

## ABSTRACT

Title of Dissertation:           READING RECOVERY CHILDREN AND EARLY LITERACY DEVELOPMENT: INVESTIGATION INTO PHONOLOGICAL AWARENESS, ORTHOGRAPHIC KNOWLEDGE, ORAL READING PROCESSING, AND READING COMPREHENSION PROCESSING

Judith Seeber Concha, Doctor of Philosophy, 2005

Dissertation directed by:     Dr. Marilyn Chambliss  
Department of Curriculum and Instruction

Marie Clay (1993) designed Reading Recovery tutoring to accelerate the early literacy development of low-performing, six-year-old children so that they achieve average levels of classroom performance. Approximately one third of the proportion of the first cohort of U.S. children who receive Reading Recovery tutoring at the beginning of a school year respond poorly to it (Gómez-Bellengé, Rodgers, & Fullerton, 2003). They fail to meet the criteria for successful performance and their Reading Recovery teachers recommend them for additional assessment and/or consideration for other supplemental instruction. An emerging program of research suggests that recommended children struggle in early literacy development.

This study compared recommended to discontinued Reading Recovery children

on phonological awareness and orthographic knowledge at pre- and post-tutoring, and oral reading processing and reading comprehension processing at post-tutoring. The sample consisted of 29 recommended children and 26 discontinued children who were taught by 16 trained Reading Recovery teachers in a single school district. This study contributes to the understanding of recommended children's early literacy development.

Analysis of phonological awareness and orthographic knowledge composite data revealed that recommended children demonstrated less overall phonological awareness and overall orthographic knowledge than discontinued children and that recommended and discontinued children combined displayed gains from pre- to post-tutoring at statistically significant levels. Analyses of the phonological awareness composite data revealed that recommended children performed at a level below discontinued children on rhyme awareness at pre-tutoring, phonological skeletal structure awareness at pre- and post-tutoring, and graphophonemic awareness with respect to beginning phonemes at pre- and post-tutoring and ending phonemes at pre-tutoring at statistically significant levels. Analyses of the orthographic knowledge composite data revealed that recommended children performed at a level below discontinued children on orthographic acceptability knowledge at pre- tutoring and spelling knowledge at post-tutoring at statistically significant levels. Analyses of oral reading processing data at post-tutoring revealed that recommended children read stories with less accuracy, more overall errors, more substitutions, less fluency, and at a slower rate than discontinued children at statistically significant levels. An analysis of reading comprehension processing data at post-tutoring revealed that the two groups comprehended the stories nearly equivalently.

READING RECOVERY CHILDREN AND EARLY LITERACY DEVELOPMENT:  
INVESTIGATION INTO PHONOLOGICAL AWARENESS, ORTHOGRAPHIC  
KNOWLEDGE, ORAL READING PROCESSING, AND READING  
COMPREHENSION PROCESSING

by

Judith Seeber Concha

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2005

Advisory Committee:

Associate Professor Marilyn Chambliss, Chair  
Professor Peter Afflerbach  
Professor Mariam Jean Dreher  
Associate Professor John O'Flahavan  
Professor Deborah Speece

©Copyright by  
Judith Seeber Concha  
2005

## TABLE OF CONTENTS

List of Tables.....	vi
List of Figures.....	x
Chapter 1: Introduction.....	1
Reading Recovery.....	2
Early Literacy Development.....	7
Phonological Awareness.....	9
Orthographic Knowledge.....	12
Oral Reading Processing.....	14
Reading Comprehension Processing.....	15
Early Literacy Instruction.....	16
Tutoring Effectiveness.....	17
Reading Recovery Tutoring Responsiveness.....	19
Research Questions.....	20
Significance.....	21
Overview of the Methodology.....	23
Rationale for the Selection of Early Literacy Measures.....	25
Definitions.....	26
Summary.....	39
Chapter 2: Early Literacy Development and Reading Recovery Tutoring....	41
Interactive Model of Early Literacy Development.....	42
Early Literacy Development.....	44
Phonological Awareness.....	45
Research on Phonological Awareness: Correlational Studies.....	46
Research on Phonological Awareness: Experimental Studies.....	51
Phonological Awareness and the Interactive Model of Early Literacy Development.....	56
Summary.....	57
Orthographic Knowledge.....	58
Research on Orthographic Knowledge.....	59
Orthographic Knowledge and the Interactive Model of Early Literacy Development.....	63
Summary.....	64
Oral Reading Processing.....	65
Research on Reading Words by Sight.....	66
Research on Reading Words by Decoding.....	68
Research on Reading Words by Analogizing.....	70
Research on Reading Words by Using Context.....	74
Oral Reading Processing and the Interactive Model of Early Literacy Development.....	76

Summary.....	77
Reading Comprehension Processing.....	77
Research on Reading Comprehension Strategies.....	78
Research on Reading Comprehension Monitoring.....	81
Reading Comprehension Processing and the Interactive Model of Early Literacy Development.....	84
Summary.....	85
Reading Recovery Tutoring.....	86
Reading Recovery Tutoring Effectiveness.....	86
Reviews of Research on Reading Recovery Tutoring Effectiveness.....	87
Research on Reading Recovery Tutoring Effectiveness.....	91
Summary.....	97
Reading Recovery Tutoring Responsiveness.....	98
Research on Reading Recovery Tutoring Responsiveness.....	99
Summary.....	105
Reading Recovery Tutoring Effectiveness and Responsive- ness and the Interactive Model of Early Literacy Development.....	105
Conclusion.....	106
Chapter 3: Methodology.....	108
Setting.....	109
Reading Recovery Schools.....	109
Reading Recovery Children.....	110
Reading Recovery Teachers.....	110
Measures.....	111
Pilot Study 1 and 2.....	111
Measurement Terms.....	116
An Observation Survey of Early Literacy Achievement.....	120
Letter Identification.....	122
Concepts About Print.....	127
Word Reading.....	128
Writing Vocabulary.....	130
Hearing and Recording Sounds in Words.....	130
Text Reading.....	131
Rhyme Detection Task.....	133
Blending Words Task.....	136
Sentence Writing and Spelling Task.....	141
Gray Oral Reading Tests-Fourth Edition.....	148
Summary.....	158
Procedures.....	160
Pre-tutoring Procedures.....	160
Post-tutoring Procedures.....	162

Data Analyses.....	163
Composite Formation.....	163
Repeated Measure Two-way Analysis of Variance.....	168
Percentages.....	169
Chi-square Test of Independence .....	169
One-way Analysis of Variance.....	170
Summary.....	171
Chapter 4: Results.....	173
Phonological Awareness.....	174
Overall Phonological Awareness.....	174
Rhyme Awareness.....	178
Phonological Skeletal Structure Awareness.....	179
Combined Syllable, Onset and Rime, and Phonemic Awareness.....	181
Graphophonemic Awareness.....	182
Summary.....	185
Orthographic Knowledge.....	186
Overall Orthographic Knowledge.....	186
Spelling Knowledge.....	190
Orthographic Pattern Knowledge.....	192
Orthographic Acceptability Knowledge.....	195
Summary.....	195
Oral Reading Processing.....	196
Oral Reading Accuracy.....	197
Oral Reading Inaccuracy.....	200
Oral Reading Behaviors.....	205
Summary.....	211
Reading Comprehension Processing.....	212
Summary.....	213
Oral Reading and Reading Comprehension Processing Composite.....	215
Summary.....	216
A Profile of an Average-performing Recommended Reading Recovery Child.....	216
Phonological Awareness.....	217
Orthographic Knowledge.....	222
Oral Reading Processing.....	227
Reading Comprehension Processing.....	233
Summary.....	234
Summary.....	234
Chapter 5: Discussion.....	238
Study Summary.....	238
Major Findings.....	240
Phonological Awareness.....	241
Orthographic Knowledge.....	245

Oral Reading Processing.....	247
Reading Comprehension Processing.....	248
Summary.....	250
Limitations.....	251
Recommendations for Instruction.....	252
Recommendations for Reading Recovery Tutoring and First-grade Instruction.....	252
Recommendations for Kindergarten Instruction.....	261
Summary.....	263
Recommendations for Research.....	264
Conclusion.....	266
Appendix A: Pilot Study 1 and 2.....	268
Appendix B: Dissertation Study Informed Consent.....	300
References.....	301



## List of Tables

Table 1	Early Literacy Knowledge and Processing, Corresponding Components, and Testing Occasions	10
Table 2	Phonological Awareness Tasks Employed in Pilot Study 1 and 2	112
Table 3	Reading Recovery Teacher Administered Measure, Assessed Early Literacy Component, and Testing Occasions	121
Table 4	Internal Consistency and Test-retest Reliability Information for An Observation Survey of Early Literacy Achievement	123
Table 5	Criterion-related Validity Information for An Observation Survey of Early Literacy Achievement	125
Table 6	Criterion-related Validity Information for An Observation Survey of Early Literacy Achievement	126
Table 7	Construct Validity Information for An Observation Survey of Early Literacy Achievement	127
Table 8	Sample Items and Passage Excerpts from An Observation Survey of Early Literacy Achievement	129
Table 9	Researcher Administered Measures, Scoring Procedures, and Assessed Early Literacy Components	134
Table 10	Sample Items and Passage Excerpts from the Rhyme Detection Task, the Blending Words Task, the Sentence Writing and Spelling Task, and the GORT-4	137
Table 11	Internal Consistency and Test-retest Reliability Information for the Rhyme Detection Task and the Blending Words Task	138
Table 12	Criterion-related Validity Information for the Rhyme Detection Task and the Blending Words Task	139
Table 13	Phonological Skeletal Structure Awareness Scoring Examples	143
Table 14	Graphophonemic Awareness Scoring Examples	145
Table 15	Test-retest Reliability Information for the GORT-4	153

Table 16	Internal Consistency and Parallel Split-half Reliability Information for the GORT-4	154
Table 17	Criterion-related Validity Information for the GORT-4	156
Table 18	Content and Construct Validity Information for the GORT-4	157
Table 19	Formation of the Overall Phonological Awareness Composite at Pre- and Post-tutoring	165
Table 20	Formation of the Overall Orthographic Knowledge Composite at Pre- and Post-tutoring	167
Table 21	Means and Standard Deviations for Proportion Correct for Overall Phonological Awareness; Rhyme Awareness; Phonological Skeletal Structure Awareness; Combined Syllable, Onset and Rime, and Phonemic Awareness; and Graphophonemic Awareness for Group, Time, and Group X Time Interaction	176
Table 22	2 (Group) X 2 (Time) Repeated Measure ANOVA for Overall Phonological Awareness	178
Table 23	Recommended and Discontinued Children's Percentages for Rhyme Awareness at Pre- and Post-tutoring	179
Table 24	Recommended and Discontinued Children's Percentages For Phonological Skeletal Structure Awareness at Pre- and Post-tutoring	180
Table 25	Recommended and Discontinued Children's Percentages for Combined Syllable, Onset and Rime, and Phonemic Awareness at Pre- and Post-tutoring	182
Table 26	Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Beginning Sounds at Pre- and Post-tutoring	184
Table 27	Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Middle Sounds at Pre- and Post-tutoring	184
Table 28	Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Ending Sounds at Pre- and Post-tutoring	185

Table 29	Means and Standard Deviations for Proportion Correct for Overall Orthographic Knowledge, Spelling Knowledge, Orthographic Pattern Knowledge, and Orthographic Acceptability Knowledge for Group, Time, and Group X Time Interaction	188
Table 30	2 (Group) X 2 (Time) Repeated Measure ANOVA for Overall Orthographic Knowledge	190
Table 31	Recommended and Discontinued Children's Percentages for Spelling Knowledge at Pre- and Post-tutoring	191
Table 32	Recommended and Discontinued Children's Percentages for Orthographic Pattern Knowledge at Pre- and Post-tutoring	193
Table 33	Recommended and Discontinued Children's Percentages for Specific Orthographic Pattern Knowledge at Post-tutoring	194
Table 34	Recommended and Discontinued Children's Percentages for Orthographic Acceptability Knowledge at Pre- and Post-Tutoring	196
Table 35	Means and Standard Deviations for GORT-4 Accuracy and Modified GORT-4 Accuracy at Post-tutoring	198
Table 36	One-way ANOVA for GORT-4 Accuracy at Post-tutoring	198
Table 37	One-way ANOVA for Modified GORT-4 Accuracy at Post-tutoring	199
Table 38	Percentages of Recommended and Discontinued Children's Accuracy, Rate, Fluency, Comprehension, and Composite Standard Scores Placed in the GORT-4 Rating Category Scale	201
Table 39	Means and Standard Deviations for Overall Errors, Substitutions, Tolds, Omissions, and Insertions on Story 1 and 2 from the GORT-4 at Post-tutoring	203
Table 40	One-way ANOVA for Overall Errors on Story 1 and 2 from the GORT-4 at Post-tutoring	204
Table 41	One-way ANOVA for Substitutions on Story 1 and 2 From the GORT-4 at Post-tutoring	204

Table 42	One-way ANOVA for Tolds on Story 1 and 2 from the GORT-4 at Post-tutoring	205
Table 43	One-way ANOVA for Omissions on Story 1 and 2 from the GORT-4 at Post-tutoring	206
Table 44	One-way ANOVA for Insertions on Story 1 and 2 from the GORT-4 at Post-tutoring	206
Table 45	Means and Standard Deviations for GORT-4 Rate, GORT-4 Fluency, Self-Corrections on Story 1 and 2 from the GORT-4, and Repetitions on Story 1 and 2 from the GORT-4 at Post-tutoring	208
Table 46	One-way ANOVA for GORT-4 Rate at Post-tutoring	209
Table 47	One-way ANOVA for GORT-4 Fluency at Post-tutoring	209
Table 48	One-way ANOVA for Self-corrections on Story 1 and 2 from the GORT-4 at Post-tutoring	211
Table 49	One-way ANOVA for Repetitions on Story 1 and 2 from the GORT-4 at Post-tutoring	212
Table 50	Means and Standard Deviations for GORT-4 Comprehension at Post-tutoring	214
Table 51	One-way ANOVA for GORT-4 Comprehension at Post-Tutoring	214

## List of Figures

Figure 1	Interactive Model of Early Literacy Development	43
Figure 2	Overall Phonological Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	218
Figure 3	Rhyme Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	219
Figure 4	Phonological Skeletal Structure Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	220
Figure 5	Graphophonemic Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	221
Figure 6	Combined Syllable, Onset and Rime, and Phonemic Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	222
Figure 7	Overall Orthographic Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	224
Figure 8	Spelling Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	225
Figure 9	Orthographic Pattern Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	226
Figure 10	Orthographic Acceptability Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring	227
Figure 11	Oral Reading Accuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring	229
Figure 12	Modified Oral Reading Accuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring	229

Figure 13	Oral Reading Inaccuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring	230
Figure 14	Oral Reading Behaviors of Rate and Fluency of an Average-performing Recommended and Discontinued Child at Post-tutoring	232
Figure 15	Oral Reading Behaviors of Self-corrections and Repetitions of an Average-performing Recommended and Discontinued Child at Post-tutoring	232
Figure 16	Reading Comprehension Processing of an Average-performing Recommended and Discontinued Child at Post-tutoring	233

## Chapter 1

### *Introduction*

Reading Recovery is a tutoring program designed to accelerate the early literacy development of low-performing, six-year-old children so that they achieve average levels of classroom performance (Clay, 1993). Since its development in New Zealand by Marie Clay, Reading Recovery has grown in popularity, as evidenced by its presence in New Zealand, Australia, Canada, the United Kingdom, the United States, and the Department of Defense Dependent Schools. Many teachers, administrators, researchers, organizations, and government agencies concerned with children's early literacy development consider Reading Recovery the prototype of tutoring programs.

Researchers have documented the effectiveness of Reading Recovery (e.g., Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994). Yet nearly one third of U.S. children selected first for Reading Recovery tutoring at the beginning of a school year do not respond well to it (Gómez-Bellengé, Rodgers, & Fullerton, 2003). It may be that this cohort of children struggle in their development of literacy knowledge and literacy processing prior to and/or upon their completion of tutoring. As a former Reading Recovery teacher and teacher leader, this possibility intrigued me.

A poor initial start in early literacy development may be influenced by many factors, including cognitive, linguistic, perceptual, biological, instructional, sociocultural, economic, and political factors (Snow, Burns, & Griffin, 1998). For example, Stanovich (1986) argued that the achievement gap between children who struggle in early literacy development and their more accomplished peers widens during the early years of schooling. Juel (1988) found that children who struggled in learning to read in first-grade

continued to struggle at the end of fourth-grade.

The dire consequences that a poor initial start in early literacy development has on later development are widely recognized. However, little is known about the poor initial start of the first cohort of Reading Recovery children who fail to respond well to their tutoring. Only a handful of studies have examined the early literacy development of these children (Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Chapman, Tunmer, & Prochnow, 2001; Clay & Tuck, 1991; Spector & Moore, 2004). In the current study, I aimed to contribute to the understanding of the early literacy development of these children and build upon these studies. To this end, I assessed their (a) phonological awareness and orthographic knowledge prior to and following their tutoring and (b) oral reading processing and reading comprehension processing following their tutoring.

Chapter 1 is organized into eight sections. Section 1 provides an overview of Reading Recovery. Section 2 summarizes early literacy development. It discusses children's literacy knowledge as it relates to phonological awareness and orthographic knowledge and children's literacy processing as it relates to oral reading processing and reading comprehension processing. Section 3 discusses early literacy instruction in regards to (a) the effectiveness of various tutoring programs, including Reading Recovery, and (b) children's responsiveness to Reading Recovery tutoring. Sections 4 through 8 present the research questions, the significance, an overview of the methodology, a rationale for the selection of early literacy measures, and the definitions of terms, respectively.

### *Reading Recovery*

Reading Recovery is designed to ameliorate the reading and writing difficulties of



low-performing, six-year-old children through intensive tutoring. This tutoring supplements classroom literacy instruction. The goal is to accelerate children's early literacy development so that they achieve average levels of classroom performance.

During the 1970s and 1980s, Marie Clay conducted several studies in New Zealand that led to the development and refinement of Reading Recovery, including field trial studies, a replication study, a lesson analysis study, and follow-up studies (Clay, 1993). Reading Recovery has been operating in the United States since 1984 when Clay visited The Ohio State University to train a small cohort of teachers and university professors. Reading Recovery has grown considerably since then. According to the National Data Evaluation Center (NDEC), an organization that analyzes data for the Reading Recovery programs in the United States, nearly 140,000 children received tutoring from approximately 17,000 teachers in 10,000 schools during the 2002-2003 school year (Gómez-Bellengé & Rodgers, 2004).

The training that Reading Recovery teachers receive extends for one academic year. During this training year, teachers attend a graduate class weekly and tutor four children daily. During the graduate class, teachers frequently tutor children behind a one-way glass or screen. As one teacher delivers the tutoring lesson on one side of the glass, the other teachers observe and discuss the lesson on the other side. Following this training year, Reading Recovery teachers participate in on-going professional development for as long as they tutor children.

In the United States Reading Recovery children receive tutoring lessons after they have received one year of literacy instruction in kindergarten. They receive 30-minute lessons each day for an average of 12 to 20 weeks. As soon as the children read and write

within the average range of their first-grade classes and demonstrate that they can continue to progress in their early literacy development, their Reading Recovery teachers discontinue tutoring and select new children to tutor.

A Reading Recovery lesson consists of individual children (a) rereading several familiar texts, (b) reading yesterday's new text while a Reading Recovery teacher takes a running record, (c) completing isolated letter identification or word analysis activities, (d) composing and writing one or two sentences, (e) reassembling the cut-up sentence(s), and (f) reading a new text that a Reading Recovery teacher introduced. Reading Recovery teachers employ various teaching procedures that correspond to these lesson components (Clay, 1993). While these teaching procedures outline specific techniques, they are not scripted. When employing specific teaching procedures, teachers make instructional decisions with the strengths and competencies of their individual children in mind.

To assess children's early literacy development and monitor their progress, Clay developed *An Observation Survey of Early Literacy Achievement* (Clay, 2002). This early literacy assessment consists of the following tasks: Letter Identification, Concepts About Print, Word Reading, Writing Vocabulary, Hearing and Recording Sounds in Words, and Text Reading (see Chapter 3 for a summary of these tasks). Reading Recovery teachers consider children's performance on this assessment to select children for tutoring and to discontinue them from tutoring.

Reading Recovery teachers select children to receive tutoring from among all first-grade children in their schools. First, classroom teachers identify the lowest-performing children in their classrooms by completing alternative rankings (see chapter 3 for a description of the alternate ranking). Next, Reading Recovery teachers individually

administer *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to the bottom 20% of children placed on these rankings. Then, Reading Recovery teachers select children with the lowest scores on Clay's assessment to receive tutoring, in accordance with the *Standard and Guidelines of the Reading Recovery Council of North America* (see RRCNA, 2001).

To discontinue tutoring, Reading Recovery teachers use "several 'relative' criteria" to determine whether their children (a) respond well to tutoring and meet the criteria for successful performance or (b) do not respond well to tutoring and fail to meet the criteria (Clay, 1993, p. 60). Reading Recovery teachers consider children's development of self-extending systems, the self-regulation of literacy processing that continues to improve as children read and/or write texts. They also consider children's performance on the tasks of *An Observation Survey of Early Literacy Achievement* (Clay, 2002). However, Clay's relative criteria do not specify exact scores or a range of scores that children must achieve on these tasks to meet the criteria for successful performance. Additionally, Reading Recovery teachers, in consultation with the children's classroom teachers, consider children's performance in literacy instruction in their classrooms.

