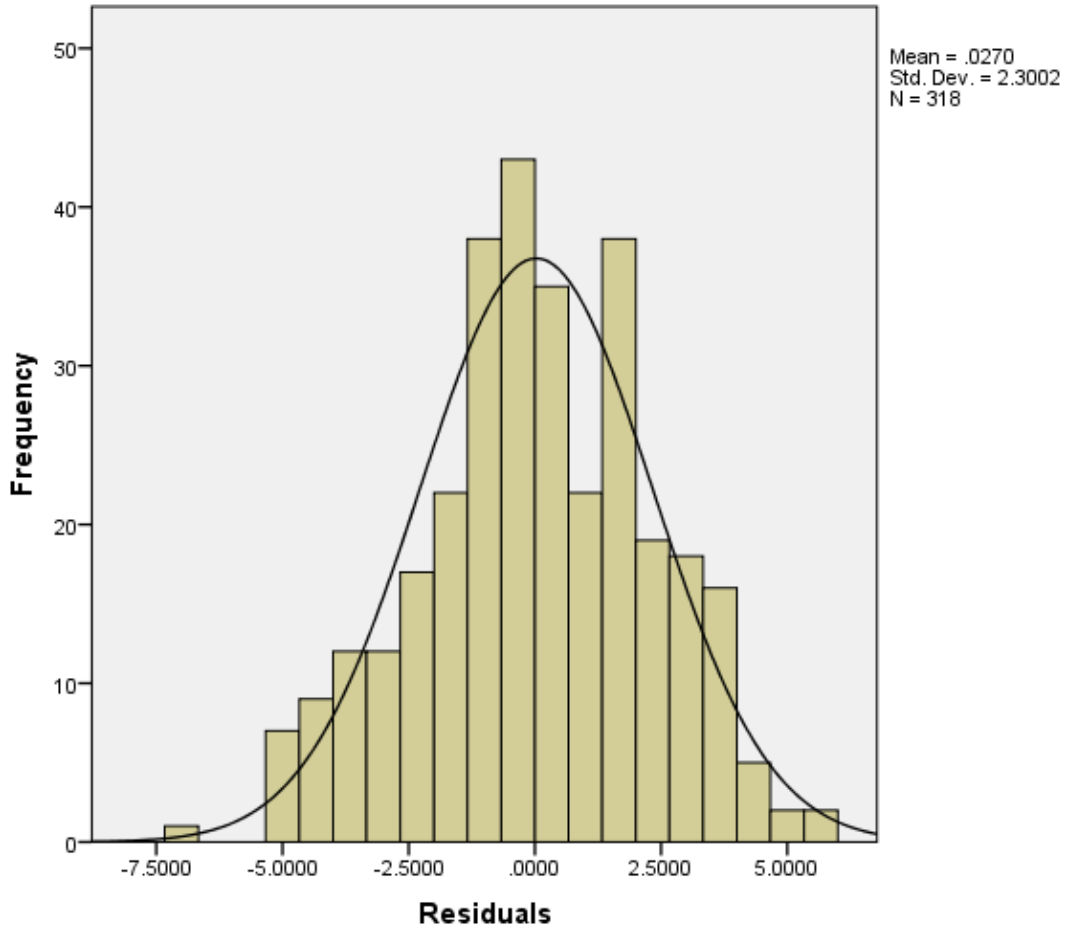
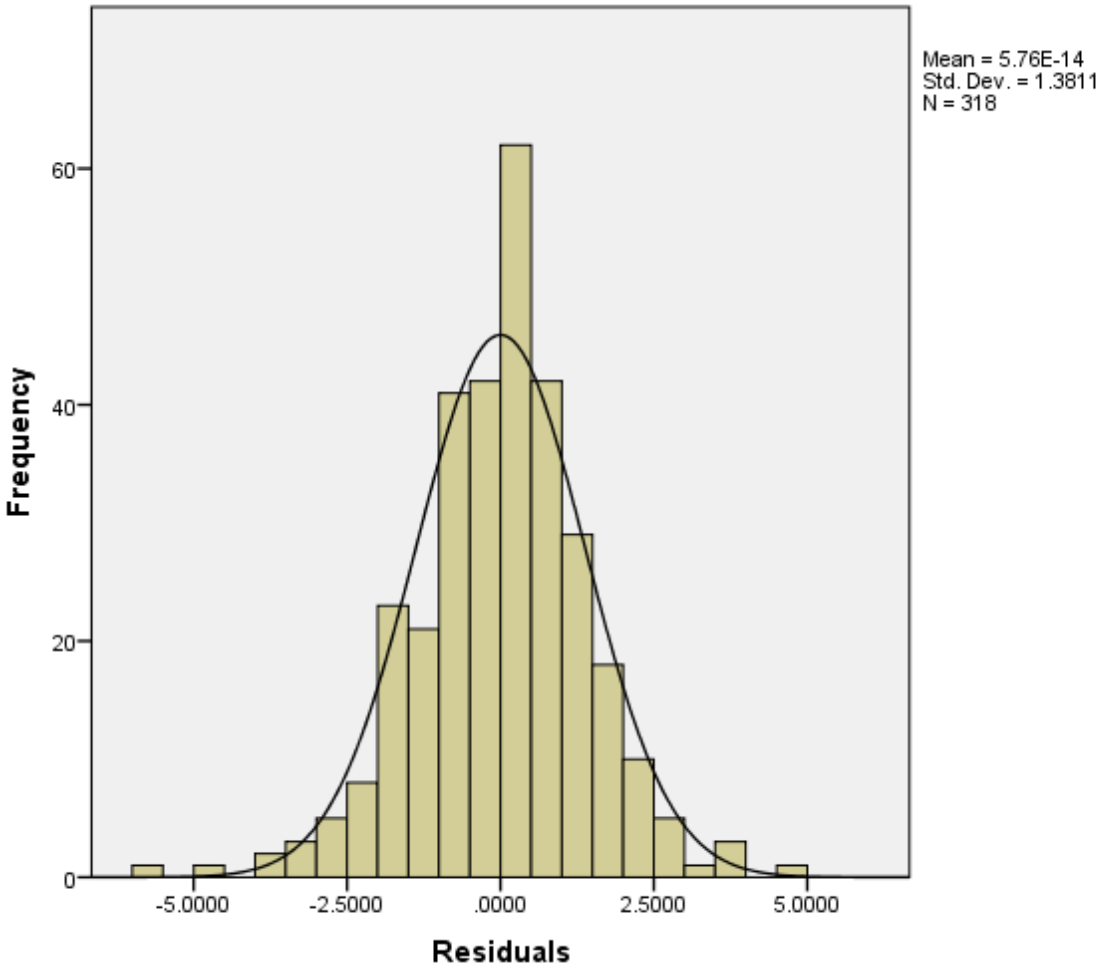


Figure 20. Histogram of the between-group model residuals



APPENDIX B-PLANNED ACTIVITY CHECK FORM

Planned Activity Check (PLA-check) for Lesson Effectiveness

Date: _____ **Site**_____

Start Time: _____ (am/pm) **End Time:** _____ (am/pm) **Length of Observation:** ___ **minutes**

Observer: _____

Directions: Note the total number of students (participants) in the classroom and the time at which the observation begins. At the end of the designated 2 minute interval, quickly scan the classroom from left to right (or right to left; whichever gives the observer a better view) and count the number of students displaying any of the off-task behaviors.