If Reading Recovery children respond well to tutoring and meet the criteria for successful performance, their Reading Recovery teachers assign the end-of-program status category of discontinued. Conversely, if Reading Recovery children do not respond well to tutoring and fail to meet the criteria, their Reading Recovery teachers (a) assign the end-of-program status category of recommended and (b) refer them to receive additional assessment and/or consideration for other supplemental instruction. According to Clay (2001), the assignment of individual children to the recommended category 'is a

positive outcome for both the child and the education system” (p. 218).

Other end-of-program status categories include moved, none of the above, and incomplete. The moved category refers to children who moved during tutoring. The none of the above category refers to children who were removed from tutoring by Reading Recovery teacher leaders due to unusual circumstances, such as parent or guardian request. The incomplete category refers to children with less than 20 weeks of tutoring because the conclusion of the school year ended their tutoring.

Reading Recovery teachers refer to the first cohort of Reading Recovery children to receive tutoring at the beginning of a school year as first-round children. First-round children generally receive tutoring for approximately 20 weeks. As individual first-round children complete their tutoring, teachers select new children to tutor. Teachers refer to this second cohort of children to receive tutoring during a school year as second-round children. The majority of second-round children complete their tutoring by the end of the school year. However, if time remains in the school year, teachers may select new children to tutor as individual second-round children complete their tutoring. If teachers select new children to tutor, they refer to this third cohort of children to receive tutoring during a school year as subsequent- or third-round children.

During the 2001-2002 school year, Reading Recovery teachers assigned 32% of all first-round children in the United States to the recommended category, whereas these teachers assigned 2% of all second-round children and 0% of all subsequent-round children to this category (Gómez-Bellengé et al., 2003, Table S26.RR). Reading Recovery teachers generally assign first-round children to the recommended category upon their completion of 20 weeks of tutoring. Reading Recovery teachers generally

assign second- and subsequent-round children to the incomplete category rather than the recommended category because they receive less than 20 weeks of tutoring due to the end of the school year.

Nearly one third of first-round Reading Recovery children in the United States do not respond well to tutoring (Gómez-Bellengé et al., 2003). That is, they fail to meet the criteria for successful performance and their Reading Recovery teachers recommend them for additional assessment and/or supplemental instruction. However, little is known about the early literacy development of these recommended children. To date, only a few studies have reported on and/or investigated their early literacy development (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004).

#### *Early Literacy Development*

Theoretical models of early literacy development, supported by programs of research from a variety of epistemological and methodological perspectives, conceptualize early literacy development in different ways (Adams, 1990; Clay, 1991, 2001; Stanovich, 1986). Many factors influence these conceptualizations, including cognitive, linguistic, perceptual, biological, instructional, sociocultural, economic, and political factors (Snow et al., 1998; Vellutino & Scanlon, 2001). Theorists propose that reading and writing are complex processes in which various knowledge sources interact simultaneously and operate in parallel while mental strategies are initiated and employed in children' minds. This parallel processing occurs as children construct and/or produce meaning during the reading and/or writing of texts, respectively. (e.g., Rumelhart, 1994; Stanovich, 1980). Theorists also propose that early literacy development progresses along a fairly predictable continuum, marked by individual differences (e.g., Stanovich, 1986).

Prior to formal literacy instruction, children advance in their language proficiency, literacy knowledge, and literacy processing due to their experiences at home and/or in school. Sulbzy and Teale (1991) referred to children's acquisition of these precursors as emergent literacy. Children's acquisition of these precursors influences their transitions from emergent to conventional literacy.

Many factors influence early literacy development. Child-based factors, such as early language impairment, relate to poor progress in early literacy development. Family-based factors, such as home literacy environment, also relate to children's struggles in early literacy development. Furthermore, community-based factors, such as low-performing schools, relate to poor progress. Although these factors correlate with early literacy development, none alone explains it (Snow et al., 1998; Scarborough, 1998).

Based on her research grounded in systematic observations of children's reading and writing behaviors, Clay (1991, 2001) authored several tenets underscoring her theory of early literacy development. First, Clay theorized that children construct literacy processing systems through the development and integration of various knowledge sources and 'in-the-head' strategies. Second, she asserted that children advance in their ability to use various knowledge sources and employ 'in-the-head' strategies with the assistance of more capable and knowledgeable others. Third, Clay theorized that the reading and writing of texts provide children with opportunities to orchestrate their developing literacy processing systems. Fourth, she proposed that the reciprocal nature of reading and writing contribute to children's early literacy development. Fifth, she asserted that individually designed diagnostic instruction accelerates the early literacy development of struggling children to close the achievement gap between them and their

more accomplished peers. Finally, Clay contended that individual differences lead children to take different pathways to early literacy development.

Reading Recovery children receive individually designed diagnostic instruction that accommodates their individual differences and accelerates their early literacy development. Yet, as previously noted, not all Reading Recovery children progress as expected. A sizable proportion of first-round children do not respond well to tutoring, as evidenced by their failure to meet the criteria for successful performance, their Reading Recovery teachers' decisions to assign them to the recommended end-of-program status category, and their Reading Recovery teachers' recommendations for them to receive additional assessment and/or supplemental instruction (Gómez-Bellengé et al., 2003).

While a host of the aforementioned precursors and factors may play a role in recommended children's early literacy development, this study focused on their (a) phonological awareness and orthographic knowledge prior to and following their tutoring and (b) oral reading processing and reading comprehension processing following their tutoring.

### *Phonological Awareness*

How do recommended children perform in comparison to discontinued children in their awareness of syllables, onsets, rimes, phonemes in spoken words prior to and following their tutoring? The current study compares recommended to discontinued children on phonological awareness in terms of overall phonological awareness in the form of a composite and the components that formulated this composite, including rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness (see Table 1).

Table 1

*Early Literacy Knowledge and Processing, Corresponding Components, and Testing Occasions*

Early Literacy Knowledge and Processing	Corresponding Component	Testing Occasion
Phonological Awareness	Overall Phonological Awareness Composite	Pre- and Post-tutoring
	Rhyme Awareness	Pre- and Post-tutoring
	Phonological Skeletal Structure Awareness	Pre- and Post-tutoring
	Combined Syllable, Onset-Rime, & Phonemic Awareness	Pre- and Post-tutoring
	Graphophonemic Awareness	Pre- and Post-tutoring
Orthographic Knowledge	Overall Orthographic Knowledge Composite	Pre- and Post-tutoring
	Spelling Knowledge	Pre- and Post-tutoring
	Orthographic Pattern Knowledge	Pre- and Post-tutoring
	Orthographic Acceptability Knowledge	Pre- and Post-tutoring



Table 1 continued

*Early Literacy Knowledge and Processing, Corresponding Components, and Testing Occasions*

---

Early Literacy Knowledge and Processing	Corresponding Component	Testing Occasion
Oral Reading Processing	Oral Reading Accuracy	Post-tutoring
	Oral Reading Inaccuracy	Post-tutoring
	Oral Reading Behaviors	Post-tutoring
Reading Comprehension Processing	Text-related Questions	Post-tutoring

---

Phonological awareness, the ability to recognize and manipulate the sound segments of syllables, onsets, rimes, and phonemes in spoken words, influences the complex process of learning to read and write. Yet attainment of phonological awareness is not sufficient for learning to read and write. Children begin their early literacy development with varying degrees of phonological awareness with some children unaware that spoken words contain syllables, onset, rimes, and phonemes. Phonemic awareness, the ability to recognize and manipulate individual sounds (i.e., phonemes) in spoken words, is a component of phonological awareness. Because of the alphabetic structure of the English language, phonemic awareness is fundamental to early literacy development. (National Reading Panel, 2000; Snow et al., 1998).

Clay designed the Reading Recovery lesson components and corresponding teaching procedures to develop children's phonemic and graphophonemic awareness. During tutoring lessons, children learn to write unfamiliar words by articulating the words slowly, listening for the phonemes in these spoken words, retrieving the phonemes and their corresponding graphemes from their lexicons, and writing the graphemes. That is, they learn to make the words 'sound right'. Children also learn to link graphemes to phonemes when they decode unfamiliar words during the reading and/or rereading of texts and the rereading of texts that they have written. Additionally, children learn to link graphemes to phonemes when they participate in isolated word analysis activities (Clay, 1993). For example, they learn to remove single onsets from familiar words, add different single onsets to make unfamiliar words, and read these unfamiliar words.

### *Orthographic Knowledge*

How do recommended children perform in comparison to discontinued children in

their knowledge of letter sequences and/or patterns in the words that they write prior to and following their tutoring? The current study compares recommended to discontinued children on orthographic knowledge in terms of overall orthographic knowledge in the form of a composite and the components that formulated this composite, including spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge (see Table 1).

Orthographic knowledge of the English language underlies early literacy development because it facilitates efficient and automatic perception of words during reading and generation of words during writing (Templeton & Morris, 2000). Orthographic knowledge includes knowledge of (a) phoneme-grapheme correspondences, (b) letter sequences and/or patterns in words, and (c) the spelling and meaning relationships among words (Henderson & Templeton, 1986). In this study, orthographic knowledge refers to children's knowledge of English letter sequences and/or patterns in words.

Clay designed the Reading Recovery lesson components and corresponding teaching procedures to develop children's orthographic knowledge. During tutoring lessons, children learn to spell unfamiliar words by retrieving letter sequences and/or patterns and their corresponding graphemes from their lexicons when they write of self-composed sentences. That is, they learn to make the words 'look right' and 'sound right'. Children also learn to search for and use letter sequences and/or patterns in unfamiliar words when they decode unfamiliar words or analogize them to familiar words during the reading and/or rereading of texts and the rereading of texts that they have written. In addition, they learn to use letter sequences and/or patterns when they participate in

isolated word analysis activities (Clay, 1993). For example, they learn to remove the rimes from familiar words, add different rimes to make unfamiliar words, and read these unfamiliar words.

### *Oral Reading Processing*

How do recommended children perform in comparison to discontinued children in their ability to read texts orally following their tutoring? The current study compares recommended to discontinued children on oral reading processing in terms of (a) oral reading accuracy; (b) oral reading inaccuracy that includes the number of overall errors, substitutions, tolds, omissions, and insertions; and (c) oral reading behaviors that includes rate, fluency, and the number of self-corrections and repetitions (see Table 1).

As children develop their ability to read words in text, they learn to read familiar words by sight. They also learn to read unfamiliar words by decoding, analogizing to other familiar words stored in their lexicons, and using context to predict, such as supportive pictures, semantic knowledge, and/or syntactic knowledge. As children develop their ability to read words in text, they read them rapidly and accurately. Children's reading accuracy is higher when they read words in text than when they read them in isolation because the oral reading processing interacts with reading comprehension processing (e.g., Nicholson, 1991).

Clay designed the Reading Recovery lesson components and corresponding teaching procedures to develop children's oral reading processing. During tutoring lessons, children learn to read familiar words by sight and unfamiliar words by decoding, analogizing, and using context when they read and/or reread texts, reread texts that they have written, and complete word analysis activities (Clay, 1993). For example, they read

words that appear frequently in texts that their teachers isolate. This process of reading isolated words quickly and repeatedly out of context builds children's sight word vocabulary so they read familiar words by sight in a fast and accurate manner. During tutoring lesson, children also learn to (a) look at the pictures and the first letters of unfamiliar words contained in the texts and (b) think of meaningful words that match both the pictures and first letters. This process of cross-checking context with first letters enables children to read unfamiliar words.

### *Reading Comprehension Processing*

How do recommended children perform in comparison to discontinued children in their ability to respond orally to text-related questions following their tutoring? The current study compares recommended to discontinued children on reading comprehension processing (see Table 1).

Reading comprehension processing involves children constructing meaning through their interactions with texts. This processing involves children (a) building mental representations by accessing the information or content contained directly in the texts and integrating this information or content with their background knowledge (b) using conceptual knowledge and knowledge of word meanings, and (c) applying reading comprehension strategies and reading comprehension monitoring (Snow et al., 1998). When children apply reading comprehension strategies, such as identifying story structure elements and drawing inferences, they engage in regulation processing to develop, maintain, and/or increase their understanding of the texts (Pearson & Duke, 2002). When children monitor their reading comprehension, they engage in evaluation and regulation processing to determine whether they understand the texts and to activate

reading comprehension strategies to develop, maintain, and/or increase their understanding (Baker & Brown, 1984).

Clay designed the Reading Recovery lesson components and corresponding teaching procedures to foster children's reading comprehension monitoring. During their tutoring lessons, children learn to monitor their reading comprehension through the development of evaluation and regulation processing when they read and/or reread texts and reread texts that they have written (Clay, 1993). For example, they learn to reread words, phrases, and/or sentences in texts to construct, confirm, and/or disconfirm meaning. They also learn to search for and use semantic cues to self-correct oral reading attempts that do not make sense.

#### *Early Literacy Instruction*

To ameliorate low-performing, six-year-old children's reading and writing difficulties, Clay designed the Reading Recovery teaching procedures to develop various aspects of their early literacy development, including phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. In a synthesis of research, Snow et al. (1998) made instructional recommendations to prevent the occurrence and/or advancement of early reading difficulties without privileging a particular instructional method or program. Although they acknowledged that some children are at an elevated risk for difficulties, their recommendations extended to all children. They argued that there is little evidence indicating that high-risk children who experience difficulties, even those with learning disabilities, require substantially different instruction than low-risk children.

Snow et al. (1998) recommended that classroom teachers provide adequate initial

instruction that focuses on “using reading to obtain meaning from print; the sublexical structure of spoken words (phonological and morphological components of words); the nature of the orthographic system (letters and their sequences in words); the specifics of frequent, regular spelling-sound relationships; frequent opportunities to read; and opportunities to write” (p. 314). Furthermore, they asserted that adequate progress beyond the initial level depends on “having established a working understanding of how sounds are represented alphabetically; sufficient practice in reading to achieve fluency with different kinds of texts written for different purposes; and control over procedures for monitoring comprehension and repairing misunderstandings” (p. 314).

Acknowledging that some children need more literacy instruction beyond the instruction provided in their classrooms, Snow et al. (1998) further recommended that reading and/or special education teachers provide these children with supplemental instruction, such as tutoring. Reading Recovery children receive supplemental instruction in the form of tutoring in addition to their regular classroom instruction. Even though Reading Recovery children receive tutoring to accelerate their early literacy development, some children do not progress as expected. Their failure to respond well to their tutoring poses inquiry into the effectiveness of tutoring.

### *Tutoring Effectiveness*

Recent reviews of tutoring studies (e.g., Elbaum, Vaughn, Hughes, & Moody, 2000) suggest that overall tutoring provided by trained teachers or community-based volunteers, such as college students, is generally effective in advancing the early literacy development of young children, as well as older children identified as struggling literacy learners. These reviews have reported generally moderate effects upon children’s

immediate completion of tutoring when comparing the performance of tutored children in treatment groups to the performance of children in comparison and/or control groups who did not receive treatment. The duration, length, and frequency of tutoring varied, ranging from two weeks to two and a half years, 15 to 60 minutes, and two to five times a week, respectively. The grade levels of tutored children also varied, ranging from first- through sixth-grades with the preponderance of children in first-grade. Several reviews also have examined the long term effects of tutoring and concluded that overall tutored children in treatment groups maintained more of their literacy gains than children in comparison and/or control groups who did not receive treatment in the years following tutoring (e.g., Wasik & Slavin, 1993). However, several reviews (e.g., Elbaum et al., 2000) have noted that some studies possessed research design limitations, such as failure to assign children randomly or match children to equivalent control groups, and have argued that these limitations inflated results. These reviews also have noted that some of the effect sizes corresponding to individual tutoring studies were small to moderate, thereby questioning the practical significance of the results (Shanahan, 1998).

Although some reviews have examined studies that investigated the effectiveness of several tutoring programs, including Reading Recovery (Elbaum et al., 2000; Wasik & Slavin, 1993), other reviews exclusively examined studies that investigated the effectiveness of Reading Recovery (Hiebert, 1994; Shanahan & Barr, 1995). These reviews not only reported effect sizes from studies published in peer-reviewed journals, but also conducted independent, secondary analyses of unpublished documents (Hiebert, 1994; Shanahan & Barr, 1995).

Overall, the reviews have regarded Reading Recovery tutoring as generally



effective based on the Reading Recovery children's literacy gains from the beginning to the end of tutoring, as evidenced by large to moderate effect sizes (e.g., Shanahan & Barr, 1995; Wasik & Slavin, 1993). Furthermore, the reviews have concluded that Reading Recovery children maintained their literacy gains in the years following tutoring; however, the reviews have noted diminishing effects from first-to fourth-grades (e.g., Shanahan & Barr, 1995; Wasik & Slavin, 1993). Although these reviews have considered Reading Recovery tutoring generally effective, they have noted that some of the studies possessed research design limitations, such as failure to analyze data and report results corresponding to all tutored children, and have argued that these limitations inflated results (e.g., Shanahan & Barr, 1995).

#### *Reading Recovery Tutoring Responsiveness*

Despite the overall consensus that Reading Recovery tutoring is effective in accelerating the lowest-performing children's early literacy development, Reading Recovery tutoring does not accelerate the early literacy development of all children who receive it. During the 2001-2002 school year, 32% of first-round Reading Recovery children in the United States who received 30 minutes of daily tutoring for approximately 20 weeks from highly trained Reading Recovery teachers failed to respond well to it (Gómez-Bellengé et al., 2003). Specifically, these Reading Recovery children did not meet the criteria for successful performance. Their Reading Recovery teachers assigned them to the recommended category and referred them to receive additional assessment and/or consideration for other supplemental instruction.

Interestingly, despite the fact that nearly one third of first-round Reading Recovery children continue to struggle in their early literacy development upon their

completion of tutoring, only four studies have examined the early literacy development of these children separate from discontinued children (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004). Specifically, these studies examined recommended children's phonemic awareness (Center et al., 1995; Chapman et al., 2001; Spector & Moore, 2004) and graphophonemic awareness (Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004), two components of phonological awareness. One study investigated recommended children's spelling knowledge (Chapman et al., 2001), a component of orthographic knowledge. Three studies examined recommended children's oral reading accuracy (Clay & Tuck, 1991; Chapman et al., 2001; Spector & Moore, 2004), a component of oral reading processing. Another study investigated recommended children's ability to respond correctly to text-related questions (Chapman et al., 2001), a component of reading comprehension processing. From among these four studies, only Spector and Moore tested for statistically significant differences between recommended and discontinued children at pre-tutoring. None of these studies tested for statistically significant differences between these two groups at post-tutoring. Additionally, only Center et al. and Chapman et al. administered measures other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002). Spector and Moore administered Clay's measure, as well as other early literacy measures independent of Reading Recovery.

### *Research Questions*

To describe the early literacy development of recommended children, I proposed the following overarching research question: How do recommended Reading Recovery children compare to discontinued Reading Recovery children on (a) phonological

awareness and orthographic knowledge prior to and following their completion of tutoring and (b) oral reading processing and reading comprehension processing following their completion of tutoring? To answer this overarching research question, I developed four subsidiary questions:

- (a) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on phonological awareness prior to and following their completion of tutoring?
- (b) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring?
- (c) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring?
- (d) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring?

### *Significance*

The current study contributes to an emerging program of research that has examined bits and pieces of recommended children's early literacy development. First, this investigation assesses multiple components of recommended children's phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing (see Table 1). Although previously conducted studies assessed one or two components of phonological awareness, orthographic knowledge, oral reading

processing, or reading comprehension processing (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004), no one study measured multiple components of all of them. Second, this investigation employs several statistical analyses at both pre- and post-tutoring to compare recommended to discontinued children on phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. Whereas only one earlier study conducted statistical analyses at pre-tutoring (Spector & Moore, 2004), no other studies conducted such analyses at both pre- and post-tutoring. Third, this study administers early literacy measures other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to recommended children, thereby assessing components of phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing not assessed by Clay's measure. Whereas two previously conducted studies employed Clay's measure (Clay & Tuck, 1991; Spector & Moore, 2004), only two other studies employed measures other than Clay's measure (Center et al., 1995; Chapman et al., 2001). Fourth, this study compares recommended children to normative sample children on oral reading processing and reading comprehension processing. Although one previously conducted study compared recommended children to normative sample children on reading comprehension processing (Chapman et al., 2001), no other study made this comparison on both oral reading processing and reading comprehension processing. Additionally, this current study contributes to the understanding of recommended children's early literacy development by explaining how their phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing relate to an interactive model of early literacy development.

### *Overview of the Methodology*

To answer the current study's research questions, I individually administered several early literacy measures to Reading Recovery children to assess their (a) phonological awareness and orthographic knowledge prior to and following their tutoring and (b) oral reading processing and reading comprehension processing following their tutoring. I conducted this study in a single school district. All the schools with Reading Recovery tutoring in this district accepted the invitation to participate with the exception of one school. At the start of the current study, the sample consisted of 60 children who were taught by 15 trained Reading Recovery teachers and one Reading Recovery teacher leader. By the end of this study, the sample consisted of 55 children: 29 recommended children and 26 discontinued children.

Prior to the current study, I conducted two pilot studies in which I administered (a) modified phonological awareness tasks to a sample of kindergarten children at the end of their school year and (b) complete phonological awareness tasks to a sample of first-round Reading Recovery children prior to and following their tutoring (see Chapter 2 for a summary and Appendix A for a complete report of these pilot studies). Based on the results from these pilot studies, I selected phonological awareness tasks for the current study that were neither too easy nor too difficult for recommended children to complete prior to and following their tutoring.