Once you have counted the number of off-task students at the end of the two minute interval, subtract this number from the total number of students in the classroom. Each scan should take a quick 15-30 seconds scan of the room. Continue this process at the end of each designated 2 minute interval for a total of 20 minutes (for 10 intervals).

Off-task Behaviors:

a. Student's headphones are either on only one ear or are sitting on the desk.

b. Student engaged in activities with any other person (outside of instructional interaction).

c. Other software program open on the desktop of the computer other than MVRC.

Number of participants

Time (e.g., Time 1=T1)	**Off-Task**	**On-Task**
T 1 _____	_____	_____
T 2 _____	_____	_____
T 3 _____	_____	_____
T 4 _____	_____	_____
T 5 _____	_____	_____
T 6 _____	_____	_____
T 7 _____	_____	_____
T 8 _____	_____	_____
T 9 _____	_____	_____
T 1 0 _____	_____	_____
Total	_____	**Total** _____

APPENDIX C-ADDITIONAL TABLES

Table 41

Students' Spellings Scored with Traditional and Non-Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
1	1	7	0.143	17	29	0.586
2	9	18	0.5	72	98	0.735
3	5	12	0.417	34	56	0.607
4	5	15	0.333	40	76	0.526
5	9	16	0.563	67	84	0.798
6	5	11	0.455	32	51	0.627
7	11	22	0.5	99	128	0.773
8	5	14	0.357	37	69	0.536
9	7	14	0.5	41	69	0.594
10	6	12	0.5	34	51	0.667
11	12	19	0.632	76	104	0.731
12	6	13	0.462	39	63	0.619
13	6	12	0.5	31	56	0.554
14	3	10	0.3	27	44	0.614
15	12	19	0.632	72	104	0.692
16	5	11	0.455	31	51	0.608
17	10	22	0.455	83	128	0.648

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
18	8	21	0.381	73	118	0.619
19	11	22	0.5	83	128	0.648
20	12	19	0.632	76	104	0.731
21	8	15	0.533	49	76	0.645
22	7	18	0.389	50	98	0.51
23	7	14	0.5	40	69	0.58
24	7	16	0.438	51	84	0.607
25	4	11	0.364	28	51	0.549
26	5	11	0.455	34	51	0.667
27	13	22	0.591	109	128	0.852
28	12	22	0.545	94	128	0.734
29	4	10	0.4	24	44	0.545
30	11	18	0.611	77	98	0.786
31	9	18	0.5	72	98	0.735
32	14	30	0.467	137	194	0.706
33	6	14	0.429	43	69	0.623
34	5	11	0.455	31	52	0.596

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
35	15	24	0.625	115	145	0.793
36	14	22	0.636	100	128	0.781
37	8	17	0.471	61	90	0.678
38	10	16	0.625	60	84	0.714
39	10	18	0.556	69	97	0.711
40	6	13	0.462	38	63	0.603
41	9	17	0.529	62	90	0.689
42	8	20	0.4	60	84	0.714
43	7	13	0.538	38	63	0.603
44	9	18	0.5	63	98	0.643
45	7	15	0.467	47	76	0.618
46	8	16	0.5	54	84	0.643
47	9	16	0.563	59	84	0.702
48	0	6	0	8	24	0.333
49	5	15	0.333	44	76	0.579
50	17	25	0.68	127	153	0.83
51	6	16	0.375	50	84	0.595

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
52	9	16	0.563	63	84	0.75
53	5	14	0.357	40	69	0.58
54	6	13	0.462	37	63	0.587
55	6	16	0.375	47	84	0.56
56	11	22	0.5	87	128	0.68
57	9	17	0.529	68	90	0.756
58	5	13	0.385	37	63	0.587
59	4	10	0.4	26	44	0.591
60	3	10	0.3	25	44	0.568
61	10	17	0.588	69	90	0.767
62	10	18	0.556	74	98	0.755
63	9	16	0.563	58	84	0.69
64	5	11	0.455	32	51	0.627
65	8	16	0.5	62	84	0.738
66	2	8	0.25	13	33	0.394
67	5	14	0.357	46	69	0.667
68	3	12	0.25	27	56	0.482

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
69	11	20	0.55	70	109	0.642
70	12	21	0.571	89	118	0.754
71	10	16	0.625	64	84	0.762
72	8	16	0.5	56	77	0.727
73	10	18	0.556	73	98	0.745
74	11	18	0.611	76	98	0.776
75	16	25	0.64	113	153	0.739
76	11	20	0.55	79	109	0.725
77	7	16	0.438	52	84	0.619
78	8	18	0.444	59	98	0.602
79	6	13	0.462	38	63	0.603
80	6	16	0.375	53	84	0.631
81	10	16	0.625	63	84	0.75
82	15	21	0.714	103	118	0.873
83	11	21	0.524	78	118	0.661
84	5	11	0.455	32	51	0.627
85	5	11	0.455	31	51	0.608

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
86	7	16	0.438	50	84	0.595
87	6	14	0.429	45	69	0.652
88	19	28	0.679	151	179	0.844
89	7	22	0.318	62	128	0.484
90	10	22	0.455	82	128	0.641
91	1	9	0.111	15	39	0.385
92	5	13	0.385	37	63	0.587
93	12	19	0.632	83	104	0.798
94	13	22	0.591	99	128	0.773
95	11	22	0.5	93	128	0.727
96	11	17	0.647	66	90	0.733
97	5	13	0.385	39	63	0.619
98	9	20	0.45	70	109	0.642
99	15	24	0.625	115	145	0.793
100	5	13	0.385	41	63	0.651
101	4	10	0.4	24	44	0.545
102	3	10	0.3	26	44	0.591