The Reading Recovery teachers selected children to receive Reading Recovery tutoring according to the guidelines set forth by RRCNA (2001). Then, I administered the Rhyme Detection task (Muter, Hulme, & Snowling, 1997), the Blending Words task (Wagner, Torgesen, & Rashotte, 1999), and the Sentence Writing and Spelling task

(DeFord, 2000) to assess the children's phonological awareness and orthographic knowledge.

After these Reading Recovery teachers delivered approximately 20 weeks of tutoring to their Reading Recovery children, they determined whether their children (a) responded well to tutoring and subsequently assigned them to the discontinued category or (b) failed to respond well to tutoring and subsequently assigned them to the recommended category. Then, I re-administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) to reassess the children's phonological awareness and orthographic knowledge. At this time, I also administered the Gray Oral Reading Test-Fourth Edition (GORT-4) (Wiederholt & Bryant, 2001) to assess the children's oral reading processing and reading comprehension processing.

To describe the recommended children's phonological awareness and orthographic knowledge at pre- and post-tutoring and oral reading processing and reading comprehension processing at post-tutoring, I compared their performance on the various early literacy measures to the comparable performance of the discontinued children. I reported means, standard deviations, and percentages. Then, I conducted a number of statistical tests on these descriptive statistics, including repeated measure two-way analyses of variance, chi-square tests of independence, and one-way analyses of variance.

To further describe the recommended children's oral reading processing and reading comprehension processing, I compared their GORT-4 accuracy, rate, fluency, comprehension, and composite oral reading and reading comprehension processing standard score performance to the discontinued children's comparable performance.

Additionally, I compared the recommended children's, as well as the discontinued children's composite oral reading and reading comprehension processing standard score performance to the GORT-4's normative sample distribution.

#### *Rationale for the Selection of Early Literacy Measures*

For the current study, I selected early literacy measures other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) for two specific reasons. First, a relationship exists between *An Observation Survey of Early Literacy Achievement* and the Reading Recovery program (Wasik & Slavin, 1993). Second, *An Observation Survey of Early Literacy Achievement* fails to assess specific components of phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing (Hiebert, 1994).

Clay designed and developed *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and Reading Recovery. Reading Recovery teachers administer Clay's assessment when selecting children to receive tutoring and discontinuing them. Thus, a relationship exists between Clay's assessment and Reading Recovery tutoring (Wasik & Slavin, 1993). Because of this relationship, I selected early literacy measures for the current study other than *An Observation Survey of Early Literacy Achievement*.

Although *An Observation Survey of Early Literacy Achievement* (Clay, 2002) assesses children's (a) letter knowledge (i.e. Letter Identification task), (b) knowledge of printed language conventions (i.e., Concepts About Print task), (c) reading vocabulary (i.e., Word Reading task), (d) writing vocabulary (i.e., Writing Vocabulary task), (e) graphophonemic awareness (i.e., Hearing and Recording Sounds in Words task), and (f) oral reading accuracy (i.e., Text Reading task), it remains limited in scope (Hiebert,

1994). The Hearing and Recording Sounds in Words task assesses children's ability to hear phonemes in spoken words, match the phonemes to the corresponding graphemes, and record the graphemes, but it fails to assess children's (a) awareness of the larger sound segments of syllables, onsets, and rimes in spoken words and (b) knowledge of orthographic letter sequences and/or patterns. The Text Reading task assesses children's oral reading processing in terms of accuracy, but it fails to assess formally oral reading processing in terms of rate and fluency (i.e., combined rate and accuracy). This task also fails to assess formally the children's reading comprehension processing. Because of these limitations, I selected early literacy measures for the current study other than *An Observation Survey of Early Literacy Achievement*.

To assess the children's phonological awareness and orthographic knowledge prior to and following their tutoring, I selected and administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000). To assess the children's oral reading processing and reading comprehension processing following their tutoring, I selected and administered the GORT-4 (Wiederholt & Bryant, 2001). Through the employment of these measures, I gained insight into the recommended children's early literacy development that *An Observation Survey of Early Literacy Achievement* (Clay, 2002) fails to provide.

### *Definitions*

*Analogizing.* Analogizing is a child's ability to read an unfamiliar word by (a) recognizing a specific orthographic letter sequence, (b) matching this letter sequence to a familiar word stored in his/her lexicon, (c) pronouncing the matched letter sequence in



the same manner as the letter sequence stored in his/her lexicon, and (d) blending the matched letter sequence to the other letters in the unfamiliar word (see definitions corresponding to blending and lexicon).

*Appeal.* An appeal is a child's request for help with a word while reading a text. The child directs the appeal to the examiner. An appeal for help from a child is followed by a told from the examiner (see definition corresponding to told).

*Blending.* Blending is a sound analysis skill that involves the oral combination of syllables, onsets, rimes, and/or phonemes to form a word (see definitions corresponding to onset, phoneme, rime, and syllable).

*Combined syllable, onset and rime, and phonemic awareness.* Combined syllable, onset and rime, and phonemic awareness is a child's ability to blend isolated syllables, onsets and rimes, or phonemes into words. Combined syllable, onset and rime, and phonemic awareness is a component of phonological awareness (see definitions corresponding to blending, onset, onset and rime awareness, phoneme, phonemic awareness, phonological awareness, rime, syllable, and syllable awareness).

*Decoding.* Decoding is a child's ability to read an unfamiliar word by (a) either sounding the graphemes and blending their corresponding phonemes and/or locating orthographic letter sequences and/or patterns to generate an approximate pronunciation and (b) subsequently searching for a meaningful word that matches the pronunciation in his/her lexicon. Decoding is also known as phonological recoding (see definitions corresponding to blending, grapheme, lexicon, phoneme, and phonological recoding).

*Deletion.* Deletion is a sound analysis skill that involves the oral elimination of a syllable, onset, rime, or phoneme from a word (see definitions corresponding to onset, phoneme

rime, and syllable).

*Digraph.* A digraph is the sequence of two graphemes that represent a single phoneme. A digraph is either a consonant digraph (e.g., *th* in *think*) or a vowel digraph (e.g., *ai* in *said*) (see definitions corresponding to grapheme and phoneme).

*Discontinued.* Discontinued is a Reading Recovery end-of-program status category that refers to a Reading Recovery child who meets the criteria for successful tutoring completion.

*Doublet.* A doublet is the sequence of two identical graphemes that represent a single phoneme. A doublet is either a consonant doublet (e.g., *tt* in *kitten*) or a vowel digraph (e.g., *oo* in *good*) (see definitions corresponding to grapheme and phoneme).

*Error.* An error is a child's inaccurate oral reading response while reading a text. An error is a substitution, a told, an omission, or an insertion. An error may or may not include a self-correction or a repetition depending upon a measure's scoring procedures (see definitions corresponding to insertion, omission, oral reading, repetition, self-correction, substitution, and told).

*First-round children.* First-round children are the first cohort of children to receive Reading Recovery tutoring at the beginning of a school year.

*Fluency.* Fluency is the number of seconds a child takes to read a text and the number of errors a child makes while reading a text (see definition corresponding to error).

*Grapheme.* Grapheme is a segment of a written word, either a single letter or a group of letters, that represents a phoneme in the spelling of a word (see definition corresponding to phoneme).

*Graphemic alternation.* Graphemic alternation is the occurrence of certain graphemes in

some spellings and the occurrence of other graphemes in other spellings due to the position of graphemes in words, as well as the phonological structures of words (e.g., the grapheme *y* tends to occur at the ends of morphemes, whereas the grapheme *i* tends to occur at the beginnings and in the middles of words (e.g., *dinosaur*) (see definitions corresponding to grapheme and morpheme).

*Graphophonemic awareness.* Graphophonemic awareness is a child's ability to hear phonemes in a spoken word, match the phonemes to the corresponding graphemes, and record the graphemes. Graphophonemic awareness is an extension of phonemic awareness. Graphophonemic awareness is a component of phonological awareness (see definitions corresponding to grapheme, phoneme, phonemic awareness, and phonological awareness).

*Incomplete.* Incomplete, also known as 'Incomplete Program at Year-End', is a Reading Recovery end-of-program status category that refers to a Reading Recovery child who receives tutoring until the end of the school year but does not complete his/her tutoring due to the end of the school year.

*Insertion.* An insertion is a child's addition of a word to a text while reading. An insertion is an error (see definition corresponding to error).

*Knowledge source.* A knowledge source is a store of information, such as phonological awareness and knowledge of the world, that interacts simultaneously and operates in parallel with other knowledge sources in a child's mind during the reading and writing of a text (see definitions corresponding to phonological awareness).

*Lexicon.* The lexicon is a child's mental dictionary that stores phonemes, graphemes, letter sequences, letter patterns, words, word meanings, and pronunciations (see

definitions corresponding to grapheme and phoneme).

*Literacy knowledge.* Literacy knowledge refers to a child's awareness and/or understanding of literacy information, such as phonological awareness and orthographic knowledge (see definitions corresponding to phonological awareness and orthographic knowledge).

*Literacy processing.* Literacy processing refers to the operations or actions occurring in a child's mind during the reading and writing of a text, including accessing knowledge sources and using strategies (see definitions corresponding to knowledge source and strategy).

*Morpheme.* A morpheme is the smallest meaning-based unit in a word (e.g., *boys* has two morphemes, *boy* and *s*). A free morpheme is a word (e.g., *boy*); a bound morpheme is an affix that attaches to a word (e.g., *s* in *boys*).

*Morphology.* Morphology refers to the study of the structure and form of words in a language, including inflection, derivation, and the formation of compounds.

*Moved.* Moved, also known as 'Moved While Being Served', is a Reading Recovery end-of-program status category that refers a Reading Recovery child who moves before his/her completion of tutoring.

*None of the Above.* None of the Above is a Reading Recovery end-of-program status category that refers to a Reading Recovery child removed from his/her tutoring by the Reading Recovery teacher leader prior to his/her completion of tutoring due to an unusual circumstance, such as a parent or guardian request (see definition corresponding to teacher leader).

*Omission.* An omission is a child's deletion of a word from a text while reading. An

omission is an error (see definition corresponding to error).

*Onset.* Onset is the initial consonant(s) that precedes the vowel in a spoken or written syllable (e.g., *sn* is the onset in *snow*; *m* is the onset in the first syllable *mel* in *mellow* and *l* is the onset in the second syllable *low* in *mellow*) (see definition corresponding to syllable).

*Onset and rime awareness.* Onset and rime awareness is the sensitivity to and/or knowledge of (a) the onset and rime in a spoken and/or written single syllable word (e.g., *sn* is the onset and *ow* is the rime in *snow*) or (b) the onset and rime in each syllable in a spoken or written multisyllabic word (e.g., *m* is the onset and *el* is the rime of the first syllable *mel* in *mellow*; *l* is the onset and *ow* is the rime in the second syllable *low* in *mellow*). Onset and rime awareness is the same as phonemic awareness when the onset and rime in a spoken word are single phonemes (e.g., *g* is the onset and *o* is the rime in *go*). Onset and rime awareness is a component of phonological awareness when a word is spoken and a component of orthographic knowledge when a word is written (see definitions corresponding to onset, orthographic knowledge, phoneme, phonemic awareness, phonological awareness, rime, and syllable).

*Oral reading accuracy.* Oral reading accuracy refers to the number of correct oral reading responses a child makes while reading a text. Oral reading accuracy may or may not include repetitions or self-corrections depending on a measure's scoring procedures. Oral reading accuracy is a component of oral reading processing (see definitions corresponding to oral reading processing, repetition, and self-correction).

*Oral reading behavior.* Oral reading behavior refers to the oral reading outcomes a child yields while reading a text. Oral reading behavior (a) includes rate and fluency and (b)

may or may not include repetitions or self-corrections depending on a measure's scoring procedures. Oral reading behavior is a component of oral reading processing (see definitions corresponding to fluency, oral reading processing, rate, repetition, and self-correction).

*Oral reading inaccuracy.* Oral reading inaccuracy refers to the number of incorrect oral reading responses or errors a child makes while reading a text. Oral reading inaccuracy (a) includes substitutions, tolds, omissions, and insertions and (b) may or may not include repetitions or self-corrections depending on a measure's scoring procedures. Oral reading inaccuracy is a component of oral reading processing (see definitions corresponding to error, insertion, omission, oral reading processing, repetition, self-correction, substitution, and told).

*Oral reading processing.* Oral reading processing is a child's ability to read familiar and unfamiliar words in a text. A child often reads familiar words in text by sight and unfamiliar words in texts by decoding, analogizing, or using context. Oral reading processing differs from word reading processing in that oral reading processing involves the reading of a text. Oral reading processing includes the following components: oral reading accuracy, oral reading inaccuracy, and oral reading behavior (see definitions corresponding to analogizing, decoding, oral reading accuracy, oral reading inaccuracy, oral reading behavior, using context, and word reading processing).

*Orthographic acceptability knowledge.* Orthographic acceptability knowledge is a child's ability to record an acceptable English letter sequence in the writing of a word.

Orthographic acceptability knowledge is a component of orthographic knowledge (see definition corresponding to orthographic knowledge).

*Orthographic knowledge.* Orthographic knowledge is a child's understanding of the English letter sequences and/or patterns in a word that s/he writes. Orthographic knowledge includes the following components: spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge (see definitions corresponding to orthographic acceptability knowledge, orthographic pattern knowledge, and spelling knowledge).

*Orthographic pattern knowledge.* Orthographic pattern knowledge is a child's ability to write an English word with one or more of the following letter patterns: vowel digraphs, double consonants, inflectional endings, consonant digraphs, r-controlled vowels, silent letters, and consonant blends. Orthographic pattern knowledge is a component of orthographic knowledge (see definitions corresponding to digraph and orthographic knowledge).

*Overall orthographic knowledge.* Overall orthographic knowledge is a composite that consists of a child's spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge (see definitions corresponding to orthographic acceptability knowledge, orthographic knowledge, orthographic pattern knowledge, and spelling knowledge).

*Overall phonological awareness.* Overall phonological awareness is a composite that consists of a child's rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness (see definitions corresponding to combined syllable, onset and rime, and phonemic awareness; graphophonemic awareness; phonological awareness; phonological skeletal structure awareness; and rhyme awareness).

*Phoneme.* Phoneme is the smallest sound segment in a spoken word that changes the meaning of the word (e.g., *snow* has three phonemes; *mellow* has four phonemes).

English contains 41 phonemes that are represented by graphemes (see definition corresponding to grapheme).

*Phonemic awareness.* Phonemic awareness is a child's ability to recognize and manipulate individual phonemes in a spoken word. Phonemic awareness encompasses onset and rime awareness when a spoken word has a single phoneme as an onset (e.g., *g* is the onset in *go*) and/or a single phoneme as a rime (e.g., *o* is the rime in *go*). Phonemic awareness is a component of phonological awareness (see definitions corresponding to onset, onset and rime awareness, phoneme, phonological awareness, and rime).

*Phonological awareness.* Phonological awareness is a child's ability to recognize and manipulate the sound segments of a spoken word, including syllables, rhymes, onsets, rimes, and phonemes. Phonological awareness includes the following components: rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness (see definitions corresponding to combined syllable, onset and rime, and phonemic awareness; graphophonemic awareness; onset; onset and rime awareness; phoneme; phonemic awareness; phonological skeletal structure awareness; rhyme; rhyme awareness; rime; syllable; and syllable awareness).

*Phonological recoding.* Phonological recoding is a child's ability to read an unfamiliar word by (a) either sounding the graphemes and blending their corresponding phonemes and/or locating orthographic letter sequences and/or patterns to generate an approximate pronunciation and (b) subsequently searching for a meaningful word that matches the



pronunciation in his/her lexicon. Phonological recoding is also known as decoding (see definitions corresponding to blending, decoding, grapheme, lexicon, and phoneme).

*Phonological skeletal structure awareness.* Phonological skeletal structure awareness is a child's ability to hear phonemes in a spoken word, match the phonemes to either consonant graphemes or vowel graphemes, and record the consonant-vowel structure.

Phonological skeletal structure awareness is a component of phonological awareness (see definitions corresponding to grapheme, phoneme, and phonological awareness).

*Rate.* Rate is the number of seconds a child takes to read a text.

*Reading comprehension processing.* Reading comprehension processing involves a child constructing meaning through his/her interactions with a text. This processing involves a child's (a) constructing a mental representation of a text by accessing the information or content contained directly in the text and integrating this information or content with his/her background knowledge, (b) accessing and using his/her conceptual knowledge and knowledge of word meanings, and (c) applying reading comprehension strategies and reading comprehension monitoring (see definitions corresponding to reading comprehension monitoring and reading comprehension strategies).

*Reading comprehension monitoring.* Reading comprehension monitoring involves a child (a) engaging in evaluation processing to determine whether s/he understands a text and (b) engaging in regulation processing to activate reading comprehension strategies to develop, maintain, and/or increase his/her understanding of a text. Reading comprehension monitoring is a component of reading comprehension processing (see definitions corresponding to reading comprehension processing and reading comprehension strategies).

*Reading comprehension strategies.* Reading comprehension strategies involve a child engaging in processing to develop, maintain, and/or increase his/her understanding of a text, such as drawing inferences and identifying story structure elements. Reading comprehension strategies is a component of reading comprehension processing (see definitions corresponding to reading comprehension processing and reading comprehension monitoring).

*Recommended.* Recommended, also known as ‘Recommended Action After a Full Program’, is a Reading Recovery end-of-program status category that refers to a Reading Recovery child who fails to meet the criteria for successful tutoring completion after receiving approximately 20 weeks of tutoring. A Reading Recovery teacher recommends the child for additional assessment and/or consideration of other instructional support.

*Repetition.* Repetition is a child saying a word, phrase, and/or sentence in a text more than once while reading. A measure may or may not consider a repetition an error depending upon its scoring procedures (see definition corresponding to error).

*Rhyme.* Rhyme is a sound segment that consists of a vowel(s) and the subsequent consonant(s) (i.e., rime) in a spoken single syllable word (e.g., *snow* rhymes with *blow*; *ow* is the rhyme and rime in *snow* and *blow*) or in each syllable of a spoken multisyllabic word (e.g., *mellow* rhymes with *bellow*; *el* and *ow* are the rhymes and rimes in *mellow* and *bellow*). A child may pronounce the rime in each syllable the same but may not necessarily spell it the same (e.g., *snow* rhymes with *go*; *ow* is the rime in *snow*, *o* is the rime in *go*) (see definitions corresponding to rime and syllable).

*Rhyme awareness.* Rhyme awareness is a child’s sensitivity to the sound categorization of two or more spoken single syllable or multisyllabic words in which the rimes in each

syllable are pronounced the same (e. g., *el* is the rime in the first syllable *mel* in *mellow* and *ow* is the rime in the second syllable *low* in *mellow*; *mellow* rhymes with *bellow*), but are not necessarily spelled the same at the end of each syllable (e.g., *ow* in the rime in *snow* and *o* is the rime in *go*; *snow* rhymes with *go*). Rhyme awareness is a component of phonological awareness (see definitions corresponding to phonological awareness, rhyme, rime, and syllable).

*Rime.* Rime is the vowel(s) and the subsequent consonant(s) in a spoken or written syllable (e.g., *ow* is the rime in *snow*; *el* is the rime in the first syllable *mel* in *mellow* and *ow* is the rime in the second syllable *low* in *mellow*). A child may spell the rime in each syllable the same but may not necessarily pronounce it the same (e.g., *ow* is the rime in *snow* and *plow*; *snow* and *plow* do not rhyme) (see definitions corresponding to syllable and rhyme).

*Running Record.* A running record is an assessment of oral reading processing developed by Marie Clay and used in Reading Recovery tutoring (see definition corresponding to oral reading processing).

*Second-round children.* Second-round children are the second cohort of children to receive Reading Recovery tutoring during a school year.

*Segmentation.* Segmentation is a sound analysis skill that involves breaking a word into individual syllables, onsets and rimes, and/or phonemes (see definitions corresponding to syllable, onset, rime, and phoneme).

*Self-correction.* A self-correction occurs when a child fixes a previously committed error in a text while reading. A measure may or may not consider a self-correction an error depending upon its scoring procedures (see definition corresponding to error).

*Self-extending system.* A self-extending system refers to a Reading Recovery child's self-regulation of his/her literacy processing that continues to improve as a child engages in text reading and/or writing activities.

*Spelling knowledge.* Spelling knowledge is a child's ability to write a word with correct spelling. Spelling knowledge is a component of orthographic knowledge (see definition corresponding to orthographic knowledge).

*Strategy.* A strategy is a mental operation or action, such as monitoring and evaluating, that a child initiates and engages in during the reading and writing of a text.

*Subsequent-round children.* Subsequent-round children are the third cohort of children to receive Reading Recovery tutoring during a school year (see definition corresponding to third-round children).

*Substitution.* A substitution is a child's replacement of a word with another word in a text while reading. A substitution is an error (see definition corresponding to error).

*Syllable.* Syllable is the largest sound segment in a spoken word that consists of a vowel(s) and any preceding and/or subsequent consonant(s) (e.g., *snow* has one syllable; *bellow* has two syllables).

*Syllable awareness.* Syllable awareness is a child's sensitivity to a syllable in a spoken single syllable word or the syllables in a spoken multisyllabic word. Syllable awareness is a component of phonological awareness (see definitions corresponding to phonological awareness and syllable).

*Teacher leader.* A teacher leader is a Reading Recovery trained teacher who oversees a Reading Recovery program one or more school districts. A teacher leader's responsibilities may include teaching Reading Recovery children, training teachers in the

Reading Recovery teaching procedures, providing continued professional development for trained Reading Recovery teachers, monitoring the progress of Reading Recovery children in a school district(s), and/or compiling and submitting outcome data to the National Data Evaluation Center (NDEC) at the end of a school year.

*Third-round children.* Third-round children are the third cohort of children to receive Reading Recovery tutoring during a school year (see definition corresponding to subsequent-round children).

*Told.* A told is a test examiner's pronunciation a word in a text for a child while the child reads the text. An examiner may provide a told when a child (a) appeals for help with a word after making an attempt to read the word and/or (b) balks at a word without making an attempt to read the word and fails to proceed in reading the text. A told is an error (see definitions corresponding to appeal and error).

*Using context.* Using context is a child's ability to read an unfamiliar word in a text by predicting it based on a picture that accompanies the text or the semantic and/or syntactic information contained in the preceding text.

*Word reading processing.* Word reading processing is a child's ability to read a word in isolation. A child may read a familiar word by sight or an unfamiliar word by decoding or analogizing. Word reading processing differs from oral reading processing in that oral reading processing involves the reading of a text (see definitions corresponding to analogizing, decoding, and oral reading processing).

### *Summary*

This current study assessed recommended Reading Recovery children's phonological awareness, orthographic knowledge, oral reading processing, and reading

comprehension processing. By administering the Rhyme Detection task (Muter et al., 1997), the Blending Word task (Wagner et al., 1999), the Sentence Writing and Spelling task (DeFord, 2000) to Reading Recovery children, I obtained data on their phonological awareness and orthographic knowledge prior to and following their tutoring. By administering the GORT-4 (Wiederholt & Bryant, 2001) to these same Reading Recovery children, I obtained data on their oral reading processing and reading comprehension processing following their tutoring. Then, I compared the recommended and discontinued children's (a) phonological awareness and orthographic knowledge prior to and following tutoring and (b) oral reading processing and reading comprehension processing following tutoring. Chapter 2 presents an interactive model of early literacy development and reviews research studies on phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. This chapter also reviews the research on Reading Recovery tutoring in relation to its effectiveness and children's responsiveness to it. Chapter 3 details this study's research methodology. Chapter 4 reports this study's results and provides a profile of an average-performing recommended child. Chapter 5 provides a study summary and includes a discussion of the major findings, limitations, and recommendations for instruction and research.

## Chapter 2

### *Early Literacy Development and Reading Recovery Tutoring*

Recommended Reading Recovery children may fail to develop adequately phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing, despite their participation in Reading Recovery tutoring. The following research questions guide the current study in an effort to provide a comprehensive description of their early literacy development:

- (a) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on phonological awareness prior to and following their completion of tutoring?
- (b) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring?
- (c) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring?
- (d) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring?

This chapter contains three main sections. Section 1 presents an interactive model of early literacy development. Section 2 discusses phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing and relates these aspects of early literacy development to the interactive model of early literacy

development that guides this study. Section 3 discusses the effectiveness of Reading Recovery tutoring and children's responsiveness to it.

*Interactive Model of Early Literacy Development*

The current investigation compares recommended to discontinued children on phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. Thus, it calls for a model that conceptualizes early literacy development. The interactive model of early literacy development that guides this study (see Figure 1) illustrates the complex processes that children engage in as they participate in early literacy activities.

In Figure 1, the center box represents the central processor in a child's mind that process information as s/he participates in an early literacy activity. When this processing occurs, various knowledge sources and mental strategies interact simultaneously and operate in parallel. The box to the left of the center box represents the early literacy activity, such as a text reading, text writing, or word analysis activity. The boxes above and below the center box represent some of the knowledge sources, such as lexical knowledge or phonological awareness, that a child brings to an early literacy activity or develops from early literacy experiences and instruction. The box to the right of the center box represents the early literacy outcome, such as phonological awareness, orthographic knowledge, oral reading processing, and/or reading comprehension processing. This model is an information-processing model. Thus, it considers the influence of cognitive, linguistic, and perceptual factors on early literacy development and neglects biological, instructional, sociocultural, economic, and political factors.

The model guiding this study is based on Ehri's (1998) interactive model of text



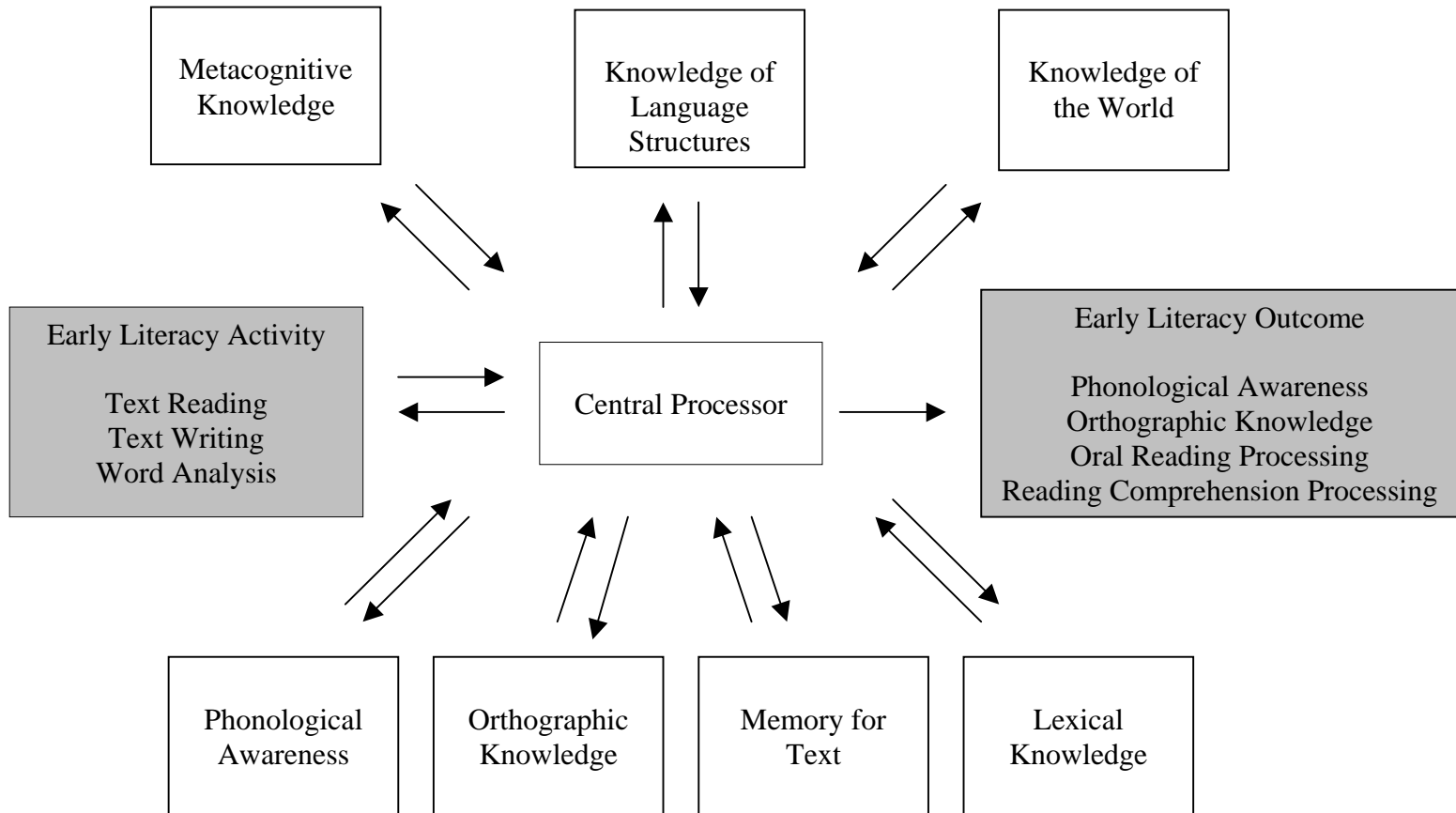


Figure 1. Interactive Model of Early Literacy Development

reading that describes how a child processes words during text reading. In her model, Ehri depicts a child's central processor surrounded by various knowledge sources and the text reading activity that inputs information into the central processor. The central processor, the text, and the knowledge sources interact simultaneously and operate in parallel to facilitate the construction of meaning. Ehri based her model on Rumelhart's (1977) model of interactive reading.

The interactive model of early literacy development that underscores this study differs from Ehri's (1998) model in four ways. First, Ehri's model considers only a text reading activity because her model reflects reading development, whereas the model guiding this study considers a text reading, text writing, or word analysis activity because the model guiding this study reflects early literacy development. Second, Ehri's model is structured such that information from the knowledge sources and the text reading activity flows into the central processor, whereas the model guiding this study is structured such that information from the knowledge sources and the early literacy activity flows into and out from the central processor. Third, the goal of Ehri's model is the construction of meaning, whereas the goal of the model guiding this study is the construction or meaning, the production of meaning, or the completion of an isolated word analysis activity. Fourth, Ehri's model does not address early literacy outcomes, whereas the model guiding this study does.

### *Early Literacy Development*

Recommended Reading Recovery children may encounter difficulties in developing specific aspects of their early literacy development, including phonological awareness, orthographic knowledge, oral reading processing, and/or reading

comprehension processing. This section discusses these four aspects of early literacy development by reviewing the research on them and explaining their role in the interactive model of early literacy development that guides this study (see Figure 1).

### *Phonological Awareness*

Phonological awareness is the ability to recognize and manipulate the various sound segments of syllables, rhymes, onsets, rimes, and phonemes in spoken words. It plays an important role in children's early literacy development as they learn to read and write in an alphabetic script, such as English (Blachman, 2000). Phonological awareness consists of the following components: syllable awareness, rhyme awareness, onset and rime awareness, phonemic awareness, and graphophonemic awareness (Goswami, 2001). Phonological awareness is one of the knowledge sources and early literacy outcomes depicted in the interactive model of early literacy development that guides this study (see Figure 1).

Numerous phonological awareness tasks that vary in cognitive demands require children to recognize and manipulate the various sound segments in spoken words. Researchers have examined individual children's performance on these tasks and found that their phonological awareness conforms to a developmental progression. That is, the ability to analyze spoken words into syllables precedes the ability to analyze spoken words into phonemes (e.g., Liberman, Shankweiler, Fischer, & Carter, 1974), with the ability to analyze onsets and rimes as an intermediary step (e.g., Treiman & Zukowski, 1996).

According to Scarborough (1998), kindergarten children's phonological awareness appears to be "more of a successful predictor of future *superior* reading than

of future reading *problems*.” (p. 85). Classificatory analyses have indicated that few children who begin kindergarten with strong phonological awareness are likely to encounter difficulty in learning to read, but many children who enter kindergarten with weak phonological awareness are unlikely to encounter difficulty in learning to read and become average readers (e.g., Bradley & Bryant, 1985). That is, kindergarten children’s performance on phonological awareness tasks does not appear to predict subsequent reading achievement very well because upon school entry many children who become average readers lack phonological awareness, and are thereby indiscernible from children who become poor readers (Scarborough, 1998).

Reciprocal causation characterizes the relationship between phonological awareness and early literacy development in that each facilitates and is facilitated by the other (Blachman, 2000). Several researchers have found that phonological awareness plays a reciprocal and causal role in children’s reading (e.g., Perfetti, Beck, Bell, & Hughes, 1987; Stuart & Coltheart, 1988) and spelling (e.g., Vandervelden & Siegel, 1995) of words. Researchers have proposed (a) that children’s awareness of syllables, onsets, rimes, and rhymes and their ability to complete simple phonological awareness tasks precedes their early development of reading and spelling and (b) that children’s awareness of phonemes and their ability to complete advanced phonological awareness tasks follows their later development (e.g., Stahl & Murray, 1998).

*Research on phonological awareness: Correlational studies.* Correlational studies have described the relationship between phonological awareness and early literacy development (Blachman, 2000). Some correlational studies have demonstrated that children’s phonological awareness is related to their achievement in early reading and

spelling (e.g., Fox & Routh, 1975). To demonstrate the relationship between phonological awareness and early reading, Fox and Routh conducted a study on children, aged three to seven years old ( $N = 50$ ). The researchers administered (a) a phonological awareness task that assessed the children's ability to repeat spoken sentences and then to segment the sentences into words, the words into syllables, and the syllables into phonemes, (b) a standardized measure of word reading, and (c) a standardized measure of reading comprehension.

Fox and Routh (1975) found moderate and statistically significant correlations corresponding to the children's ability to (a) segment syllables into phonemes and their performance on the word reading measure ( $r = .50, p < .01$ ) and (b) segment words into syllables and their performance on this same measure of word reading ( $r = .47, p < .01$ ). These concurrent correlations suggest that the children's phoneme segmentation ability has a slightly stronger relationship to word reading than their syllable segmentation ability. The researchers also found moderate and statistically significant correlations between the children's ability to (a) segment syllables into phonemes and their performance on the reading comprehension measure ( $r = .37, p < .01$ ) and (b) segment words into syllables and their performance on this same reading comprehension measure ( $r = .39, p < .01$ ). These concurrent correlations indicate that the children's phoneme and syllable segmentation abilities have a similar relationship to reading comprehension. Additional studies on the concurrent relationship between phonological awareness and early reading and spelling achievement support Fox and Routh's findings (e.g., Calfee, Lindamood, & Lindamood, 1973).

Many correlational studies have described the predictive relationship between

phonological awareness and later reading and spelling achievement (e.g., Share, Jorm, Maclean, & Matthews, 1984). In a longitudinal study that determined the predictive power of phonemic awareness, Share et al. examined Australian children's attributes relating to school, home, and family. At the beginning of kindergarten, Share et al. individually administered a set of measures that assessed the children's motor skills, oral language, behavior, and early literacy development, including phonemic awareness (N = 543). At this time, they also (a) distributed questionnaires to the children's parents or guardians to obtain information about the children's home literacy environment, family practices (e.g., hours of television viewing), and family information (e.g., family size), and (b) reviewed school documents to obtain additional family information (e.g., father's occupation). Then, Share et al. administered (a) measures of early reading achievement (e.g., letter naming) to these same children (N = 525) at the end of kindergarten and (b) measures of early reading and spelling achievement at the end of first-grade (N = 479).

Share et al. (1984) found that the children's performance on a phoneme segmentation task at the beginning of kindergarten was the second-strongest predictor of reading achievement (i.e., composite reading score) at the end of kindergarten ( $r = .66, p < .001$ ) and the strongest predictor of reading and spelling achievement (i.e., composite reading and spelling score) at the end of first-grade ( $r = .62, p < .001$ ). These predictive correlations indicate that children's phonemic segmentation ability strongly predicts reading and spelling achievement during the first two years of early literacy instruction.

Additional studies on the predictive power of phonological awareness support Share et al.'s (1984) findings. Some studies have shown that kindergarten children's phonemic awareness predicts their first-grade reading achievement (e.g., Lundberg,

Olofsson, & Wall, 1980; Mann, 1993), as well as spelling achievement (e.g., Lundberg et al., 1980). Other studies have shown that kindergarten children's graphophonemic awareness predicts their first-grade reading achievement (e.g., Mann, 1993; Mann, Tobin, & Wilson, 1987). Furthermore, additional studies have shown that preschool children's phonological awareness predicts their later reading and spelling achievement (e.g., Bradley & Bryant, 1983; Bryant, MacLean, Bradley, & Crossland, 1990). Interestingly, researchers have conducted some of these studies in languages other than English (e.g., Lundberg et al., 1980).

Goswami and Bryant (1990) theorized that phonological awareness consists of the distinct components or factors of rhyme awareness and phonemic awareness and that both factors make separate contributions to the development of early reading and spelling. To test Goswami and Bryant's theory, researchers conducted studies that identified the factors underlying phonological awareness and determined whether rhyme awareness predicts reading and spelling achievement better than phonemic awareness or vice versa (e.g., Muter, Hulme, Snowling, & Taylor, 1998).

In their longitudinal study, Muter et al. (1998) administered several phonological awareness tasks (e.g., rhyme production task and phoneme deletion task) to the children (N = 38) on three testing occasions, each one year apart, as the children advanced in age from four to six years old. The children, aged four years old, had not received formal literacy instruction by the first testing occasion. The children, aged five and six years old, had received formal literacy instruction by the second and third testing occasions, respectively. Additionally, Muter et al. administered several measures of reading and spelling achievement to the children on the second and third testing occasions.

To identify the factors underlying phonological awareness, Muter et al. (1998) conducted principal component analyses on the children's scores on the phonological tasks at Times 1, 2, and 3 and found two separate factors: a phoneme factor and a rhyme factor. The children's scores on the first testing occasion yielded a rhyme and a phoneme factor, accounting for 48% and 29% of the observed variance, respectively. Their scores on the second testing occasion also produced a rhyme and a phoneme factor, accounting for 20% and 49% of the observed variance, respectively. The children's scores on the third testing occasion yielded a rhyme and a phoneme factor, accounting for 57% and 19% of the observed variance, respectively. Muter et al. argued that the fluctuations in the percentages of observed variances reflect the changes in the children's development of phonological awareness over the three years. These findings suggest that a rhyme and a phoneme factor underlie phonological awareness, thereby supporting Goswami and Bryant's (1990) theory. Other researchers have also concluded that phonological awareness consists of separate factors corresponding to the different levels of language analysis required by phonological awareness tasks (e.g., Høien, Lundberg, Stanovich, & Bjaalid, 1995).

Furthermore, Muter et al. (1998) determined the predictive potency of rhyme awareness relative to phonemic awareness by correlating the children's scores on the phonological awareness tasks with their scores on the reading and spelling measures as the children advanced in age. First, the researchers found (a) low correlations between the children's scores at four years old on the rhyme awareness tasks and their scores at five years old on the measures of reading ( $r = .27$ ) and spelling achievement ( $r = .16$ ) and (b) moderate correlations between the children's scores at four years old on the phonemic



awareness tasks and their scores aged five years old on the same measures of reading ( $r = .36$ ) and spelling achievement ( $r = .54, p < .001$ ). Second, Muter et al. found (a) low correlations between the children's scores at four years old on the rhyme awareness tasks and their scores at six years old on the measures of reading ( $r = .07$ ) and spelling achievement ( $r = .04$ ) and (b) low to moderate correlations between the children's scores at four years old on the phonemic awareness tasks and their scores at six years old on measures of reading ( $r = .29$ ) and spelling achievement ( $r = .40, p < .01$ ). Third, the researchers found (a) low correlations between the children's scores at five years old on the rhyme awareness tasks and their scores at six years old on measures of reading ( $r = -.01$ ) and spelling achievement ( $r = .22$ ) and (b) moderate correlations between the children's scores at five years old on the phonemic awareness tasks and their scores at six years old on the measures of reading ( $r = .62, p < .001$ ) and spelling achievement ( $r = .60, p < .001$ ). These findings indicate that phonemic awareness predicts later reading and spelling achievement better than rhyme awareness.

Muter et al.'s (1998) findings support Goswami and Bryant's (1990) theoretical position that rhyme and phonemic awareness maintain a predictive relationship to early reading and spelling development and that phonemic awareness maintains a stronger predictive relationship with reading and spelling achievement than rhyme awareness. Additional studies on the predictive power of phonological awareness support Muter et al.'s findings (Muter & Snowling, 1998; Nation & Hulme, 1997).

*Research on phonological awareness: Experimental studies.* Unlike correlational studies, experimental studies have evaluated the effectiveness of phonological awareness instruction that causes children's phonological awareness to improve and their reading

and/or spelling achievement to advance (Blachman, 2000). Yet Troia's (1999) critical review of these intervention studies revealed that they contain several methodological limitations (e.g., failure to employ random assignment), thereby placing the research findings, as well as their interpretations and generalizability, in question. Although some experimental studies have evaluated the effects of instruction in phonological awareness on the improvement of phonological awareness and the advancement of reading and spelling achievement, many more have evaluated the effects of instruction in two specific components of phonological awareness, phonemic and graphophonemic awareness (National Reading Panel, 2000).

The National Reading Panel (2000) recently conducted a quantitative meta-analysis that evaluated the effects of phonemic and graphophonemic awareness instruction on the improvement of phonemic and graphophonemic awareness and the advancement of reading and spelling achievement. The panel examined 52 intervention studies that produced 96 comparisons of children in treatment groups who received phonemic and/or graphophonemic awareness instruction and children in control groups who received alternative forms of instruction or no instruction. Although the majority of the children spoke English, some of the children spoke other alphabetic scripts, such as Norwegian or Dutch. The grade levels of the children in these studies ranged from preschool to sixth-grade. The delivery of instruction differed across studies with children receiving tutoring, small group instruction, or whole class instruction. The children in these studies formed three distinct groups: children who were progressing normally in reading, children who were at risk for developing reading difficulties, and children who were identified as possessing reading difficulties. The instructional delivery varied across

studies with some children receiving instruction in the manipulation of phonemes in spoken words and other children receiving instruction in the manipulation of phonemes in spoken words with letters as markers for phonemes. Furthermore, the length of instruction varied, ranging from 2.5 to 67 hours.