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
103	3	11	0.273	26	51	0.51
104	9	22	0.409	78	128	0.609
105	9	20	0.45	62	90	0.689
106	5	13	0.385	32	63	0.508
107	1	9	0.111	14	39	0.359
108	4	11	0.364	30	51	0.588
109	4	13	0.308	34	63	0.54
110	1	8	0.125	12	60	0.2
111	5	16	0.313	39	84	0.464
112	2	10	0.2	16	44	0.364
113	2	8	0.25	17	33	0.515
114	4	11	0.364	29	51	0.569
115	4	11	0.364	27	51	0.529
116	4	10	0.4	24	44	0.545
117	4	7	0.571	31	51	0.608
118	7	14	0.5	49	69	0.71
119	12	22	0.545	86	128	0.672

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
120	5	11	0.455	32	51	0.627
121	10	19	0.526	69	104	0.663
122	4	10	0.4	22	44	0.5
123	7	15	0.467	50	76	0.658
124	5	13	0.385	38	63	0.603
125	9	19	0.474	70	104	0.673
126	4	10	0.4	20	44	0.455
127	1	8	0.125	14	33	0.424
128	9	16	0.563	63	84	0.75
129	8	16	0.5	52	84	0.619
130	4	10	0.4	23	44	0.523
131	5	15	0.333	37	76	0.487
132	9	15	0.6	51	76	0.671
133	6	15	0.4	42	76	0.553
134	13	20	0.65	78	109	0.716
135	3	9	0.333	20	35	0.571
136	11	19	0.579	78	104	0.75

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
137	8	16	0.5	60	84	0.714
138	6	13	0.462	41	63	0.651
139	12	19	0.632	85	104	0.817
140	8	14	0.571	50	69	0.725
141	6	14	0.429	44	69	0.638
142	3	10	0.3	25	44	0.568
143	11	21	0.524	81	118	0.686
144	6	15	0.4	46	76	0.605
145	8	15	0.533	53	76	0.697
146	12	22	0.545	99	128	0.773
147	5	11	0.455	27	51	0.529
148	9	20	0.45	79	109	0.725
149	2	10	0.2	23	45	0.511
150	1	8	0.125	11	33	0.333
151	7	13	0.538	45	63	0.714
152	6	14	0.429	38	69	0.551
153	0	6	0	3	24	0.125

(continued)

Table 41 (continued)

Students' Spellings Scored with Traditional Scoring Pre-Test of Test Three

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
154	10	22	0.455	84	128	0.656
155	17	27	0.63	140	166	0.843
156	2	9	0.222	19	39	0.487
157	22	34	0.647	188	231	0.814
158	9	22	0.409	82	128	0.641
159	14	25	0.56	124	153	0.81

Table 42

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
1	11	22	0.5	86	128	0.672
2	17	31	0.548	160	204	0.784
3	6	13	0.462	43	63	0.683
4	12	22	0.545	95	128	0.742
5	14	22	0.636	104	128	0.813
6	7	17	0.412	40	69	0.58
7	20	28	0.714	158	180	0.878
8	11	22	0.5	86	128	0.672
9	10	20	0.5	77	128	0.602
10	6	14	0.429	37	69	0.536
11	13	22	0.591	97	128	0.758
12	11	27	0.407	77	166	0.464
13	11	25	0.44	110	153	0.719
14	7	20	0.35	63	109	0.578
15	18	30	0.6	165	193	0.855
16	9	16	0.563	68	84	0.81
17	14	22	0.636	108	128	0.844

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
18	21	29	0.724	162	187	0.866
19	19	28	0.679	156	180	0.867
20	21	31	0.677	163	204	0.799
21	9	22	0.409	83	128	0.648
22	15	26	0.577	131	159	0.824
23	11	22	0.5	88	128	0.688
24	12	21	0.571	94	118	0.797
25	7	13	0.538	46	63	0.73
26	13	31	0.419	147	204	0.721
27	20	31	0.645	173	204	0.848
28	18	29	0.621	144	187	0.77
29	16	22	0.727	112	128	0.875
30	17	25	0.68	128	153	0.837
31	13	22	0.591	101	128	0.789
32	21	30	0.7	169	187	0.904
33	11	20	0.55	79	109	0.725
34	6	13	0.462	42	63	0.667

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
35	21	32	0.656	176	217	0.811
36	19	29	0.655	150	187	0.802
37	14	22	0.636	100	128	0.781
38	14	21	0.667	93	118	0.788
39	17	23	0.739	118	136	0.868
40	6	13	0.462	38	63	0.603
41	19	31	0.613	163	204	0.799
42	14	22	0.636	103	128	0.805
43	13	22	0.591	99	128	0.773
44	12	22	0.545	87	128	0.68
45	8	22	0.364	83	128	0.648
46	17	25	0.68	129	153	0.843
47	21	33	0.636	183	223	0.821
48	13	22	0.591	100	128	0.781
49	13	22	0.591	98	128	0.766
50	27	39	0.692	231	283	0.816
51	14	22	0.636	97	128	0.758

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
52	16	29	0.552	140	187	0.749
53	7	15	0.467	52	76	0.684
54	7	22	0.318	69	128	0.539
55	8	20	0.4	70	109	0.642
56	15	22	0.682	110	128	0.859
57	17	30	0.567	153	194	0.789
58	10	22	0.455	74	128	0.578
59	13	22	0.591	95	128	0.742
60	6	13	0.462	38	63	0.603
61	20	32	0.625	172	217	0.793
62	16	22	0.727	111	128	0.867
63	12	22	0.545	92	128	0.719
64	10	22	0.455	80	128	0.625
65	15	24	0.625	116	145	0.8
66	4	15	0.267	35	76	0.461
67	10	22	0.455	86	128	0.672
68	6	12	0.5	34	56	0.607

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
69	12	22	0.545	90	128	0.703
70	18	29	0.621	160	187	0.856
71	18	30	0.6	155	195	0.795
72	10	22	0.455	76	128	0.594
73	10	20	0.5	89	109	0.817
74	18	31	0.581	166	204	0.814
75	14	24	0.583	118	145	0.814
76	17	31	0.548	160	204	0.784
77	15	29	0.517	140	187	0.749
78	11	22	0.5	84	128	0.656
79	11	22	0.5	94	128	0.734
80	11	22	0.5	98	128	0.766
81	16	31	0.516	145	204	0.711
82	22	33	0.667	200	232	0.862
83	14	25	0.56	119	153	0.778
84	10	20	0.5	75	109	0.688
85	5	13	0.385	40	63	0.635