The National Reading Panel's (2000) analysis of the effect sizes of the individual comparisons between the treatment and control groups revealed that phonemic and graphophonemic awareness instruction improved children's phonemic and graphophonemic awareness, as evidenced by a large, statistically significant effect size ( $d = .86$ ). Additional analyses of effect sizes revealed that this instruction produced moderate, statistically significant effects on reading ( $d = .53$ ) and spelling ( $d = .59$ ) achievement. Furthermore, the panel's analysis revealed that the effects on phonemic and graphophonemic awareness instruction were greater for children who were progressing normally in reading and for children who were at risk for developing reading difficulties than for older children who were identified with reading difficulties. In addition, this instruction yielded greater effects for children in preschool and kindergarten than for children in first- through sixth-grade.

In addition to finding substantial effects corresponding to phonemic and graphophonemic awareness instruction, the National Reading Panel (2000) determined that when phonemic and graphophonemic awareness instruction adhered to certain conditions, the instruction produced larger effects than instruction that fails to adhere to such conditions. The panel found that phonemic and graphophonemic awareness instruction (a) that focused one or two skills (e.g., phoneme segmentation) yielded larger effects than instruction that focused on more than two skills, (b) that ranged from 5-18

hours produced greater effects than shorter or longer treatments, (c) that incorporated the use of letters as markers for sounds proved to be more effective than instruction that did not use letters, and (d) that used small group instruction produced greater effects than tutoring or classroom instruction.

One of the intervention studies included in the National Reading Panel's (2000) meta-analysis was a study conducted by Iversen and Tunmer (1993). The primary purpose of this study was to determine whether Reading Recovery tutoring would be more effective if Reading Recovery teachers provided their children with instruction in phonemic awareness. The researchers formed three matched groups of first-grade children (N = 96) identified as at-risk for reading difficulties: (a) a standard Reading Recovery group who received standard Reading Recovery tutoring (n = 32), (b) a modified Reading Recovery group who received modified Reading Recovery tutoring (n = 32), and (c) a standard intervention group who received supplemental small group instruction in early literacy (n = 32). The three groups of children were matched on their letter knowledge, graphophonemic awareness, geographical location, socioeconomic status, and type of classroom reading program. The researchers also included a classroom control group that consisted of first-grade children from the same classrooms as the children in the two Reading Recovery groups who were judged by their classroom teachers to be average readers.

Reading Recovery teachers taught the children in the standard and modified Reading Recovery groups, whereas reading specialists with master's degrees in reading taught the children in the standard intervention group. The classroom teachers taught the children in the control group. The modified Reading Recovery group received explicit

instruction in phoneme-grapheme correspondences by manipulating plastic letters as markers for sounds. The other groups did not receive this instruction. The teachers individually administered the tasks of *An Observation Survey of Early Literacy Achievement* (Clay, 2002), a word reading measure, a pseudoword phonological recoding measure, a phoneme deletion task, and a phoneme segmentation task. The teachers administered these measures (a) at the beginning of the school year prior to treatment, (b) in middle of the school year at the end of treatment, and (c) at the end of the school year. Iversen and Tunmer (1993) only included the matched triplets that contained Reading Recovery children who met the criteria for successful performance upon their exit from tutoring.

Upon the completion of treatment, the Iversen and Tunmer (1993) determined that the modified Reading Recovery group performed better than the classroom control on the phoneme segmentation task at a statistically significant level. At the end of the school year, the researchers found that the modified Reading Recovery group performed better than the standard Reading Recovery group on Clay's (2002) Text Reading task, measure of oral reading accuracy at a statistically significant level. These findings indicate that instruction in phonemic awareness improved the modified Reading Recovery group's phonemic awareness and advanced their reading achievement more than the classroom control group and standard Reading Recovery group, respectively.

Additional intervention studies support Iversen and Tunmer's (1993) finding that instruction in phonemic awareness improves phonemic awareness and reading achievement. In some of these studies, children of varying ages received (a) individual instruction in long-term tutoring programs (e.g., Hatcher, Hulme, & Ellis, 1994), (b)

individual instruction in short-term training programs (e.g., Bradley & Bryant, 1983, 1985), (c) small group instruction (e.g., Byrne & Fielding-Barnsley, 1991, 1993, 1995), and (d) whole class instruction (e.g., Lundberg, Frost, & Petersen, 1988). Interestingly, the children in these studies spoke different alphabetic scripts, such as English and Dutch).

*Phonological awareness and the interactive model of early literacy development.*

As illustrated in Figure 1, phonological awareness is one of many knowledge sources and early literacy outcomes featured in the interactive model of early literacy development that guides the current study. This model can be used to understand the development of a child's phonological awareness. When a child writes an unfamiliar word, his/her central processor may process information from the text writing activity. A child's central processor may also employ strategies, such as monitoring the writing process. Furthermore, a child's central processor may access information from the many knowledge sources to write an unfamiliar word. For instance, a child's central processor may access phonological awareness to recognize individual phonemes and match graphemes to these phonemes to write an unfamiliar word. A child's central processor may draw upon lexical knowledge to identify and retrieve phonemes and corresponding graphemes to write an unfamiliar word. A child's central processor may also access metacognitive knowledge to monitor (a) the identification and retrieval of phonemes and graphemes stored in the lexicon and (b) the writing of the graphemes in an unfamiliar word. A child's phonological awareness is observed as an early literacy outcome.

Clay (1993) designed the Reading Recovery lesson components and corresponding teaching procedures to promote growth in phonological awareness. A

Reading Recovery child receives explicit instruction in phonological awareness from the onset of his/her tutoring. For example, when a child wants to write an unfamiliar word in a self-composed sentence, s/he learns to recognize phonemes in a spoken word and to represent these phonemes by writing their corresponding graphemes. Before a child writes these graphemes, the Reading Recovery teacher provides a phonological framework by drawing sound boxes, a box for each phoneme in the unfamiliar word (Clay, 1993).

*Summary.* Phonological awareness, specifically phonemic and graphophonemic awareness, plays a critical role in children's early literacy development of an alphabetic script, such as English. Researchers have examined (a) the relationship between phonological awareness and reading and spelling achievement, (b) the predictive potency of phonological awareness, (c) the factors underlying phonological awareness, and (d) the effectiveness of phonemic and graphophonemic awareness instruction in improving phonemic and graphophonemic awareness and advancing reading and spelling achievement. Phonological awareness is one of many knowledge sources and early literacy outcomes featured in the interactive model of early literacy development that guides the current study (see Figure 1). The current study compares recommended to discontinued Reading Recovery children prior to and following their tutoring on (a) overall phonological awareness in the form of pre- and post-tutoring composites and (b) the following components that formulated these composites: rhyme awareness; phonological skeletal structure awareness; syllable, onset and rime, and phonemic awareness; and graphophonemic awareness.

### *Orthographic Knowledge*

Three overlapping layers of knowledge characterize children's English spelling development: an alphabetic layer, a pattern layer, and a meaning layer (Henderson & Templeton, 1986; Templeton & Morris, 2000). The alphabetic layer involves children using their phonological knowledge of phoneme-grapheme correspondences to spell words. The pattern layer involves children using their orthographic knowledge of letter sequences and/or patterns to spell words. In the current study, I referred to this pattern layer as orthographic knowledge. The meaning layer involves children using their morphological knowledge of derivational relations to spell words and understanding how these relations determine meaning and influence pronunciation. Orthographic knowledge is one of the knowledge sources and early literacy outcomes depicted in the interactive model of early literacy development that guides this study (see Figure 1).

Although researchers agree that these three layers of knowledge develop over time (e.g., Henderson & Templeton, 1986; Treiman, 1993), they disagree about how to characterize their development. Some researchers analyzed children's misspellings and developed stage models. Stage models propose that children progress through qualitatively different stages over the course of their spelling development (Ehri, 1986; Gentry, 1982; Henderson, 1985). That is, during the early stages of spelling development, children develop and use their knowledge of phoneme-grapheme correspondences (i.e., alphabetic layer) to spell words, whereas during the later stages, the children develop and use their orthographic knowledge (i.e., pattern layer) and morphological knowledge (i.e., meaning layer). Stage models suggest that the pattern and meaning layers of knowledge are inaccessible to children in their early stages of spelling development. Other



researchers observed children's spelling behaviors and analyzed their misspellings and verbal self-reports and developed strategy models. Strategy models propose that children develop and use a variety of strategies, such as lexical retrieval of a word's spelling, and different types of knowledge, such as knowledge of a conventional spelling rule, over the course of their spelling development (e.g., Rittle-Johnston & Siegler, 1999; Treiman & Bourassa, 2000b).

Treiman and Bourassa (2000b) argued that although the stage models support the developmental nature of learning to spell, they fail to capture fully the phonological, orthographic, and morphological complexities of spelling development. According to stage models, children in their beginning stages of spelling development primarily use their phonological knowledge of phoneme-grapheme correspondences. If so, these children's misspellings would not contain orthographic letter sequences and/or patterns. However, Treiman and her colleagues have conducted several research studies that indicate simple orthographic knowledge emerges early in children's spelling development; thereby questioning stage models of spelling development (e.g., Cassar & Treiman, 1997; Treiman, 1993).

*Research on orthographic knowledge.* In a descriptive and naturalistic study that consisted of several investigations, Treiman (1993) examined children's orthographic knowledge of English letter sequences and/or patterns and their phonological knowledge of phoneme-grapheme correspondences in separate investigations. Results from the investigations that examined orthographic knowledge revealed that kindergarten, first-, and second-grade children's knowledge of letter sequences and/or patterns emerges early in their spelling development.

In one of these investigations, Treiman (1993) collected first-grade children's writings (N = 43) from the same classroom over the course of two years. She collected writings from one cohort (n = 26) during the first year and writings from a second cohort (n = 17) the following year. The children wrote stories and then dictated to their teacher what they had written. The teacher recorded the children's dictated words on the children's papers. However, the teacher did not (a) discuss with the children how their misspellings differed from correct spellings, (b) provide direct instruction in spelling, (c) stress the importance of correct spelling, or (d) spell words for the children. Although Treiman gathered children's writing over the course of each school year, she analyzed their misspellings from the beginning and end of each school year.

Upon completion of her analysis, Treiman (1993) concluded that first-grade children understood that some vowels and consonants are spelled with more than one grapheme. The children's misspellings included the following orthographic patterns: vowel doublets (e.g., *oo* in *good*), vowel digraphs (e.g., *ai* in *said*), vowel plus final e (e.g., *i* and *e* in *bike*), consonant doublets (e.g., *tt* in *kitten*), and consonant digraphs (e.g., *th* in *think*).

Furthermore, Treiman (1993) concluded that the children understood graphemic alternations; that is, children understood that certain graphemes tend to appear in some spellings, while other graphemes tend to appear in other spellings. Treiman found that the children understood that the position of the graphemes in the words govern graphemic alternations. The children demonstrated this understanding in four separate ways. First, the children understood that vowel digraphs ending in *w* (e.g., *ow*, *aw*, *ew*) and *y* (e.g., *ay*, *ey*, *oy*) commonly appear before vowels (e.g., *power*) and at the ends of morphemes (e.g.,

*obey*), whereas vowel digraphs ending in *i* (e.g., *ai, ei, oi*) and *u* (e.g., *au, eu, ou*) commonly appear before consonants within morphemes (e.g., *maid*). Second, they knew that the single grapheme *y* generally occurs at the ends of morphemes (e.g., *myself*), whereas the single grapheme *i* generally occurs at the beginnings and in the middles of words (e.g., *dinosaur*). Third, the children understood that the consonant digraph *ck* occurs in the middles of morphemes (e.g., *package*) and at the ends of morphemes (e.g., *sick*), whereas the single grapheme *c* or *k* occurs at the beginnings of words. Fourth, they understood that double consonants tend to occur in the middles of words (e.g., *pepper*) and at the ends of words (e.g., *egg*), whereas single consonants tend to occur at the beginnings of words. Treiman also found that children understood that the phonological structures of words govern graphemic alternations. The children understood that double consonants tend to follow stressed short vowels (e.g., *pepper*) and precede unstressed vowels (e.g., *mommy*), whereas single consonants tend to follow long vowels (e.g., *baby*). These findings indicate that first-grade children incorporated orthographic knowledge of specific letter sequences and/or patterns in their misspellings at the beginning and end of the school year without explicit instruction.

In two other investigations, Treiman (1993) examined children's ability to recognize orthographic letter patterns in print. In both investigations, the children completed an orthographic constraints test that consisted of pronounceable nonword pairs (e.g., *ckun, nuck*). Each nonword pair tested one constraint or regularity of English (e.g., *ck* occur in the middle or at the end of words). One of the nonwords in each pair conformed to an orthographic pattern and the other nonword did not. The children selected the nonwords that looked most like real words.

In her first investigation, Treiman (1993) assessed the orthographic knowledge of kindergarten, first-, and second-grade children who were progressing normally in reading toward the end of the school year. These children were Caucasian and from middle to upper-middle income homes. In her second investigation, Treiman assessed the orthographic knowledge of first-grade children at the beginning and end of the school year. These children were ethnically diverse and from lower to lower-middle income homes. None of the children in either study received explicit instruction in the orthographic patterns.

The kindergarten, first-, and second-grade children in Treiman's (1993) first investigation produced a mean percentage of correct responses of 56.4%, 62.3%, and 83.2%, respectively. The performance of these children increased from first- to second-grade and the difference reached statistical significance. Furthermore, the children's performance in each grade level differed from what would be expected by chance and the differences reached statistical significance. In her second investigation, the first-grade children produced a mean percentage of correct responses of 49.8% and 67.6% at the beginning and end of the school year, respectively. The increase in the children's performance from the beginning to the end of the school year was statistically significant. These findings from both investigations indicate that children from different socioeconomic backgrounds possess orthographic knowledge at the end of the school year. Furthermore, the findings from the second investigation indicate that first-grade children increase their orthographic knowledge from the beginning to the end of the school year without explicit instruction.

Overall, Treiman's (1993) findings suggest that children begin to develop

orthographic knowledge during kindergarten, first-, and second-grade; thereby challenging stage models of spelling development. Her findings further suggest that learning about orthographic knowledge is a developmental process; children advance their understanding of orthographic knowledge as they progress in school. Treiman argued that because orthographic knowledge was not explicitly taught in school and was probably not taught at home, the children developed it through early literacy experiences that involved exposure to orthographic letter sequences and/or patterns in printed words. Additional research conducted by Treiman and her colleagues support Treiman's original findings. Cassar and Treiman (1997) and Treiman, Berch, and Weatherston (1993) found that kindergarten, first-, and second-grade possess knowledge of the letter sequences and/or patterns, specifically double vowels and consonants. Treiman and Bourassa (2000a) found that kindergarten, first-, and second-grade children's misspellings contain acceptable orthographic letter sequences and/or patterns.

*Orthographic knowledge and the interactive model of early literacy development.*

As seen in Figure 1, orthographic knowledge is one of many knowledge sources and early literacy outcomes featured in the interactive model of early literacy development that guides this study. This model can be used to understand the development of a child's orthographic knowledge. When a child writes an unfamiliar word, his/her central processor may process information from the text writing activity. A child's central processor may also employ strategies, such as monitoring the writing process. Furthermore, a child's central processor may access information from the many knowledge sources to write an unfamiliar word. For example, a child's central processor may access orthographic knowledge to identify the orthographic letter sequences and/or

patterns or words containing these letter sequences and/or patterns to write an unfamiliar word. A child's central processor may draw upon lexical knowledge to identify and retrieve orthographic letter sequences and/or patterns or words containing these letter sequences and/or patterns to write an unfamiliar word. A child's central processor may also access metacognitive knowledge to monitor (a) the identification and retrieval of the letter sequences and/or patterns stored in the lexicon and (b) the writing of the graphemes in the unfamiliar word. A child's orthographic knowledge is observed as an early literacy outcome.

Clay (1993) designed the Reading Recovery lesson components and corresponding teaching procedures to foster growth in orthographic knowledge. A Reading Recovery child receives explicit instruction in orthographic knowledge. For example, when a child wants to write an unfamiliar word in a self-composed sentence, s/he learns to think of a familiar word and the corresponding letter sequence and/or pattern that sounds and looks the same as the unfamiliar word. Before a child writes the letter sequence and/or pattern, the Reading Recovery teacher provides a orthographic framework by drawing letter boxes, a box for each grapheme in the unfamiliar word (Clay, 1993).

*Summary.* Orthographic knowledge is central to children's early literacy development. Treiman and her colleagues have conducted a program of research demonstrating that (a) orthographic knowledge emerges early in children's spelling development; earlier than proposed by stage models of spelling development, and (b) children's learning of orthographic knowledge is a developmental process. Orthographic knowledge is one of many knowledge sources and early literacy outcomes featured in the

interactive model of early literacy development that guides this study (see Figure 1). The current study compares recommended to discontinued Reading Recovery children prior to and following their tutoring on (a) overall orthographic knowledge in the form of pre- and post-tutoring composites and (b) the following components that formulated these composites: spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge.

### *Oral Reading Processing*

Oral reading processing refers to children's ability to recognize familiar words and/or problem-solve unfamiliar words in text. When children instantaneously recognize familiar words in text, they read them automatically and accurately without conscious attention. When children problem-solve unfamiliar words in text, they read them with deliberate and conscious attention. As children develop their oral reading processing, they problem-solve fewer words because they continuously increase the number of familiar words stored in their lexicons. When children develop automatic and accurate oral reading processing, they direct their attention to reading comprehension processing (LaBerge & Samuels, 1974).

The development of children's oral reading processing enables them to read familiar and unfamiliar words. They learn to read familiar words by sight and unfamiliar words by (a) decoding them, (b) analogizing them to familiar words, or (c) by using context to predict them (Ehri, 1998). Not surprisingly, children read familiar words faster than they read unfamiliar words (e.g., Ehri & Wilce, 1983).

Some developmental theories of oral reading processing suggest that children develop their ability to read unfamiliar words in a series of sequential stages (e.g., Chall,

1996; Ehri, 1998; Frith, 1985). In the initial stage, children apply salient, visual symbols that elicit the words' pronunciations and/or meanings. During this initial stage, they may read unfamiliar words by using context. In the intermediate stage, children add the application of phoneme-grapheme correspondences. They begin this intermediate stage by linking some phonemes to their graphemes (e.g., initial and/or final graphemes) and end it by linking all phonemes to their graphemes. During this intermediate stage, they may read unfamiliar words by decoding and/or using context. In the final stage, children add the application of orthographic letter sequences and/or patterns. During this final stage, they may read unfamiliar words by analogizing to familiar words, as well as by decoding and/or using context. According to stage models, children engage in qualitatively different oral reading processing abilities at different stages with the earlier stages available to children as they advance into the later stages.

Other developmental theories of oral reading processing propose that children develop their abilities to read unfamiliar words concurrently rather than in a series of stages. That is, children's phonological knowledge (e.g., rhyme awareness) interacts continuously with their knowledge of orthographic letter sequences and/or patterns (e.g., rimes) (e.g., Goswami, 1998). Children may read unfamiliar words by decoding, analogizing and/or using context. According to interactive models, the interplay between phonological and orthographic knowledge assists children in their development of oral reading processing.

*Research on reading words by sight.* When children read familiar words by sight, they access word-specific information stored in their lexicons from previous literacy experiences. They locate information about the words' meanings, pronunciations, and



spellings of words and retrieve this information without conscious attention and deliberate processing. During this process, children bypass all the other words in their lexicons, including the words with similar meanings, pronunciations, and spellings (Ehri, 1998). Interestingly, first-grade children retain words in their lexicons after reading them as few as four times (Reitsma, 1983). When children read unfamiliar words by sight, they (a) read them without pauses between phonemes, letter sequences and/or patterns, and syllables, (b) read them faster than equally-decodable nonwords, (c) pronounce irregularly spelled words correctly rather than phonetically decoding them, and/or (d) distinguish identically pronounced words with correct spellings from nonwords with homophonic spellings (Ehri, 1998, 2004). As children advance in their early literacy development, they continuously add words to their lexicons; thereby increasing the number of words they read by sight. Eventually, they read the majority of words, if not all words, by sight (Ehri, 1998).

Ehri and Wilce (1983) examined children's ability to read familiar words by sight. These researchers theorized that if children read familiar words by sight rather than decoding, then they should read familiar, decodable words faster than unfamiliar, equally-decodable nonwords. Ehri and Wilce recorded the reaction times of skilled ( $n = 8$ ) and unskilled ( $n = 8$ ) first-grade readers and skilled ( $n = 8$ ) and unskilled ( $n = 8$ ) second-grade readers. The researchers recorded the number of seconds that these children required to read familiar, decodable words (e.g., man, five) and unfamiliar, equally-decodable nonwords (e.g., guz, baf). The skilled and unskilled readers in both grade levels required more time to read the unfamiliar, equally-decodable nonwords than the familiar, decodable words, as evidenced by their mean scores (i.e., number of seconds).

Although the researchers displayed the children's mean scores in a graph, they failed to provide the exact means and corresponding standard deviations. This finding indicates that skilled and unskilled readers in first- and second-grade read familiar words by sight.

Additional research studies support Ehri and Wilce's (1983) findings that children read unfamiliar words by sight. Some studies have examined children's ability to read familiar words automatically (e.g., Guttentag & Haith, 1978). Other studies have investigated children's ability to pronounce irregularly spelled words correctly rather than phonetically decoding them (e.g., Adams & Huggins, 1985). Additional studies have examined children's ability to distinguish identically pronounced words with correct spellings from nonwords with homophonic spellings (e.g., Reitsma, 1983).

*Research on reading words by decoding.* Children read unfamiliar words by decoding, also known as phonological recoding. They decode unfamiliar words by (a) sounding graphemes and blending their corresponding phonemes to generate approximate pronunciations and (b) searching their lexicons for meaningful words that match these pronunciations. Because English is not completely phonetic, children often try several pronunciations before producing a recognizable word that matches the phoneme-grapheme correspondences and meaning. As children become more skilled in their reading, they typically decode words by (a) locating the orthographic letter sequences and/or patterns (e.g., rimes) to generate pronunciations and (b) searching their lexicons for meaningful words that fit these pronunciations. That is, they pronounce the letter sequences and/or patterns as units without parsing them into individual graphemes (Ehri, 1998). Children's ability to decode words becomes faster and less overt as they become more skilled in their reading of texts.

Over the course of the school year, Cohen (1975) analyzed first-grade children's (N = 50) oral word reading errors, specifically non-response errors, word substitutions, and nonword substitutions, that they made during the reading of instructional and non-instructional texts. Then, Cohen made inferences about the children's development of decoding skills. She also categorized the children's word and nonword substitutions according to their degree of graphic approximation to the unfamiliar words in the texts, for example a substitution that shared first and/or last letters or at least half of the letters with the word in the text. Cohen selected children from two classrooms who were nonreaders at the beginning of the school year.

At the onset of Cohen's (1975) investigation, the children produced more non-response errors than word and nonword substitutions when they encountered unfamiliar words in texts. The children's percents of total non-response errors on both instructional and non-instructional texts were 61% and 67%, respectively, whereas their percents of total word substitutions were 10% and 8%, respectively; and their percents of total nonword substitutions were 7% and 8%, respectively. These findings indicate that children who begin first-grade as nonreaders possessed weak decoding skills, as evidenced by their low percentages of word and nonword substitutions.

By the middle of the investigation, the children produced approximately the same number of non-response errors as word and nonword substitutions. More precisely, the children's percents of total non-response errors on instructional and non-instructional texts were 31% and 32%, respectively, whereas their percents of total word substitutions were 33% and 29%, respectively; and their percents of total nonword substitutions were 25% and 31%, respectively. Half of the children's word and nonword substitutions

resembled the unfamiliar words in the texts, sharing at least half of the same letters. These findings suggest that by the middle of first-grade children develop substantial decoding skills.

By the end of the investigation, the children slightly decreased the number of non-response errors and stabilized the number of word and nonword substitutions. The children's percents of total non-response errors on instructional and non-instructional materials were 28% and 27%, respectively. Their percents of total word and nonword substitutions ranged between 26% and 35% with many of these word and nonword substitutions sharing at least half of the letters with the unfamiliar words in the texts. These findings indicate that by the end of first-grade children stabilize their ability to read unfamiliar words by decoding.

Additional research studies that have analyzed children's oral reading errors support Cohen's (1975) findings that children's ability to decode words conforms to a developmental progression (e.g., Biemiller, 1970). Other studies that have examined children's performance on nonword reading tasks also suggest that children's decoding skills follow a developmental sequence (e.g., Vandervelden & Siegel, 1995).

*Research on reading words by analogizing.* Analogizing is another way children read unfamiliar words. Children analogize an unfamiliar word to a familiar word by (a) recognizing a specific orthographic letter sequence in the unfamiliar word, (b) matching this letter sequence to a familiar word or words stored in the lexicon, (c) pronouncing the matched letter sequence in the same manner as the letter sequence in the familiar word stored in the lexicon, and (d) blending the matched letter sequence to the other letters in the unfamiliar word. When children blend the matched letter sequence to the other letters

in the unfamiliar word, they understand that they pronounce the matched letter sequences in the unfamiliar word the same as they pronounce the letter sequence in the familiar word stored in their lexicons (Ehri, 1998; Goswami, 1998).

Goswami (1986) conducted one of the first studies that examined children's ability to read unfamiliar words and nonwords by analogy without prior instruction in analogy use. Based on their performance on a word reading test, she divided kindergarten, first-, and second-grade children into three groups: advanced beginning readers (N = 17), beginning readers (N = 18), and nonreaders (N = 18). She assessed the children's ability to read words and nonwords by analogy on two testing occasions.

Goswami (1986) assessed the children's ability to read unfamiliar words when given three types of unfamiliar words: target words, common letter words, and control words. First, she assessed the children's ability to read target words and nonwords *by analogy* in the following three conditions: (a) with the presence of clue words (e.g., *skin*) that shared the same onsets with the target words (e.g., *skip*) and nonwords (e.g., *skib*), (b) with the presence of clue words (e.g., *skin*) that shared the same rimes with the target words (e.g., *chin*) and nonwords (e.g., *hin*), and (c) without the presence of clue words. Second, she also assessed the children's ability to read common letter words and nonwords by other means than analogy in the same three conditions. These common letter words and nonwords contained the same letters as the clue words but were not in the same sequence as the clue words. Thus, the children could not read them by analogy. Third, Goswami assessed the children's ability to read control words and nonwords by other means than analogy in the same three conditions. These control words and nonwords were target and common letter words and nonwords for a different clue word.

Thus, the children could not read them by analogy. To control for initial knowledge of words, Goswami compared the children's scores on the first testing occasion to their scores on the second testing occasion.

From the first to the second testing occasion, the advanced beginning and beginning readers improved more in their ability to read target words and nonwords with the presence of clue words that shared the same onsets and rimes *by analogy* than in their ability to read (a) common letter words and nonwords with the presence of clue words that shared the same onsets and rimes by other means than analogy and (b) control words and nonwords with the presence of clue words that shared the same onsets and rimes by other means than analogy at statistically significant levels. However, these advanced beginning and beginning readers generally improved more in their ability to read (a) common letter words and nonwords and (b) control words and nonwords without the presence of clue words by means other than analogy than in their ability to read target words and nonwords without the presence of clue words *by analogy*. These findings suggest that beginning readers read unfamiliar words and nonwords by analogy when provided clue words that share the same onsets and rimes with the target words and nonwords.

In addition, the advanced beginning and the beginning readers read more target words and nonwords with the presence of clue words that shared the same *rimes* by analogy than target words and nonwords with the presence of clue words that shared the same *onsets* by analogy on the second testing occasion. The advanced beginning and beginning readers produced higher combined mean scores corresponding to the target words and nonwords that shared the same rimes than their combined mean scores

corresponding to the target words and nonwords that shared the same onsets at a statistically significant level. This finding suggests beginning readers read unfamiliar words by analogy when provided clue words that share the same rimes more accurately than they read those that share the same onsets.

The nonreaders failed to read any of the words and nonwords in all three conditions on the first testing occasion, thereby demonstrating an inability to read target words and nonwords by analogy (a) with the presence of clue words that share the same onsets, (b) with the presence of clue words that share the same rimes, and (c) without the presence of clue words. On the second testing occasion, the nonreaders displayed a limited ability to read six target words and nonwords by analogy with the presence of clue words that shared the same rimes, as evidenced by their mean score of .89 ( $SD = 1.37$ ). The children's scores in this condition differed from what would be expected by chance at a statistically significant level. However, these nonreaders demonstrated an inability to read six target words and nonwords by analogy (a) with the presence of clue words that shared the same onsets and (b) without the presence of clue words, as evidenced by their mean scores of .11 ( $SD = .32$ ) and 0.00 ( $SD = 0.00$ ), respectively. These findings suggest that nonreaders possess a limited ability to read unfamiliar words and nonwords by analogy with the presence of clue words that share the same rimes with the target words and nonwords.

Goswami (1986) has conducted additional studies that support her original findings that (a) children read unfamiliar words by analogy when provided clue words that share the same onset and rimes and (b) children read unfamiliar words by analogy that share the same rimes more accurately than they read those that share the same onsets

(e.g., Goswami, 1988). Although Goswami has assessed children's ability to read unfamiliar words by analogy when provided clue words, she has also assessed this ability during the reading of texts and found similar results (e.g., Goswami, 1990). Other researchers have conducted studies that support the importance of rimes in the reading of unfamiliar words by analogy (e.g., Bowey & Hansen, 1994). Although some studies have found that children require minimal competence in decoding to read unfamiliar words by analogy because of the processes of segmenting and blending (e.g., Ehri & Robbins, 1992), other studies challenge these findings (e.g., Goswami, 1986).

*Research on reading words by using context.* Children also read unfamiliar words by predicting them based on context, such as pictures that accompany the text, as well as semantic and/or syntactic information contained in the preceding text (Ehri, 1998). Children use context to predict unfamiliar words during the reading of words in text, as opposed to reading words in isolation. Thus, children read words more accurately in context than in isolation (Adams & Huggins, 1985).

Biemiller (1970) analyzed the oral reading errors that first-grade children (N = 42) made during reading instruction throughout the school year and on a reading assessment administered at the end of the school year, and then made inferences about their ability to read unfamiliar words by using context. The children in this study came from two different classrooms with the majority of the children from one classroom reading above grade level and the majority of the children from the other classroom reading below grade level by the end of the school year.

Biemiller (1970) analyzed the children's development of contextually constrained errors, defined as substitutions, omissions, and insertions characterized by the use of



syntactic and semantic information obtained from the preceding text and pictures that accompanied the text. He also analyzed their development of graphically constrained errors, defined as substitutions characterized by the use of phoneme-grapheme information. Then, he calculated percentages of contextually and graphically constrained errors from among all substitutions, omissions, and insertions to establish trends corresponding to three phases of reading development.

In the first phase, the children mainly produced contextually constrained errors (74%) with a few graphically constrained errors (19%). In the second phase, the children continued to produce contextually constrained errors (76%) and increased the number of graphically constrained errors (39%). In the third phase, the children continued to increase their use of both contextually (83%) and graphically constrained errors (44%). Although the children increased their use of contextually constrained errors throughout the three phases, only the increase between the second and third phase reached statistical significance. These results indicate that first-grade children access a substantial amount of contextual information to read unfamiliar words at the beginning of the school year, and continue to access this information throughout the remainder of the school year.

Because the two categories of contextually and graphically constrained errors were not mutually exclusive, Biemiller (1970) placed some of the children's substitutions into both categories. Additionally, because some of the children's errors were not characterized by the use of either (a) syntactic and semantic information or (b) phoneme-grapheme information, Biemiller excluded some of their substitutions, omissions, and insertions from the two categories. Thus, the percentages of contextually and graphically constrained errors in each phase of reading development did not sum to 100%.

Several analyses of children's oral reading errors over time have found that children use context by accessing semantic and/or syntactic information contained in the preceding text or from pictures that accompany the text to predict unfamiliar words. Whereas Biemiller (1970) analyzed children's combined use of semantic and syntactic information, other studies analyzed (a) children's separate use of semantic and syntactic information (e.g., Clay, 1968; Cohen, 1975; Weber, 1970).

*Oral reading processing and the interactive model of early literacy development.*

Oral reading processing is one of the many early literacy outcomes in the interactive model of early literacy development that guides this study (see Figure 1). This model can be used to understand the development of a child's oral reading processing. When a child reads an unfamiliar word by decoding, analogizing, and/or using context, his/her central processor may process information from the text reading activity. A child's central processor may also employ various strategies, such as monitoring the reading process and searching for and using semantic, syntactic, and graphophonic information. Furthermore, a child's central processor may access information from the many knowledge sources to read an unfamiliar word. For example, a child's central processor may access phonological awareness and orthographic knowledge. A child's central processor may draw upon lexical knowledge to identify and retrieve relevant information about an unfamiliar word's phonemes, graphemes, and letter sequences and/or patterns from the lexicon. To aid in the reading of an unfamiliar word, a child's central processor may activate memory for text. A child's central processor may also access knowledge of language structures, such as semantic and syntactic information, to support the reading of an unfamiliar word. A child's central processor may also draw upon conceptual, factual,

experiential, and schematic knowledge of the world to assist in the reading of an unfamiliar word. The child's central processor may also access metacognitive knowledge to evaluate and regulate the application of the various knowledge sources and the accurate reading of an unfamiliar word. A child's oral reading processing is observed as an early literacy outcome.

Clay (1993) designed the Reading Recovery lesson components and corresponding teaching procedures to develop oral reading processing. From the start of tutoring, a Reading Recovery child receives explicit instruction in oral reading processing. For example, when a child encounters an unfamiliar word in a text, his/her Reading Recovery teacher provides instruction in how to read the unfamiliar word by decoding, analogizing, and/or using context.

*Summary.* The development of oral reading processing makes a vital contribution to children's early literacy development. Research studies have found that children read familiar words by sight and unfamiliar words by decoding, analogizing, and using context. Research studies have examined children's oral reading of words in text (e.g., Biemiller, 1970; Cohen 1975) and in isolation (e.g., Ehri & Wilce, 1985; Goswami, 1986). Oral reading processing is one of several outcomes featured in the interactive model of early literacy development that guides this study (see Figure 1). The current study compares recommended to discontinued Reading Recovery children on oral reading processing in regards to oral reading accuracy, oral reading inaccuracy, and oral reading behaviors following their completion of tutoring.

#### *Reading Comprehension Processing*

Reading comprehension processing is a complex process in which children

construct meaning through their interactions with texts. Reading comprehension processing involves children constructing mental representations by accessing the information or content contained directly in them and integrating this information or content with their background knowledge. The outcome of constructing these mental representations is reading comprehension (Kintsch 1998; van Dijk & Kintsch, 1983). Reading comprehension also involves children's using their conceptual knowledge and knowledge of word meanings (Davis, 1968). Furthermore, reading comprehension involves children's application of reading comprehension strategies (Pearson & Duke, 2002) and reading comprehension monitoring (Baker & Brown, 1984).

When children apply reading comprehension strategies, such as drawing inferences or identifying story structure elements, they engage in processing to develop an understanding of the texts that they read (e.g., Baumann & Bergeron, 1993). When children monitor their reading comprehension, they engage in evaluation and regulation processing to determine whether they understand the texts and to activate reading comprehension strategies to develop, maintain, and/or increase their understanding (Baker & Brown, 1984). Research studies have suggested that young and unskilled readers develop evaluation processing before they develop regulation processing (e.g., Kinnunen & Vauras, 1995).

*Research on reading comprehension strategies.* In a quasi-experimental study of first-grade children (N = 74), Baumann and Bergeron (1993) examined the effects of instruction in the comprehension strategy of identifying story structure elements on enhancing children's reading comprehension processing. At the end of the school year, Baumann and Bergeron assigned four classrooms of children to one of following groups:

story mapping 1 group ( $n = 17$ ), story mapping 2 group ( $n = 20$ ), Directed Reading-Thinking (DRTA) group ( $n = 19$ ), or control group ( $n = 18$ ). The children received six consecutive instructional sessions with each session featuring a specific narrative text. The children in each group read the same texts. They read these texts silently as their teacher read them out loud (i.e., a listening-silent reading procedure). The children in the story mapping 1 group read the stories and received instruction in the comprehension strategy of identifying story structure elements. The children in the story mapping 2 group read the stories and received instruction in the comprehension strategy of identifying story structure elements, as well as writing stories from story maps. The children in the DRTA group read the stories and received instruction in the comprehension strategy of making predictions. The children in the control group read and discussed the stories without receiving instruction in a comprehension strategy. Baumann and Bergeron administered (a) an important idea of a parsed story test (range 0 - 30), (b) a wh-question test (e.g., who and where) (range 0 - 10), (c) an important story element recognition test (range 0 - 7), and (d) a two week delayed wh-question test (range 0 - 10). Additionally, the researchers selected four children from each group ( $n = 16$ ) and assessed their ability to retell a story orally.

The story mapping 1 group produced mean scores of 17.77 ( $SD = 3.19$ ), 8.00 ( $SD = 1.87$ ), 4.94 ( $SD = 1.30$ ), and 8.77 ( $SD = 1.30$ ) on each of the four tests, respectively. Similarly, the story mapping 2 group produced mean scores of 16.65 ( $SD = 2.78$ ), 7.60 ( $SD = 1.47$ ), 5.10 ( $SD = 1.12$ ), and 8.10 ( $SD = 1.65$ ) on these same tests, respectively. The DRTA group produced lower mean scores of 13.42 ( $SD = 3.44$ ), 6.06 ( $SD = 2.16$ ), 3.39 ( $SD = 1.38$ ), and 6.17 ( $SD = 1.89$ ) on these same four tests, respectively. The control

group produced the lowest mean scores of 11.28 ( $SD = 4.00$ ), 6.32 ( $SD = 2.29$ ), 4.58 ( $SD = 1.07$ ), and 7.90 ( $SD = 2.00$ ) on these same tests, respectively. The story mapping 1 and 2 groups outperformed the children in the control group on all four tests at statistically significant levels. These findings suggest that instruction in the comprehension strategy of identifying story structure elements promotes first-grade children's reading comprehension of narrative texts more than no instruction in a comprehension strategy immediately and two weeks following instruction. Additionally, the children in the story mapping 1 and 2 groups outperformed the children in the DRTA group on all four tests with performance on the first two tests reaching statistically significant levels. These findings suggest that instruction in the comprehension strategy of identifying story structure elements promotes first-grade children's reading comprehension of narrative texts more than instruction in the comprehension strategy of making predictions immediately following instruction.

The subset of children who retold the story orally recalled 36.6%, 43.8%, 35.7%, and 21.4% of the central story elements in the story mapping 1, story mapping 2, DRTA, and control groups, respectively. In support of the first statistical finding, these percentages indicate that instruction in the comprehension strategy of identifying story structure elements advances first-grade children's reading comprehension more than no instruction in a comprehension strategy. Furthermore, these percentages suggest that instruction in identifying story structure elements and making predictions advances first-grade children's reading comprehension more than no instruction in a comprehension strategy. To support the second statistical finding, these percentages indicate that instruction in identifying story structure elements advances first-grade children's reading

comprehension more than instruction in making predictions. However, this finding applies only to the children in the story mapping 2 group (43.8%) compared to the children in the DRTA group (35.7%). The children in story mapping 1 group (36.6%) performed nearly equivalent to the children in the DRTA group (35.7%).

Additional studies confirm Baumann and Bergeron's (1993) statistical findings that suggest instruction in reading comprehension strategies advances children's reading comprehension (e.g., Brown, Pressley, Van Meter, & Schuder, 1996). However, researchers have conducted a limited number of these studies on children in kindergarten, first-, and second-grade (see Pearson & Duke, 2002 for a review). Interestingly, a few studies have found that instruction in reading comprehension strategies promotes children's listening comprehension (e.g., Morrow, 1984).

*Research on reading comprehension monitoring.* Researchers have investigated children's reading comprehension monitoring by (a) designing texts that contain obstacles, such as text contradictions that violate previously expressed ideas, assumed to interfere with the children's understanding and (b) assessing their ability to detect them (e.g., Kinnunen et al., 1998). In addition, researchers have assessed children's ability to employ various standards to monitor their reading comprehension, such as semantic standards to monitor the construction of meaning by detecting contradictions that violate previously expressed ideas and detect falsehoods that violate prior knowledge (Baker & Brown, 1984).

In a study of Finnish-speaking, first-grade children, Kinnunen et al. (1998) investigated the children's reading comprehension monitoring approximately mid-way through their school year. Kinnunen et al. randomly selected the children from among

children in four first-grade classrooms ( $N = 132$ ). The researchers gathered data on the children's detection of (a) lexical obstacles, syntactical obstacles, and falsehoods that violated prior knowledge embedded in sentences and (b) lexical obstacles, syntactical obstacles, and contradictions that violated previously expressed ideas embedded in passages, theorized to interfere with the children's construction of meaning. From these data the researchers made inferences about the children's reading comprehension monitoring. The children read these sentences and passages on computers that recorded (a) the amount of time children spent reading target words in sentences and target sentences in passages and (b) the number of times the children looked back from the target the words and sentences or returned to the target words or sentences from subsequent words or sentences. Kinnunen et al. computed a comprehension monitoring mean consistency score, defined as the number of detected obstacles in relation to the total number of obstacles. They computed this score for the time spent reading the target words or sentences and the number of target-related lookbacks in sentences or passages.

When reading target words in sentences, the children ( $N = 127$ ) produced mean consistency scores of 62% and 20% corresponding to the time spent reading target words and the number of target-related lookbacks, respectively. This finding suggests that when reading sentences, first-grade children monitor their reading comprehension with their consistency of monitoring being higher when measured as time spent reading target words than when measured as the number of target-related lookbacks. Additionally, these children produced mean number of seconds (range not specified) of 7.39 ( $SD = 4.31$ ), 6.15 ( $SD = 3.29$ ), 5.22 ( $SD = 3.36$ ), and 3.15 ( $SD = 1.82$ ) corresponding to lexical obstacles, syntactical obstacles, falsehoods, and no intended obstacles, respectively. The



children spent more time on (a) lexical obstacles than on syntactical obstacles, (b) syntactical obstacles than on falsehoods, and (c) falsehoods than on no intended obstacles, and these differences reached statistical significance. These results suggest that first-grade children spend the most amount of their reading comprehension monitoring time on lexical and syntactical obstacles. The children also produced the mean number of target-related lookbacks (range 0 - 1) of .22 ( $SD = .32$ ), .28 ( $SD = .38$ ), .23 ( $SD = .34$ ), and .08 ( $SD = .18$ ) corresponding to lexical obstacles, syntactical obstacles, falsehoods, and no intended obstacles, respectively. The children produced more lookbacks when encountering lexical obstacles, syntactical obstacles, and falsehoods than when encountering no intended obstacles at statistically significant levels. These findings suggest that first-grade children (a) produce approximately the same number of lookbacks corresponding to lexical obstacles, syntactical obstacles, and falsehoods and (b) produce fewer lookbacks corresponding to no intended obstacles.