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
86	12	22	0.545	83	128	0.648
87	9	22	0.409	89	128	0.695
88	24	31	0.774	172	203	0.847
89	13	22	0.591	97	128	0.758
90	13	22	0.591	108	128	0.844
91	5	14	0.357	37	69	0.536
92	13	22	0.591	107	128	0.836
93	17	25	0.68	130	153	0.85
94	17	27	0.63	135	153	0.882
95	17	31	0.548	160	203	0.788
96	15	22	0.682	97	128	0.758
97	7	13	0.538	42	63	0.667
98	14	22	0.636	105	128	0.82
99	19	30	0.633	152	193	0.788
100	10	21	0.476	85	118	0.72
101	4	10	0.4	27	44	0.614
102	9	16	0.563	55	84	0.655

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
103	5	11	0.455	35	51	0.686
104	12	27	0.444	122	166	0.735
105	23	39	0.59	215	282	0.762
106	9	22	0.409	85	128	0.664
107	7	16	0.438	53	84	0.631
108	9	16	0.563	61	84	0.726
109	15	29	0.517	148	186	0.796
110	0	6	0	6	24	0.25
111	3	15	0.2	37	121	0.306
112	13	22	0.591	99	128	0.773
113	6	13	0.462	37	63	0.587
114	11	25	0.44	102	145	0.703
115	12	22	0.545	90	128	0.703
116	9	22	0.409	74	128	0.578
117	6	13	0.462	42	63	0.667
118	8	20	0.4	84	109	0.771
119	13	22	0.591	102	128	0.797

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
120	10	19	0.526	71	104	0.683
121	15	22	0.682	107	128	0.836
122	4	12	0.333	28	56	0.5
123	14	22	0.636	97	128	0.758
124	10	22	0.455	87	128	0.68
125	14	22	0.636	100	128	0.781
126	3	10	0.3	25	44	0.568
127	5	12	0.417	28	56	0.5
128	14	25	0.56	120	153	0.784
129	12	22	0.545	86	128	0.672
130	8	22	0.364	73	128	0.57
131	6	15	0.4	42	76	0.553
132	10	19	0.526	74	104	0.712
133	9	20	0.45	71	109	0.651
134	13	22	0.591	106	128	0.828
135	5	13	0.385	31	63	0.492
136	13	22	0.591	95	128	0.742

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
137	13	22	0.591	96	128	0.75
138	7	14	0.5	45	69	0.652
139	16	24	0.667	116	145	0.8
140	13	22	0.591	104	128	0.813
141	9	22	0.409	75	128	0.586
142	4	11	0.364	28	51	0.549
143	14	28	0.5	136	179	0.76
144	9	22	0.409	81	128	0.633
145	10	22	0.455	90	128	0.703
146	18	31	0.581	167	203	0.823
147	7	17	0.412	60	90	0.667
148	13	22	0.591	109	128	0.852
149	5	12	0.417	37	56	0.661
150	4	11	0.364	25	51	0.49
151	11	18	0.611	74	98	0.755
152	9	20	0.45	75	128	0.586
153	4	10	0.4	26	44	0.591

(continued)

Table 42 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Three

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
154	11	22	0.5	102	128	0.797
155	21	28	0.75	162	179	0.905
156	3	13	0.231	29	63	0.46
157	28	44	0.636	267	324	0.824
158	9	15	0.6	56	76	0.737
159	17	20	0.85	102	109	0.936

Table 43

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
1	4	13	0.308	36	67	0.537
2	4	11	0.364	40	54	0.741
3	3	11	0.273	31	54	0.574
4	3	9	0.333	28	41	0.683
5	7	17	0.412	50	72	0.694
6	6	14	0.429	51	72	0.708
7	7	21	0.333	74	130	0.569
8	7	17	0.412	60	69	0.87
9	2	8	0.25	15	36	0.417
10	5	17	0.294	58	96	0.604
11	9	21	0.429	88	130	0.677
12	7	15	0.467	65	79	0.823
13	5	14	0.357	52	72	0.722
14	6	20	0.3	77	124	0.621
15	6	12	0.5	50	62	0.806
16	5	12	0.417	43	62	0.694
17	8	18	0.444	69	106	0.651

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
18	9	18	0.5	72	106	0.679
19	10	20	0.5	93	117	0.795
20	10	21	0.476	99	130	0.762
21	6	18	0.333	61	106	0.575
22	8	18	0.444	74	106	0.698
23	5	14	0.357	48	72	0.667
24	10	21	0.476	84	130	0.646
25	8	21	0.381	90	130	0.692
26	5	18	0.278	72	106	0.679
27	9	20	0.45	99	124	0.798
28	9	21	0.429	89	130	0.685
29	5	12	0.417	37	62	0.597
30	10	21	0.476	93	130	0.715
31	9	20	0.45	93	124	0.75
32	8	15	0.533	62	79	0.785
33	6	16	0.375	63	87	0.724
34	5	12	0.417	39	63	0.619

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
35	11	21	0.524	102	130	0.785
36	11	21	0.524	90	130	0.692
37	8	21	0.381	73	130	0.562
38	10	18	0.556	84	106	0.792
39	7	18	0.389	70	106	0.66
40	7	21	0.333	87	130	0.669
41	9	21	0.429	92	130	0.708
42	9	21	0.429	91	130	0.7
43	9	21	0.429	82	130	0.631
44	8	21	0.381	81	128	0.633
45	6	12	0.5	45	62	0.726
46	10	21	0.476	98	130	0.754
47	8	21	0.381	89	130	0.685
48	1	12	0.083	30	62	0.484
49	7	14	0.5	51	72	0.708
50	10	20	0.5	101	124	0.815
51	6	14	0.429	45	72	0.625

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
52	9	21	0.429	83	130	0.638
53	7	21	0.333	82	130	0.631
54	6	18	0.333	67	106	0.632
55	4	10	0.4	36	50	0.72
56	10	21	0.476	103	130	0.792
57	10	21	0.476	102	130	0.785
58	5	12	0.417	39	62	0.629
59	5	14	0.357	44	72	0.611
60	7	21	0.333	87	130	0.669
61	8	14	0.571	60	72	0.833
62	10	21	0.476	92	130	0.708
63	9	20	0.45	88	124	0.71
64	1	10	0.1	23	50	0.46
65	9	21	0.429	92	130	0.708
66	2	13	0.154	32	67	0.478
67	8	21	0.381	88	130	0.677
68	4	10	0.4	30	50	0.6