Although the children monitored their reading comprehension when reading target words in sentences, they encountered difficulty when reading target sentences in passages. The children ( $N = 90$ ) produced mean consistency scores of 23% and 2% corresponding to the time spent reading target sentences and the number of target-related lookbacks, respectively. This finding indicates that when reading passages, first-grade children monitor their reading comprehension infrequently with their consistency of monitoring being higher when measured as time spent reading target sentences than when measured as the number of target-related lookbacks.

No other research studies have looked at the reading comprehension monitoring of children below third grade. However, some studies have examined listening

comprehension monitoring of children below third grade (e.g., Baker 1984). Further, although some studies have looked at the general monitoring of the reading process of children below third grade (e.g., Juliebo, Malicky, & Norman, 1998), no studies have looked at reading comprehension monitoring specifically.

*Reading comprehension processing and the interactive model of early literacy development.* Reading comprehension processing is one of several early literacy outcomes in the interactive model of early literacy development that guides this study (see Figure 1). This model can be used to understand the development of a child's reading comprehension processing. When a child constructs meaning of a text, his/her central processor may process information from the text reading activity. A child's central processor may also employ various strategies, such as reading comprehension monitoring and reading comprehension strategies. Furthermore, a child's central processor may access information from the many knowledge sources. For example, a child's central processor may access lexical knowledge to identify and retrieve information about word meanings. To aid in reading comprehension processing at the word, phrase, sentence, and/or text level, a child's central processor may build and activate memory for text. A child's central processor may access knowledge of language structures, such as semantic and syntactic information, to use contextual information in the construction of meaning and/or the monitoring of reading comprehension. A child's central processor may draw upon conceptual, factual, experiential, and/or schematic knowledge of the world to aid in the construction of meaning and reading comprehension monitoring. To evaluate and regulate the application of the various knowledge sources and the construction of meaning, a child's central processor may also access

metacognitive knowledge. A child's reading comprehension processing is observed as an early literacy outcome.

Clay (1993) designed the Reading Recovery lesson components and corresponding teaching procedures to develop reading comprehension processing in regards to reading comprehension monitoring. A Reading Recovery child receives explicit instruction in reading comprehension monitoring from the start of his/her tutoring. For example, a child learns to evaluate and regulate his/her reading comprehension by noticing his/her oral reading errors and subsequently searching for and using semantic cues to self-correct them. However, instruction in reading comprehension monitoring is only one aspect of reading comprehension processing. Reading comprehension processing also involves a child (a) constructing mental representations by accessing the information or content contained directly in the text and integrating this information or content with his/her background knowledge, (b) accessing and using conceptual knowledge and knowledge of word meanings, and (c) applying reading comprehension strategies. Although a child receives instruction in reading comprehension monitoring, s/he does not receive instruction in these other aspects of reading comprehension processing because Clay's teaching procedures do not foster the development of them.

*Summary.* Reading comprehension is the ultimate goal of reading. Reading comprehension processing involves the application of reading comprehension strategies and reading comprehension monitoring. Researchers have conducted a limited number of studies on reading comprehension strategies (e.g., Baumann & Bergeron, 1993) and reading comprehension monitoring (i.e., Kinnunen et al., 1998) of children below third-

grade. Yet the results of these studies suggest that these children (a) improve their reading comprehension processing when provided instruction in a reading comprehension strategy and (b) monitor their reading comprehension. Reading comprehension processing is one of several early literacy outcomes in the interactive model of early literacy development that guides this study (see Figure 1). The current study compares recommended and discontinued Reading Recovery children on reading comprehension processing by examining their ability to respond correctly to text-related questions following their tutoring.

### *Reading Recovery Tutoring*

Reading Recovery is a tutoring program designed by Clay (1993) to accelerate the early literacy development of the low-performing, six-year-old children so that they achieve average levels of classroom performance (Clay, 1993). In the United States, first-round Reading Recovery receive daily, 30-minute tutoring lessons from trained Reading Recovery teachers for approximately 20 weeks. Yet some of these first-round children do not respond well to their tutoring and fail to meet the criteria for successful performance. This section discusses Reading Recovery tutoring by examining the effectiveness of it and children's responsiveness to it. This section also connects the interactive model of early literacy development that guides this study (see Figure 1) to the effectiveness of Reading Recovery tutoring and children's responsiveness to it.

### *Reading Recovery Tutoring Effectiveness*

Reviews of research have conducted independent secondary analyses on the effectiveness of Reading Recovery tutoring by examining pamphlet summaries and/or unpublished technical reports (e.g., Wasik & Slavin, 1993). A meta-analysis of reading

tutoring programs featured several research studies on Reading Recovery (Elbaum et al., 2000). Additionally, research studies published in a highly respected journal have investigated the effectiveness of Reading Recovery tutoring (Center et al., 1995; Pinnell et al., 1994). The interactive model of early literacy development that guides this study can be used to explain the effectiveness of Reading Recovery tutoring (see Figure 1).

*Reviews of research on Reading Recovery tutoring effectiveness.* Wasik and Slavin (1993) conducted a secondary analysis on the effectiveness of Reading Recovery tutoring by analyzing data from two unpublished technical reports. The researchers compared the combined performance of recommended and discontinued Reading Recovery children to the performance of non-tutored, low-achieving children in first-grade in two cohorts (i.e., a pilot cohort and a first year cohort) on the tasks of *An Observation Survey of Early Literacy Achievement* (Clay, 2002) upon the immediate completion of tutoring at the end of first-grade. Wasik and Slavin reported (a) moderate to large effect sizes ( $ES = .57 - 1.03$ ) corresponding to Clay's Concepts about Print, Writing Vocabulary, Hearing and Recording Sounds in Words, and Text Reading tasks and (b) moderate to low effect sizes ( $ES = .13 - .40$ ) corresponding to Clay's Letter Identification and Word Reading tasks due to ceiling effects. These findings indicate that Reading Recovery tutoring accelerated the early literacy development of Reading Recovery children relative to other non-tutored, low-achieving children in first-grade upon the immediate completion of tutoring on early literacy measures designed by Clay.

To assess the immediate and longitudinal effectiveness of Reading Recovery tutoring, Wasik and Slavin (1993) also analyzed data on these same two cohorts of children from two other unpublished technical reports. The researchers compared the

combined performance of recommended and discontinued Reading Recovery children to the performance of non-tutored, low-achieving children in first-grade on the Clay's (2002) Text Reading task upon the immediate completion of tutoring at the end of first-grade, one year following tutoring at the end of second-grade, and two years following tutoring at end of third-grade. Upon the immediate completion of tutoring, the researchers reported large effect sizes corresponding to the Reading Recovery children and non-tutored, low-achieving children for the pilot cohort ( $ES = .72$ ) and the first year cohort ( $ES = .78$ ), favoring Reading Recovery. However, the researchers found that the corresponding effect sizes diminished from the immediate completion of tutoring, to one year following tutoring, to two years following tutoring for the pilot cohort ( $ES = .72, .29, .14$ ), as well as the first year cohort ( $ES = .78, .46, .25$ ). These findings suggest that Reading Recovery tutoring accelerated the reading development of Reading Recovery children relative to other non-tutored, low-achieving children in first-grade upon the immediate completion of tutoring, but failed to maintain this acceleration one and two years after tutoring on a measure of oral reading accuracy developed by Clay. The two groups of children became more similar over time.

Hiebert (1994) also conducted a secondary analysis on the immediate effectiveness of Reading Recovery tutoring by examining data in technical reports and pamphlet summaries from three large Reading Recovery training sites. She calculated the mean reading level of recommended and discontinued Reading Recovery children on Clay's (2002) Text Reading task for each year of Reading Recovery implementation in several Reading Recovery sites. According to Hiebert, a high percentage of these children attained an average reading level between first- and second-grade by the end of first-

grade. This finding suggests that Reading Recovery tutoring accelerated the reading development of Reading Recovery children to average reading levels upon the immediate completion of tutoring, as evidenced by their performance on a measure of oral reading accuracy designed by Clay.

In their independent secondary analysis on the effectiveness of Reading Recovery tutoring, Shanahan and Barr (1995) also examined unpublished technical reports. The researchers compared discontinued Reading Recovery children to (a) non-tutored, low-achieving children in first-grade and (b) non-tutored, average-achieving children in first-grade by examining their gain scores on Clay's (2002) Text Reading task from first-through third-grade. In regards to the non-tutored, low-achieving children, Shanahan and Barr found that the Reading Recovery children and the non-tutored, low-achieving children progressed at approximately the same rate (a) from the end of first-grade to the end of second-grade and (b) from the end of second-grade to the end of third-grade. In regards to the non-tutored, average-achieving children, the researchers found that the Reading Recovery children progressed at a slower rate than the non-tutored, average-achieving children from the end of first-grade to the end of second-grade. However, the two groups progressed at approximately the same rate from the end of second-grade to the end of third-grade. These findings indicate that the reading development of Reading Recovery children and non-tutored, low-achieving children progressed at the same rate during second- and third-grade, as evidenced by their gain score performance on a measure of oral reading accuracy designed by Clay. The findings also suggest that although the reading development of Reading Recovery children progressed at a slower rate than non-tutored, average-achieving children during second-grade, the two groups

progressed at a similar rate during third-grade, as evidenced by their gain score performance on measure of oral reading accuracy designed by Clay.

In their recent meta-analysis, Elbaum et al. (2000) reviewed experimental and quasi-experimental studies that investigated the effectiveness of Reading Recovery tutoring relative to other first-grade tutoring programs upon the immediate completion of tutoring. The researchers calculated a moderate mean weighted effect size for Reading Recovery tutoring ( $ES = .66$ ) and a low mean weighted effect size for other first-grade tutoring ( $ES = .29$ ); the difference between these two effect sizes reached statistical significance. These findings indicate that Reading Recovery tutoring accelerates the early literacy development of first-grade children better than other tutoring programs. However, Elbaum et al. noted that some of the Reading Recovery studies included in their analysis had inflated results because the researchers in these studies failed to report data on all the Reading Recovery children who received tutoring.

Although these reviews regarded Reading Recovery as a generally effective tutoring program, they offered criticisms concerning the studies' research designs. Some of the Reading Recovery studies failed to (a) assign children randomly or match children to equivalent control groups, (b) account for regression to the mean, (c) check fidelity of treatment, (d) use standardized measures, (e) use measures independent of Reading Recovery, (f) use residual gain scores, (g) account for attrition in longitudinal analyses, and (h) include all tutored children in data analyses (e.g., Shanahan & Barr, 1995). Consequently, the results from some of the Reading Recovery tutoring studies and the independent, secondary analyses overestimated the effects of Reading Recovery tutoring.

Secondary analyses on the effectiveness of Reading Recovery tutoring examined



unpublished technical reports and pamphlet summaries. Additionally, Elbaum et al.'s (2000) meta-analysis examined experimental and quasi-experimental studies that investigated the effectiveness of Reading Recovery tutoring and compared these results to the results from other experimental and quasi-experimental studies that investigated the effectiveness of other first-grade tutoring programs. The next section describes two experimental studies that investigated the effectiveness of Reading Recovery tutoring. A respected journal published these studies.

*Research on Reading Recovery tutoring effectiveness.* Pinnell et al. (1994) conducted an experimental study that compared the effectiveness of Reading Recovery tutoring to three other treatments offered to other low-achieving, first-grade children (N = 324). The Reading Recovery group received Reading Recovery tutoring from certified teachers who received the standard, one year Reading Recovery training. The instruction focused on teaching reading and writing strategies. The Reading Success group received tutoring modeled after Reading Recovery tutoring from certified teachers who received a two-week Reading Recovery training. The instruction focused on teaching reading and writing strategies and the teachers used Reading Recovery materials. The Direct Instruction Skills Plan group received tutoring from certified teachers who received three days of training in reading skills instruction. The instruction focused on sequential teaching of skills and the application of these skills in reading and writing contexts, using a variety of teacher selected teaching techniques. The Reading and Writing group received small group instruction from certified teachers who had previously received the standard, one year Reading Recovery training. The instruction focused on teaching reading and writing strategies, and the teachers used Reading Recovery materials. The

control groups received supplemental instruction from Chapter 1 teachers who received no training, aside from the typical Chapter 1 training. The instruction focused on decoding skills and the development of a core of familiar words. The Reading Recovery, Reading Success, and Direct Instruction Skills Plan groups received daily tutoring for 30 minutes from October to February, whereas the Reading and Writing and control groups received daily instruction for 30-45 minutes from October to May.

Pinnell et al. (1994) identified 10 school districts with four schools in each district. One school in each district had Reading Recovery. The researchers randomly assigned the remaining three schools in each district to one of the following treatment groups: Reading Success, Direct Instruction Skills Plan, and Reading and Writing. The teachers identified the 10 lowest-performing children in each school and randomly assigned four of these children to the treatment assigned to their school. The six remaining children in each school formed a randomized control group for that corresponding treatment. The researchers combined these control groups to form four control groups, one corresponding to each treatment.

Pinnell et al. (1994) conducted a time and content analysis by viewing videotapes of the children in the treatment and control groups and recording the amount of time the children spent on the following: activities involving children reading texts, activities involving children writing texts, and activities not involving reading or writing texts, such as isolated word analysis activities. The Reading Recovery group produced mean percentages of time of 60.2%, 25.3%, and 14.5% on reading, writing, and other activities, respectively. Although the Reading Recovery group's percentages paralleled the Reading Success group's percentages (62.3%, 28.8%, and 8.9%), the Reading Recovery group's

percentages were greater than the Direct Instruction Skills Plan group (29.9%, .3%, and 69.8%), the Reading and Writing group (26.8%, 23.4%, and 49.8%), and the combined control group across all four treatments (21.0%, 3.1%, and 75.9%). These qualitative results suggest that the instructional activities in Reading Recovery tutoring lessons center on the reading and writing of texts.

Pinnell et al.'s (1994) statistical analyses involved effect size estimates that compared each treatment group to their own control group in their own schools, combined across all schools in the 10 school districts. Upon the immediate completion of tutoring in February, the Reading Recovery group produced (a) a large standardized effect size estimate on Clay's (2002) Text Reading task ( $ES = 1.50$ ) and (b) medium estimates on Clay's Hearing and Recording Sounds in Words task ( $ES = .65$ ) and two standardized measures of reading achievement ( $ES = .49, .51$ ). All of these estimates reached statistical significance. These findings indicate that Reading Recovery tutoring accelerated the early literacy development of low-achieving, first-grade children better than its respective control group upon the immediate completion of tutoring on early literacy measures dependent and independent of Reading Recovery.

In comparison to the Reading Recovery group's standardized effect size estimates, the Reading Success group produced (a) medium estimates on Clay's (2002) Text Reading ( $ES = .45$ ) and Hearing and Recording Sounds in Words tasks ( $ES = .45$ ) and (b) small estimates on the two standardized measures of reading achievement ( $ES = .04, .27$ ), with the estimates reaching statistical significance on Clay's tasks. The Direct Instruction Skills Plan group produced low estimates on all four measures ( $ES = -.03, -.05, .25, .14$ ). The Reading and Writing group produced (a) a medium estimate on Clay's

Text Reading task ( $ES = .41$ ) that reached statistical significance and (b) small estimates on the remaining three measures ( $ES = .14, .23, .23$ ). These comparisons suggest that Reading Recovery tutoring with its year-long training when compared to its respective control group produced larger standardized effect size estimates than (a) Reading Recovery tutoring with an abbreviated Reading Recovery training, (b) a tutoring program with a different instructional focus, and (c) small group instruction with year-long Reading Recovery training when compared to their respective control groups.

At the end of first-grade in May, the Reading Recovery group produced a low standardized effect size estimate ( $ES = .19$ ) on one of the standardized measure of reading achievement administered in February. This finding indicates that Reading Recovery tutoring did not accelerate the early literacy development of low-achieving, first-grade children better than its respective control group three months following the completion of tutoring on a standardized measure of reading achievement independent of Reading Recovery. It also suggests that the effect of Reading Recovery tutoring diminished from the immediate completion of tutoring to three months following tutoring.

In comparison to the Reading Recovery group's standardized effect size estimates, the Reading Success group ( $ES = -.14$ ) and the Direct Instruction Skills Plan group ( $ES = -.05$ ) also produced low estimates on this same standardized measure. These comparisons suggest that Reading Recovery tutoring with its year-long training when compared to its respective control group produced slightly larger estimates than (a) Reading Recovery tutoring with an abbreviated Reading Recovery training and (b) a tutoring program with a different instructional focus when compared to their respective

control groups. In an additional comparison to the Reading Recovery group's standardized effect size estimate, the Reading and Writing group produced a higher estimate ( $ES = .34$ ). This comparison indicates that Reading Recovery tutoring with its year-long training when compared to its respective control group produced a smaller estimate than small group instruction with year-long Reading Recovery training when compared to its respective control group. The researchers (a) noted that the Reading and Writing group received instruction from October to May, whereas the Reading Recovery group received instruction from October to February and (b) offered this extended instructional time as an explanation for the Reading and Writing group's superior performance.

At the beginning of second-grade in October, the Reading Recovery group produced a small and large standardized effect size estimate on Clay's (2002) Hearing and Recording Sounds in Words task ( $ES = .35$ ) and Text Reading task ( $ES = .75$ ), respectively, with both estimates reaching statistical significance. These findings indicate that Reading Recovery tutoring accelerated the early literacy development of low-achieving, first-grade children better than its respective control group eight months following the completion of tutoring on early literacy measures designed by Clay.

In comparison to the Reading Recovery group's standardized effect size estimates, the Reading Success group ( $ES = .00, .07$ ) and the Direct Instruction Skills Plan group ( $ES = -.25, .06$ ) produced low estimates on Clay's (2002) Hearing and Recording Sound in Words and Text Reading tasks. These comparisons suggest that Reading Recovery tutoring with its year-long training when compared to its respective control group produced larger estimates than (a) Reading Recovery tutoring with an

abbreviated Reading Recovery training and (b) a tutoring program with a different instructional focus when compared to their respective control groups. In an additional comparison of the Reading Recovery group's standardized effect size estimate, the Reading and Writing group produced an estimate on the Hearing and Recording Sounds in Words task ( $ES = .29$ ) that approached the Reading Recovery group's estimate. This comparison suggests that Reading Recovery tutoring with its year-long training when compared to its respective control group produced slightly larger estimates than small group instruction with year-long Reading Recovery training when compared to its respective control group on the Hearing and Recording Sounds in Words task. Furthermore, the Reading and Writing group produced a low estimate ( $ES = .32$ ) on the Text Reading task. This comparison suggests that Reading Recovery tutoring with its year-long training when compared to its respective control group produced larger estimates than small group instruction with year-long Reading Recovery training when compared to its respective control group. Even though the training of the Reading Recovery group and the Reading and Writing group was the same, the instructional format differed (i.e., tutoring vs. small group), as well as the length of instruction.

Center et al. (1995) research supports Pinnell et al.'s (1994) findings that Reading Recovery children perform better than non-tutored, low-achieving children in a control group. Center et al. compared Reading Recovery children to non-tutored, low-achieving children assigned randomly to a control group that received regular classroom instruction, as well as "resource assistance typically available to at-risk readers after 1 year of school" (Center et al., 1995, p. 246). Similar to Pinnell et al., Center et al. found that the Reading Recovery group performed better than the control group on all measures of early literacy

development, including measures of oral reading accuracy, word reading, orthographic knowledge, phonemic awareness, syntactic awareness, phonological recoding upon the immediate completion of tutoring and four months after tutoring. The children's performance reached statistical significance on all these measures except (a) the phonemic and syntactic measures at the immediate completion and (b) the syntactic awareness and phonological recoding measures four months following tutoring.

Center et al.'s (1995) results further support Pinnell et al.'s (1994) finding that the effects of Reading Recovery tutoring diminish over time. A comparison of effect size estimates corresponding to the Reading Recovery group and the control group at the completion of tutoring and at four months following tutoring revealed a decrease in effects sizes from the immediate completion of tutoring to four months following tutoring. Furthermore, the Reading Recovery group failed to perform better than the control group on all measures of early literacy development at statistically significant levels with the exception of Clay's (2002) Text Reading task 12 months following the completion of tutoring. These findings demonstrate the diminishing effects of Reading Recovery tutoring.

Although additional studies report results that appear to support Center et al.'s (1995) and Pinnell et al.'s (1994) findings, these studies failed to assign children randomly or match children to equivalent control groups (e.g., Clay, 1993). This research design flaw renders the results of these other studies inclusive because the children may have obtained accelerated progress in early literacy development without Reading Recovery tutoring.

*Summary.* Secondary analyses have demonstrated the effectiveness of Reading

Recovery tutoring in accelerating the early literacy development of Reading Recovery children upon the immediate completion of tutoring (Hiebert, 1994; Shanahan & Barr, 1995; Wasik & Slavin, 1993). Yet these analyses have also shown the diminishing effects of Reading Recovery in the time following tutoring (Shanahan & Barr, 1995; Wasik & Slavin, 1993). Elbaum et al.'s (2000) meta-analysis found that Reading Recovery tutoring accelerated the early literacy development of first-grade children better than other tutoring programs at statistically significant levels. Furthermore, experimental studies have demonstrated that Reading Recovery tutoring accelerates the early literacy development of Reading Recovery children more than non-tutored, low-achieving, first-grade children in other treatment groups (e.g., Pinnell et al., 1994), as well as control groups (e.g., Center et al., 1995; Pinnell et al. 1994) upon their immediate completion of tutoring and a brief time following tutoring. Yet experimental studies have shown that the effects of Reading Recovery tutoring diminish over time (e.g., Center et al., 1995; Pinnell et al., 1994).