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
69	6	21	0.286	75	130	0.577
70	7	18	0.389	74	106	0.698
71	8	21	0.381	89	130	0.685
72	6	19	0.316	68	117	0.581
73	6	15	0.4	48	179	0.268
74	5	14	0.357	54	72	0.75
75	11	25	0.44	113	171	0.661
76	9	20	0.45	93	124	0.75
77	6	20	0.3	61	124	0.492
78	5	11	0.455	42	54	0.778
79	6	12	0.5	47	62	0.758
80	7	14	0.5	51	72	0.708
81	8	21	0.381	88	130	0.677
82	10	21	0.476	95	130	0.731
83	3	9	0.333	27	41	0.659
84	8	21	0.381	90	130	0.692
85	9	21	0.429	89	130	0.685

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
86	10	21	0.476	88	130	0.677
87	8	21	0.381	91	130	0.7
88	12	21	0.571	105	130	0.808
89	7	14	0.5	51	72	0.708
90	0	6	0	14	26	0.538
91	4	11	0.364	29	54	0.537
92	7	21	0.333	81	130	0.623
93	7	20	0.35	94	124	0.758
94	10	21	0.476	97	130	0.746
95	9	21	0.429	90	130	0.692
96	9	21	0.429	83	130	0.638
97	9	21	0.429	93	130	0.715
98	7	18	0.389	78	106	0.736
99	9	18	0.5	81	106	0.764
100	9	21	0.429	103	130	0.792
101	1	11	0.091	22	54	0.407
102	4	16	0.25	53	87	0.609

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
103	6	17	0.353	64	96	0.667
104	10	21	0.476	99	130	0.762
105	8	16	0.5	71	87	0.816
106	6	14	0.429	55	72	0.764
107	2	10	0.2	19	50	0.38
108	6	21	0.286	79	130	0.608
109	9	21	0.429	92	130	0.708
110	0	6	0	11	26	0.423
111	6	21	0.286	81	130	0.623
112	2	8	0.25	19	36	0.528
113	6	21	0.286	67	130	0.515
114	7	21	0.333	75	130	0.577
115	8	21	0.381	87	130	0.669
116	2	12	0.167	33	62	0.532
117	6	15	0.4	56	79	0.709
118	7	14	0.5	56	76	0.737
119	11	21	0.524	92	130	0.708

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
120	7	14	0.5	50	72	0.694
121	7	13	0.538	54	67	0.806
122	1	10	0.1	23	50	0.46
123	9	18	0.5	79	106	0.745
124	4	21	0.19	68	130	0.523
125	5	12	0.417	42	62	0.677
126	5	13	0.385	41	67	0.612
127	2	13	0.154	27	67	0.403
128	8	21	0.381	88	130	0.677
129	5	11	0.455	40	54	0.741
130	7	21	0.333	87	130	0.669
131	8	21	0.381	87	130	0.669
132	7	13	0.538	51	67	0.761
133	8	21	0.381	81	130	0.623
134	6	19	0.316	80	117	0.684
135	0	6	0	14	26	0.538
136	11	21	0.524	96	130	0.738

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
137	9	21	0.429	97	130	0.746
138	3	14	0.214	47	72	0.653
139	8	14	0.571	60	72	0.833
140	8	14	0.571	52	72	0.722
141	7	17	0.412	67	124	0.54
142	4	14	0.286	43	72	0.597
143	10	21	0.476	97	130	0.746
144	7	13	0.538	55	67	0.821
145	9	21	0.429	95	130	0.731
146	9	18	0.5	81	106	0.764
147	6	21	0.286	72	130	0.554
148	5	14	0.357	41	72	0.569
149	7	21	0.333	79	130	0.608
150	3	14	0.214	29	72	0.403
151	7	14	0.5	54	72	0.75
152	6	12	0.5	46	62	0.742
153	1	12	0.083	26	62	0.419

(continued)

Table 43 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Pre-Test of Test Sixteen

Student Id	Pre-Test with Traditional Scoring Correct	Pre-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Pre-Test with Non-Traditional Scoring Correct	Pre-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
154	6	13	0.462	52	67	0.776
155	12	21	0.571	108	130	0.831
156	1	8	0.125	22	36	0.611
157	13	25	0.52	128	171	0.749
158	5	14	0.357	48	72	0.667
159	10	18	0.556	84	106	0.792

Table 44

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
1	11	21	0.524	98	130	11
2	10	19	0.526	93	117	10
3	9	21	0.429	92	130	9
4	10	19	0.526	92	117	10
5	11	21	0.524	101	130	11
6	11	21	0.524	100	130	11
7	12	25	0.48	118	171	12
8	10	21	0.476	88	130	10
9	7	21	0.333	88	130	7
10	10	21	0.476	96	130	10
11	9	21	0.429	96	130	9
12	11	21	0.524	100	130	11
13	10	21	0.476	96	130	10
14	9	21	0.429	90	130	9
15	12	25	0.48	118	117	12
16	12	21	0.571	103	130	12
17	11	25	0.44	109	171	11

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
18	10	20	0.5	100	124	10
19	15	25	0.6	132	171	15
20	15	21	0.714	109	130	15
21	10	25	0.4	109	171	10
22	11	21	0.524	104	130	11
23	11	21	0.524	104	130	11
24	12	25	0.48	117	171	12
25	3	11	0.273	35	54	3
26	12	22	0.545	102	141	12
27	15	22	0.682	130	171	15
28	11	21	0.524	99	130	11
29	12	22	0.545	104	141	12
30	14	21	0.667	110	130	14
31	9	22	0.409	104	141	9
32	11	21	0.524	103	130	11
33	9	21	0.429	97	130	9
34	10	21	0.476	91	130	10

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
35	11	21	0.524	105	130	11
36	9	19	0.474	94	117	9
37	12	25	0.48	111	171	12
38	8	19	0.421	93	117	8
39	8	19	0.421	81	117	8
40	5	16	0.313	58	87	5
41	9	15	0.6	65	79	9
42	14	25	0.56	127	171	14
43	12	25	0.48	124	171	12
44	10	21	0.476	83	130	10
45	10	21	0.476	95	130	10
46	14	25	0.56	125	156	14
47	13	21	0.619	106	130	13
48	5	12	0.417	43	62	5
49	9	21	0.429	92	130	9
50	13	22	0.591	112	141	13
51	10	21	0.476	96	130	10

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
52	12	21	0.571	106	130	12
53	10	21	0.476	96	130	10
54	6	15	0.4	62	79	6
55	7	21	0.333	89	130	7
56	11	21	0.524	107	130	11
57	14	21	0.667	113	130	14
58	8	21	0.381	87	130	8
59	15	25	0.6	119	171	15
60	9	21	0.429	97	130	9
61	12	20	0.6	106	214	12
62	11	21	0.524	103	130	11
63	10	20	0.5	96	124	10
64	0	6	0	4	26	0
65	10	21	0.476	96	130	10
66	4	13	0.308	56	96	4
67	10	21	0.476	101	130	10
68	6	14	0.429	39	72	6

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
69	10	25	0.4	96	171	10
70	10	21	0.476	103	130	10
71	11	22	0.5	107	141	11
72	12	21	0.571	92	130	12
73	10	21	0.476	92	130	10
74	10	21	0.476	104	130	10
75	11	21	0.524	110	130	11
76	13	25	0.52	122	171	13
77	11	21	0.524	95	130	11
78	8	18	0.444	75	106	8
79	4	11	0.364	35	54	4
80	12	22	0.545	107	174	12
81	12	22	0.545	100	171	12
82	10	21	0.476	97	130	10
83	11	25	0.44	111	171	11
84	10	21	0.476	95	130	10
85	7	21	0.333	77	130	7

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
86	6	14	0.429	52	72	6
87	9	21	0.429	93	130	9
88	13	25	0.52	122	171	13
89	6	21	0.286	78	130	6
90	12	25	0.48	129	171	12
91	5	11	0.455	36	54	5
92	10	21	0.476	95	130	10
93	11	25	0.44	120	171	11
94	11	19	0.579	95	117	11
95	12	21	0.571	101	130	12
96	7	18	0.389	78	106	7
97	8	21	0.381	89	130	8
98	9	20	0.45	91	124	9
99	9	20	0.45	99	124	9
100	12	21	0.571	96	130	12
101	8	21	0.381	73	130	8
102	8	21	0.381	88	130	8

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
103	6	14	0.429	48	72	6
104	12	22	0.545	107	141	12
105	13	25	0.52	131	171	13
106	9	21	0.429	98	130	9
107	1	8	0.125	23	36	1
108	9	21	0.429	92	130	9
109	6	14	0.429	54	72	6
110	0	6	0	3	26	0
111	8	21	0.381	91	130	8
112	11	21	0.524	103	130	11
113	6	14	0.429	54	72	6
114	6	12	0.5	44	62	6
115	10	21	0.476	95	130	10
116	5	21	0.238	58	130	5
117	6	12	0.5	52	62	6
118	8	17	0.471	72	96	8
119	5	12	0.417	48	62	5

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
120	9	18	0.5	74	106	0.698
121	10	21	0.476	94	130	0.723
122	4	12	0.333	35	62	0.565
123	8	15	0.533	65	79	0.823
124	5	21	0.238	81	130	0.623
125	5	21	0.238	70	130	0.538
126	4	13	0.308	38	67	0.567
127	5	12	0.417	37	62	0.597
128	9	18	0.5	73	106	0.689
129	6	12	0.5	48	62	0.774
130	8	21	0.381	83	130	0.638
131	5	21	0.238	81	130	0.623
132	8	14	0.571	57	72	0.792
133	4	15	0.267	53	79	0.671
134	10	21	0.476	94	130	0.723
135	4	14	0.286	28	72	0.389
136	8	21	0.381	93	130	0.715

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
137	10	21	0.476	99	130	0.762
138	8	14	0.571	56	72	0.778
139	9	21	0.429	94	130	0.723
140	10	21	0.476	94	130	0.723
141	9	21	0.429	93	130	0.715
142	4	14	0.286	41	72	0.569
143	9	19	0.474	92	117	0.786
144	9	19	0.474	83	117	0.709
145	9	22	0.409	106	141	0.752
146	11	21	0.524	102	130	0.785
147	8	21	0.381	90	130	0.692
148	9	18	0.5	78	106	0.736
149	7	21	0.333	88	130	0.677
150	5	12	0.417	38	62	0.613
151	8	14	0.571	55	72	0.764
152	8	21	0.381	87	130	0.669
153	6	18	0.333	60	106	0.566

(continued)

Table 44 (continued)

Students' Spellings Scored with Traditional and Non-Traditional Scoring for Post-Test of Test Sixteen

Student Id	Post-Test with Traditional Scoring Correct	Post-Test with Traditional Scoring Total	% Correct Traditional Scoring Pre	Post-Test with Non-Traditional Scoring Correct	Post-Test with Non-Traditional Scoring Total	% Correct Non-Traditional Scoring Pre
154	8	18	0.444	77	106	0.726
155	14	22	0.636	116	141	0.823
156	5	18	0.278	67	106	0.632
157	14	25	0.56	134	171	0.784
158	10	21	0.476	98	130	0.754
159	14	25	0.56	128	171	0.749

REFERENCES

Abu-Hamoura, B., Urso, A., & Mather, N. (2013). The application of standardized assessments and CBM measures in a case study of a student with a severe reading disability. *Reading & Writing Quarterly, 29*(1), 44-63. doi:10.1080/10573569.2013.741950

Allred, R. A. (1990). Gender differences in spelling achievement in grades 1 through 6. *The Journal of Educational Research, 83,* 187-193. doi: 10.1080/00220671.1990.10885955

Bell, B. A., Ene, M., & Schoeneberger, J. (2013). A multilevel model primer using SAS PROC MIXED. *SAS Global Forum 2013.* SAS Institute Inc., Cary. 1-19.

Beretvas, S. N. (2010). *Cross-classified and multiple membership models*. In J. Hox & J. K. Roberts (Eds.), The handbook of advanced multilevel analysis (pp. 313-334). New York, NY: Routledge.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology, 46,* 151-172. doi:10.1016/j.jsp.2007.02.007

Borman, G. D., Slavin, R. E., Cheung, A. C. K., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of success for all. *American Educational Journal, 44,* 701-731. doi:10.3102/0002831207306743

Bruck, M. (1993). Component spelling skills of college students with childhood diagnoses of dyslexia. *Learning Disability Quarterly, 16,* 171-184.

Brus, B. T., & Voeten, M. (1973). EeÂn-minuut-test. [One minute test of reading.] Nijmegen, The Nether-lands: Berkhout.

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical Linear Models in Social and Behavioral Research: Applications and Data Analysis Methods* (First Edition). Newbury Park, CA: Sage Publications.

Burns, M.S, Griffin, P., & Snow, C.E. (Eds.). (1999). *Starting out right: A guide to promoting children's reading success*. Washington, DC: National Academy Press. Retrieved from: http://www.nap.edu/html/sor/

Cataldo, S., & Ellis, N. (1988). Interactions in the development of spelling, reading and phonological skills. *Journal of Research in Reading, 11*, 86-109. doi:10.1111/j.1467-9817.1988.tb00153.x

Chamberlain, S. P. (2004). The effects of the No Child Left Behind Act on diverse learners. *Intervention in School & Clinic*, 40, 96-105.

Clay, M. M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.

Comaskey, E. M., Savage, R. S., & Abrami, P. (2009). A randomised efficacy study of web-based synthetic and analytic programmes among disadvantaged urban kindergarten children. *Journal of Research in Reading, 32*, 92-108. doi:10.1111/j.1467-9817.2008.01383.x

Conrad, N. J. (2008). From reading to spelling and spelling to reading: Transfer goes both ways. *Journal of Educational Psychology, 100*, 869-878. doi:10.1037/ a0012544

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49,* 36-45.

Di Stasio, M. R., Savage, R., & Abrami, P. C. (2012). A follow-up study of the ABRACADABRA web-based literacy intervention in grade 1. *Journal of Research in Reading, 35*(1), 69-86. doi:10.1111/j.1467-9817.2010.01469.x

Dunn, L.M. & Dunn, L. M. (1996). *Peabody Picture Vocabulary Test.* (3rd ed). Circle Pines, MN: American Guidance Service.

Ehri, L. C. (1986). Sources of difficulty in learning to spell and read. In D. Routh (Ed.), Advances in developmental and behavioral pediatrics (Vol. 7, pp. 121-195).

Ehri, L.C., & Wilce, L. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly*, *22,* 47-65.

Ehri, L.C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22,* 356-365

Ehri, L. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders*, *20*(3), 19-49.

Ellis, N., & Cataldo, S. (1990). The role of spelling in learning to read. *Language and Education*, *4*(1), 1-28. doi:10.1080/09500789009541270

Ellis, N., & Large, B. (1988). The early stages of reading: A longitudinal study. *Applied Cognitive Psychology, 2*(1), 47-76. doi:10.1002/acp.2350020106

Entwisle, D. R., & Alexander, K. L. (1988). Factors affecting achievement test scores and marks of black and white first graders. *The Elementary School Journal, 88*, 449-471.

Fasting, R. B., & Lyster, S. H. (2005). The effects of computer technology in assisting the development of literacy in young struggling readers and spellers. *European Journal of Special Needs Education, 20*(1), 21-40. doi:10.1080/0885625042000319061

Foorman, B. R., & Petscher, Y. (2010). Development of spelling and differential relations to text reading in grades 3–12. *Assessment for Effective Intervention, 36*(1), 7-20. doi: 10.1177/1534508410379844

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16,* 203-212. doi:10.1111/0938-8982.00020

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 495-515). London, UK: Academic Press.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Allinder, R. M. (1991). Effects of expert system advice within curriculum-based measurement on teacher planning and student achievement in spelling. *School Psychology Review, 20,* 49–66.

Fulk, B. M., & Stormont-Spurgin, M. (1995). Spelling interventions for students with disabilities: A review. *The Journal of Special Education, 28,* 488-513. doi:10.1177/002246699502800407

Garson, G. D. (2013). *Hierarchical linear modeling: Guide and applications*. Los Angeles: SAGE Publications.

Geelhoed, J., Reitsma, P. (1999) PI-dictee. Harcourt, Amsterdam, Netherlands.

Gentry, J.R. (1982). An analysis of developmental spelling in GNYS AT WRK, *The Reading Teacher, 36,* 192–200.

Gentry, J. R. (1984). Developmental aspects of learning to spell. *Academic Therapy*, *20*(1), 11-19.

Gentry, J. R. (1987). *Spel-- is a four-letter word*. Portsmouth, N.H: Heinemann.

Graham, S., Harris, K. R., Mason, L., Fink-Chorzempa, B., Moran, S., & Saddler, B. (2008). How do primary grade teachers teach handwriting: A national survey. *Reading and Writing: An Interdisciplinary Journal*, *21*, 49–69.

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading *Harvard Educational Review, 81,* 710-744.

Harper, H., Helmer, J., Lea, T., Chalkiti, K., Emmett, S., & Wolgemuth, J. R. (2012). ABRACADABRA for magic under which conditions? Case studies of a web-based literacy intervention in the Northern Territory. Australian Journal of Language and Literacy, *35*(1), 33–50.

Heck, R. H., & Thomas, S. L. (2000). *An introduction to multilevel modeling techniques.* Mahwah, NJ: Sage.

Heck, R. H., Thomas, S., L., & Tabata, L. N. (2014). *Multilevel and longitudinal modeling with IBM SPSS* (2nd ed.). New York, NY: Taylor & Francis.

Henderson, E. (1985). *Teaching spelling* (1st ed.). Boston, MA: Houghton Mifflin.

Henderson, E.H. (1990). *Teaching spelling* (2nd ed.). Boston, MA: Houghton Mifflin.

Hernandez, D. J. 2011. *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*, Baltimore, MD: Annie E. Casey Foundation.

Howell, R. D., Erickson, K., Stanger, C., & Wheaton, J. E. (2000). Evaluation of a computer-based program on the reading performance of first grade students with potential for reading failure. *Journal of Special Education Technology, 15*(4), 5-14. Retrieved from http://ezproxy.library.arizona.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eft&AN=507725934&site=ehost-live

Hox, J. J. (1995). *Applied multilevel Analysis.* Amsterdam: TT-Publikaties.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*(4), 243-255. Retrieved from http://ezproxy.library.arizona.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=1986-29616-001&site=ehost-live

Katusic, S., Barbaresi, W., Colligan, R., Bagniewski, S., & Weaver, A. (2006). Written language learning disorder: Incidence in a population-based birth cohort, 1976-1982, rochester, minnesota. *Journal of Developmental and Behavioral Pediatrics, 27*, 439-440. doi:10.1097/00004703-200610000-00047

Kempen, E. (2012). In Salkind N. (Ed.), *Exploring research* (8th ed.). New York, NY: Pearson.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions* (pp. 3-90). San Francisco, CA: Jossey-Bass.

Kreft, G. G., & Leeuw, J. d. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Lefly, D. L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia, 41,* 143-162.

Marsh, H. W., Hau, K. T., & Kong, C. K. (2002). Multilevel causal ording of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. *American Educational Research Journal, 39,* 727-763.

Masterson, J. J., & Apel, K. (2010). The spelling sensitivity score: Noting developmental changes in spelling knowledge. *Assessment for Effective Intervention, 36*(1), 35-45. doi:10.1177/1534508410380039

MindPlay Educational Software for Reading. (2015). *MindPlay Virtual Reading Coach.* Retrieved from: http://www.mindplay.com/programs/mindplay-virtual-reaching-coach/.

Moats, L. C. (2005). How Spelling Supports Reading. *American Educator, Winter 2005/06,* 12-43.

Mommers, M. J. C., & Boland, T. (1987). The development of decoding skills, reading comprehension and spelling skills in Dutch primary schools: A longitudinal investigation. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Morris, D., & Perney, J. (1984). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology*, *86,* 221-234.

Morris, D., Tyner, B., & Perney, J. (2000). Early steps: Replicating the effects of a first-grade reading intervention program. *Journal of Educational Psychology, 92*, 681-693. doi:10.1037/0022-0663.92.4.681

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (National Institute of Health Publication No. 00-4769). Washington, DC: National Institute of Child Health and Human Development.

Neuman, S. B., Copple, C., Bredekamp, S., & National Association for the Education of Young Children, Washington, DC. (2000). Learning to read and write: Developmentally appropriate practices for young children National Association for the Education of Young Children. *Young Children. 53*(4), 30–46.

Otaiba, S. A, & Hosp, J. L. (2010). Spell it out: The need for detailed spelling assessment to inform instruction. *Assessment for Effective Intervention, 36*(1), 3-6. doi:10.1177/1534508410384478

Raudenbush, S.W. and Bryk, A.S. (2002). *Hierarchical Linear Models* (Second Edition). Thousand Oaks, CA: Sage Publications.

Read, C. (1975). *Children's categorization of speech sounds in English*. Urbana, IL: National Council of Teachers of English.

Reid, K. (1988). *Teaching the learning disabled: A cognitive developmental approach*. Needham, MA: Allyn & Bacon.

Ritchey, K. D., & Goeke, J. L. (2006). Orton-gillingham and orton-gillingham-based reading instruction: A review of the literature. *The Journal of Special Education, 40*, 171-183. doi:10.1177/00224669060400030501

Ritchey, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. *Reading and Writing, 21*(1), 27-47. doi:10.1007/s11145-007-9063-0

Roa, C., Prakash, P., & Malatesha Joshi, R. (2006). Development of scoring system for kannada spelling assessment. *National Academy of Psychology, 51*, 49-51.

Savage, R. S., Erten, O., Abrami, P. C., Hipps, G., Comaskey, E., & van Lierop, D. (2010). ABRACADABRA in the hands of teachers: The effectiveness of a web-based literacy intervention in Grade 1 language arts programs. *Computers & Education, 55*, 911– 922. doi: 10.1016/j.compedu.2010.04.002

Schrank, F.A., McGrew, K.S., Mather, N., & Woodcock, R.W. (2014). *Woodcock-Johnson IV.* Rolling Meadows, IL: Riverside Publishing.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford; New York: Oxford University Press.

Slavin, R.E., & Madden, N.A. (Eds.) (2001). *Success for All: Research and reform in elementary education*. Mahwah, NJ: Erlbaum.

Slavin, R.E., Madden, N.A., Dolan, L., & Wasik, B.A. (1996). *Every child, every school: Success for All*. Thousand Oaks, CA: Corwin Press

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Los Angeles, CA: Sage.

Snow, C. E., Griffin, P. & Burns, M. S. (Eds.) (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*. San Francisco: Jossey-Bass.

Snowling, M., & Perin, D. (1983) The development of phoneme segmentation skills in young children. In J. Sloboda (ed.) *The Acquisition of Symbolic Skills.* London, UK: Plenum Press.

Spaulding, R. L., & Papageorgiou, M. R. (1977). *Observation of the coping behaviour of children in elementary schools*. San Jose, CA: San Jose State University.

Tabachnick, G. G., and Fidell, L. S. (2007). *Experimental designs using ANOVA.* Belmont, CA: Duxbury.

Tabachnick, B. G. (March, 2008). Multivariate statistics: An introduction and some applications. Invited workshop presented to the American Psychology-Law Society, Jacksonville, FL.

Treiman, R., & Bourassa, D. C. (2000). The development of spelling skills. *Topics in Language Disorders*, *20*(3), 1-18. doi:10.1097/00011363-200020030-00004

Treiman, R. (1998) Why spelling?: The benefits of incorporating spelling into beginning reading instruction. In J. Metsala & L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 289-313). Mahwah, NJ: Erlbaum.

Twisk, J. W. R. (2006). *Applied multilevel analysis*. Cambridge, UK: Cambridge University Press.

National Assessment of Educational Progress (NAEP). (2013). *Reading Assessment*. U.S. Department of Education, Institute of Education Sciences.

Van Daal, V., & Reitsma, P. (2000). Computer-assisted learning to read and spell: Results from two pilot studies. *Journal of Research in Reading, 23,* 181-193. doi:10.1111/1467-9817.00113

Van den Bos, K. P., Scheepstra, A. J. M., & Lutje Spelberg, H. C. (1993) Het lezen van pseudowoorden en bestaande woorden [The reading of pseudo words and real words]. *Tijdschrift voor Orthopedagogiek, 32,* 158-168.

Weiner, S. (1994). Effects of phonemic awareness training on low- and middle-achieving first graders' phonemic awareness and reading ability. *Journal of Reading Behavior*, *26,* 277-300.

Wilkinson, G.S. *(*1993*). The Wide Range Achievement Test.* Wilmington, DE: Wide Range Inc.

Young, K. (2007). Developmental stage theory of spelling: Analysis of consistency across four spelling-related activities. *Australian Journal of Language and Literacy, 30,* 203–220.

Zirpoli, T. J., & Melloy, K. J. (1997). *Behavior management: Applications for teachers and parents*. Upper Saddle River, N.J.: Merrill.