#### *Reading Recovery Tutoring Responsiveness*

Although reviews of research, a meta-analysis on tutoring programs, and experimental studies have found Reading Recovery a generally effective tutoring program, some children do not respond well to this individually designed diagnostic instruction. During the 2001-2002 school year, nearly one third of first-round Reading Recovery children in the United States did not to respond well to tutoring and failed to meet the criteria for successful performance. Subsequently, their Reading Recovery teachers assigned them to the recommended end-of-program status (Gómez-Bellengé et al., 2003). Interestingly, only a handful of research studies have examined the



performance of these recommended children (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004). These studies offer some insight into recommended children's early literacy development. The interactive model of early literacy development that guides this study can be used to explain Reading Recovery children's tutoring responsiveness (see Figure 1).

*Research on Reading Recovery tutoring responsiveness.* In their descriptive study of Reading Recovery children in New Zealand, Clay and Tuck (1991) examined recommended children who did not meet the criteria for successful performance ( $n = 140$ ), as well as discontinued children who met the criteria ( $n = 140$ ) and incomplete children who had incomplete tutoring programs due to the end of the school year ( $n = 140$ ). The researchers formed these groups from the total number of Reading Recovery children ( $n = 9,860$ ) in New Zealand in 1988. After the researchers identified all the recommended children ( $n = 390$ ) during the 1988 school year, they randomly selected the discontinued children ( $n = 988$ ) and the incomplete children ( $n = 906$ ) from among all the discontinued ( $n = 6,494$ ) and incomplete children ( $n = 2,976$ ) during the 1988 school year. Then, the researchers requested sets of lesson records from the Reading Recovery teachers for all three groups of children. Because complete sets of lesson records were unavailable for all the children in the groups, the researchers identified all the recommended children with complete sets of lesson records ( $n = 150$ ), and matched the discontinued children ( $n = 150$ ) and incomplete children ( $n = 150$ ) to the recommended children, forming triplet sets, controlling for school and teacher difference. The researchers reduced the 150 triplet sets to 140 due to some incomplete lesson record sets and the unavailability of recommended children.

Clay and Tuck (1991) investigated the services that the recommended children received upon their completion of tutoring by distributing questionnaires to their Reading Recovery teachers. The Reading Recovery teachers referred 82% of the recommended children to specialists for support services with some of the children receiving multiple and concurrent services. Specifically, the teachers reported that (a) 28% of the children received instruction from a special education teacher, (b) 23% of the children received instruction from a reading teacher, (c) 36% of the children received reading assistance from others in the school (e.g., teacher aide), and (d) 11.5% of the children received services from other school professionals (e.g., speech therapist). The remaining 13% of the recommended children did not receive support services because (a) the children's parents or guardians denied referral consent, (b) the children produced high scores on early literacy measures, (c) the children left the school, and/or (d) the school neglected to complete the referral process. These findings suggest that the majority of recommended children received support services.

In addition to studying the support services that recommended children received, Clay and Tuck (1991) also separately examined the recommended children's growth in achievement on Clay's (2002) Writing Vocabulary task, Hearing and Recording Sounds in Words task, Text Reading task, and on a measure of word reading by reporting their entry performance relative to their exit performance in raw scores or stanines. An examination of the recommended children's growth revealed considerable variation, as well as some overlap between the recommended and discontinued children. Comparisons of the children's growth revealed that the recommended children generally performed below the discontinued children on these four measures prior to and following tutoring.

These findings indicate that the recommended children lag behind the discontinued children in their early literacy development prior to and following their tutoring.

Two years after the recommended children completed their tutoring and their Reading Recovery teachers referred them to specialists for support services, Clay and Tuck (1991) conducted a small follow-up study on some these recommended children in the Auckland area ( $n = 44$ ). The Reading Recovery teachers administered a word reading measure, an orthographic knowledge measure, and an oral reading accuracy measure to these recommended children. The researchers grouped the data “in levels determined by the relationship of the Burt Word Reading Score [the word reading measure] to the age of the child - 1) below the norms, 2) more than two years below age, 3) more than one year below age, 4) at or above age level plus or minus one year” (Clay & Tuck, 1991, p. 40). The researchers found that 14% of the children performed below the norms, 50% performed more than two years below age level, 20% performed more than one year below age level, and 16% performed at or above age level plus or minus one year. This finding indicates that the majority of recommended children performed below age level two years after their completion of tutoring.

In their examination of the appropriateness of the 20-week tutoring limit for recommended children, Clay and Tuck (1991) found that Reading Recovery children spend an average of 12 to 15 weeks in tutoring with an upper limit of approximately 20 weeks. The researchers also found that children with the lowest entry scores on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and a measure of word reading require more tutoring time to meet the criteria for successful performance than children with higher entry scores. A correlational analysis produced low correlations ( $r =$

.04, -.26) corresponding to the recommended children's weeks in tutoring and entry scores, indicating that the number of weeks recommended children spend in tutoring is unrelated to their entry scores.

Clay and Tuck (1991) also established a predictive relationship between children's entry scores and their criterion performance. That is, Clay and Tuck found that recommended children generally produced lower entry scores on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and on a measure of word reading than discontinued children and were less likely to meet the criteria for successful performance than discontinued children. This finding indicates that the probability of failing to meet the criteria for successful performance is highest for children with low entry scores.

Other researchers have also investigated recommended children (Center et al., 1995; Chapman et al., 2001; Spector & Moore, 2004). Although these researchers did not design their studies with the primary purpose of examining and/or describing these recommended children, they provided information about their early literacy development by reporting mean scores and corresponding standard deviations (Center et al., 1995; Chapman et al., 2001), as well as reporting results from statistical analyses (Spector & Moore, 2004).

Although Center et al. (1995) evaluated the effectiveness of Reading Recovery tutoring, they also compared the early literacy development of a subset of recommended children to a subset of discontinued children prior to their tutoring. The subset of recommended children consisted of children who performed below grade level on most early literacy measures 12 months after tutoring ( $n = 8$ ), whereas the subset of discontinued children consisted of children who performed on grade level on all early

literacy measures 12 months after tutoring ( $n = 8$ ).

Prior to Reading Recovery tutoring, the subset of recommended children produced mean scores substantially below the subset of discontinued children on measures of phonemic awareness, syntactic awareness, and phonological recoding. Center et al. (1995) did not test for statistically significant differences between the two groups on these measures. These descriptive statistics suggest that the subset of recommended children begin their tutoring with less phonemic and syntactic awareness and weaker phonological recoding skills than the subset of discontinued children. Although Center et al. administered other early literacy measures of word reading, oral word accuracy, and orthographic knowledge to these two subsets of children, they failed to report the mean scores and corresponding standard deviations.

In a longitudinal study that examined the relation between the effectiveness of Reading Recovery tutoring and the development of phonological processing skills, Chapman et al. (2001) reported the pre- and post-tutoring means and standard deviations of recommended children on numerous early literacy measures on seven testing occasions. By reporting these mean scores, the researchers allowed for mean score comparisons of the recommended children ( $n = 6$ ) to the discontinued children ( $n = 26$ ). The researchers excluded the recommended children's mean scores from their statistical analyses.

On the testing occasions prior to and immediately following tutoring, the recommended children produced mean scores below the discontinued children on several measures that assessed letter knowledge, onset and rime awareness, phonemic awareness, graphophonemic awareness, word reading, phonological recoding, analogical transfer, oral

reading accuracy, and orthographic knowledge with the exception of the reading comprehension measure. On this reading comprehension measure the recommended children produced (a) the same mean score of 0.00 as the discontinued children on the testing occasion prior to tutoring and (b) approximately the same low mean score as the discontinued children immediately following tutoring. These findings indicate that the recommended children performed below the discontinued children on all early literacy measures prior to and immediately following tutoring with the exception of the reading comprehension measure. These findings also suggest that the recommended children produced the same or approximately the same low mean scores as the discontinued children on a measure of reading comprehension prior to and immediately following tutoring. In addition, the recommended children failed to achieve age-appropriate norms on (a) a standardized measure of word reading and (b) a standardized measure of oral reading accuracy and reading comprehension prior to and immediately following tutoring, indicating that the recommended children performed below age-appropriate levels.

Most recently, in an investigation that examined the relationship between Reading Recovery children's entry scores on various early literacy measures and their responsiveness to tutoring, Spector and Moore (2004) reported the pre-tutoring performance of recommended ( $n = 55$ ) and discontinued ( $n = 51$ ) children. Prior to tutoring, the recommended children performed below the discontinued children on the six tasks of *An Observation Survey of Early Literacy Achievement* (Clay, 2002), as well as on a phonemic awareness measure, a verbal short-term memory measure, and a rapid automatized naming measure. The differences in these mean scores reached statistical

significance on the phonemic awareness measure ( $d = .57$ ) and the verbal short-term memory measure ( $d = .65$ ). These findings indicate that the recommended children's early literacy development lags behind discontinued children prior to tutoring with phonemic awareness and verbal short-term memory capacity reaching statistical significance.

*Summary.* Overall, recommended children have produced lower mean scores than discontinued children prior to tutoring (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004) and immediately following tutoring (Chapman et al., 2001; Clay & Tuck, 1991) on various early literacy measures. Recommended children also performed below age level norms prior to tutoring (Chapman et al., 2001) and immediately following tutoring (e.g., Chapman et al., 2001; Clay & Tuck, 1991) on various measures of early literacy achievement.

*Reading Recovery Tutoring Effectiveness and Responsiveness and the Interactive Model of Early Literacy Development*

The interactive model of early literacy development that guides this study (see Figure 1) can be used to understand Reading Recovery tutoring effectiveness and children's responsiveness to it. This model features various knowledge sources and a child's central processor. Over the course of tutoring, a Reading Recovery child may advance in his/her development of the various knowledge sources as well as his/her processing ability. If the development of the various knowledge sources and processing occurs, a child may respond well to tutoring and meet the criteria for successful performance. If the development in one or more of the various knowledge sources and/or processing does not occur, a child may respond poorly to tutoring and fail to meet the

criteria for successful performance. This model also illustrates the interaction among a child's central processor, an early literacy activity, and the various knowledge sources. A Reading Recovery child may respond well or fail to respond well to tutoring, depending upon the occurrence of this interaction. Thus, the effectiveness of Reading Recovery tutoring, defined by a child's tutoring responsiveness, may be influenced by a child's development of various knowledge sources, a child's development of processing, and the occurrence of the interaction among a child's central processor, an early literacy activity, and the various knowledge sources.

### *Conclusion*

The research studies reviewed in this chapter indicated that phonological awareness (e.g., Iversen & Tunmer, 1993), orthographic knowledge (e.g., Treiman, 1993), oral reading processing (e.g., Cohen, 1975), and reading comprehension processing (e.g., Baumann & Bergeron, 1993; Kinnunen et al., 1998) underlie children's early literacy development. Reviews of research (e.g., Shanahan & Barr, 1995) and research studies (e.g., Pinnell et al., 1994) have found Reading Recovery tutoring generally effective. Yet nearly one third of first-round Reading Recovery children in the United States do not respond well to their tutoring and fail to meet the criteria for successful performance (Gómez-Bellengé et al., 2003). To date, four studies have examined the early literacy development of these recommended children prior to and following their tutoring (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004). Overall, these studies have indicated that recommended children perform below discontinued children, as evidenced by the comparison of their mean score performance on various early literacy measures prior to and following their



tutoring. The interactive model of early literacy development that guides this study (see Figure 1) can be used to understand a child's development of phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. This model can also be used to understand the effectiveness of Reading Recovery tutoring and a child's responsiveness to it. To contribute to the understanding of recommended children's early literacy development and expand upon the studies that have examined recommended children's early literacy development, I individually assessed recommended children's (a) phonological awareness and orthographic knowledge prior to and following their completion of tutoring and (b) oral reading processing and reading comprehension processing following their completion of tutoring. Then, I compared the recommended and discontinued children's pre- and post-tutoring performance. Chapter 3 describes the current study's setting, measures, procedures, and data analyses.

## Chapter 3

### *Methodology*

The current study compared recommended and discontinued Reading Recovery children on (a) phonological awareness and orthographic knowledge prior to and following their completion of tutoring and (b) oral reading processing and reading comprehension processing following their completion of tutoring. This study posed four research questions:

- (a) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on phonological awareness prior to and following their completion of tutoring?
- (b) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring?
- (c) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring?
- (d) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring?

Chapter 3, organized into four sections, describes the current study's research methodology. Section 1 describes the setting, including the Reading Recovery schools, children, and teachers. Section 2 summarizes two pilot studies, defines several measurement terms, and describes the early literacy measures. Section 3 presents the pre-

and post-tutoring procedures. Section 4 outlines the data analysis procedures, including repeated measure two-way analyses of variance, chi-square tests of independence, and one-way analyses of variance.

### *Setting*

#### *Reading Recovery Schools*

Eight schools with the Reading Recovery program in a large school district in a Mid-Atlantic state participated in the current study. Although this district had nine schools with the Reading Recovery program, only eight schools participated; one school declined the invitation to participate. Each of these eight schools varied in the number of years of Reading Recovery implementation, ranging from one to nine years. Furthermore, each of these eight schools had one to four Reading Recovery teachers, depending upon a school's first-grade enrollment and determination of need. From among the eight schools, seven schools supported their Reading Recovery programs through Title 1 funds, and one school supported its Reading Recovery program through school-based funds.

The school district's Reading Recovery site coordinator and Reading Recovery teacher leader verified that these eight schools with the Reading Recovery program conformed to the 'Standards and Guidelines for Operation of a Site', as outlined in *Standards and Guidelines of the Reading Recovery Council of North America* (RRCNA, 2001). The site coordinator and the teacher leader shared the responsibility of overseeing the school district's Reading Recovery program. The site coordinator managed the administrative aspects of the program (e.g., budget), whereas the teacher leader managed the instructional aspects (e.g., teacher training and/or continuing professional development).

### *Reading Recovery Children*

At the onset, this study included 60 first-round Reading Recovery children. At the completion of this study, 55 first-round Reading Recovery children remained. Three children failed to finish their tutoring because they moved (5.0%). Two additional children failed to finish their tutoring because the Reading Recovery teacher leader removed them from their tutoring (3.3%) due to a parent request and a kindergarten placement. From among these 55 children, 29 children failed to respond well to tutoring (48.3%) and their Reading Recovery teachers assigned them to the recommended end-of-program status category, and 26 children responded well to tutoring (43.3%) and their Reading Recovery teachers assigned them to the discontinued end-of-program status category.

From among the 55 Reading Recovery children, 41 children (75%) were Caucasian, 9 children (16%) were African American, 3 children (5%) were Hispanic, 1 child (2%) was Asian, and 1 child (2%) was American Indian. The majority of the children were male (75%). The pre-tutoring age of the children ranged from five years ten months to seven years four months, whereas the post-tutoring age of the children ranged from six years three months to seven years nine months. The recommended and discontinued children had the same mean age of 6 years 8 months. The majority of the children's native language was English (93%). Less than half of the children participated in the Department of Agriculture's National School Lunch Program (42%).

### *Reading Recovery Teachers*

Fifteen Reading Recovery teachers and one teacher leader tutored the Reading Recovery children. Due to a lack of knowledge of and proficiency in the Reading

Recovery teaching procedures, I excluded two Reading Recovery teachers who were in their training year and their children from this study. The Reading Recovery teachers' and the teacher leader's general teaching experience ranged from six to 26 years, and their Reading Recovery teaching experience ranged from two to nine years. Each Reading Recovery teacher and the teacher leader tutored four children on a daily basis.

### *Measures*

In this section, I summarize the two pilot studies that guided my selection of phonological awareness tasks for this study. Next, I define some measurement terms. Then, I describe the measure that the Reading Recovery teachers administered to the Reading Recovery children: *An Observation Survey of Early Literacy Achievement* (Clay, 2002). Finally, I describe the measures that I administered to the Reading Recovery children: the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), the Sentence Writing and Spelling task (DeFord, 2000), and the Gray Oral Reading Tests-Fourth Edition (GORT-4) (Wiederholt & Bryant, 2001). I report the reliability and validity information of these measures.

#### *Pilot Study 1 and 2*

To select the phonological awareness tasks for this study, I conducted two pilot studies (see Appendix A). The primary aim of these pilot studies was to select phonological awareness tasks that were neither too easy nor too difficult for Reading Recovery children to complete prior to and following their tutoring.

In the first pilot study, I individually administered seven phonological awareness tasks to kindergarten children ( $N = 40$ ) upon their completion of their school year (see Table 2). Due to assessment session time constraints, I modified these tasks by (a)

Table 2

*Phonological Awareness Tasks Employed in Pilot Study 1 and 2*

Pilot Study 1 Modified Task	Citation	Pilot Study 2 Complete Task	Citation
Modified Rhyme Detection	Calfee et al., 1972	Rhyme Detection <sup>ab</sup>	Muter et al., 1997
Modified Rhyme Oddity	Bradley & Bryant, 1983	Rhyme Oddity <sup>a</sup>	Dodd et al., 2000
Modified Rhyme Production	MacLean et al., 1987 Stanovich et al., 1984		
Modified Sound Matching <sup>a</sup>	Wagner et al., 1999	Sound Matching <sup>a</sup>	Wagner et al., 1999
Modified Auditory Blending	Roswell & Chall, 1997	Blending Words <sup>ab</sup>	Wagner et al., 1999
Modified Phoneme Segmentation	Yopp, 1988	Sentence Writing and Spelling <sup>b</sup>	DeFord, 2000
Modified Phoneme Deletion	Bruce, 1964		

<sup>a</sup> Indicates a norm-referenced task. <sup>b</sup> Indicates the tasks selected for the current study.

selecting a reduced number of task and/or practice items, (b) developing practice items, and/or (c) adding picture support. When I modified these tasks, I rendered any normative information invalid and compromised the tasks' reliability and validity. I decided that the task difficulty information obtained from the children's performance on these modified tasks took precedence over any normative data information and the preservation of the tasks' reliability and validity.

The children's mean scores on the modified Rhyme Detection task, the modified Rhyme Oddity task, the modified Rhyme Production task, the modified Sound Matching task, and the modified Auditory Blending task revealed that the children responded correctly to more than half of the items on each task. Based on criteria that I developed, these mean scores indicated that these tasks were neither too easy nor too difficult for the kindergarten children at the end of their school year. Conversely, the children's mean scores on the modified Phoneme Segmentation task and the modified Phoneme Deletion task revealed that the children responded correctly to less than half of items on these two measures. According to my criteria, these mean scores indicated that these tasks were too difficult for the kindergarten children at the end of their school year.

Prior to the second pilot study, I replaced three of the modified tasks from the first pilot study that were not norm-referenced tasks with norm-referenced tasks that assessed the same underlying constructs. I replaced the modified Rhyme Detection task with Muter et al.'s (1997) Rhyme Detection task. I replaced the modified Rhyme Oddity task with Dodd, Crosbie, McIntosh, Teitzel, and Ozanne's (2000) Rhyme Oddity task. I also replaced the modified Auditory Blending task with Wagner et al.'s (1999) Blending Words task (see Table 2).

In the second pilot study, I individually administered five phonological awareness tasks to Reading Recovery children ( $N = 29$ ) prior to and following their tutoring. I administered the Rhyme Detection task (Muter et al., 1997), the Rhyme Oddity task (Dodd et al., 2000), the Sound Matching task (Wagner et al., 1999), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000). All the tasks were norm-referenced tasks, except DeFord's Sentence Writing and Spelling task (see Table 2).

The children's mean scores on the Rhyme Detection task (Muter et al., 1997) revealed that the entire sample of Reading Recovery children, as well as the two subgroups of recommended and the discontinued children, responded correctly to more than half of the items at pre- and post-tutoring. Based on the criteria that I developed relative to the recommended children's performance, the recommended children's mean scores indicated that this task was not too difficult for them to complete at pre-tutoring, but too easy for them at post-tutoring. Although the recommended children post-tutoring mean score hit the too easy mark, the corresponding standard deviation indicated a high degree of variability. Thus, I included this task in the current study.

The children's mean scores on the Rhyme Oddity task (Dodd et al., 2000) revealed that the entire sample of Reading Recovery children, as well as the two subgroups, responded correctly to approximately half of the items at pre- and post-tutoring. Based on the criteria that I developed relative to the recommended children's performance, the recommended children's mean scores indicated that this task was neither too easy nor too difficult for them to complete at pre- and post-tutoring. However, the mean scores for the entire sample of Reading Recovery children and the two



subgroups did not increase from pre- to post-tutoring, and the corresponding standard deviations indicated a high degree of variability at pre- and post-tutoring. Thus, I did not select this task for the current study. Extraneous cognitive requirements of memory demands and comparison skills may explain the lack of increase in mean scores from pre- to post-tutoring and the large standard deviations.

The children's mean scores on the Sound Matching task (Wagner et al., 1999) revealed that the Reading Recovery children, as well as the two subgroups, responded correctly to less than half or approximately half of the items at pre-tutoring and more than half of the items at post-tutoring. Based on the criteria that I developed relative to the recommended children's performance, the recommended children's mean scores indicated that this task was neither too easy nor too difficult for them to complete at pre- and post-tutoring. However, the standard deviation corresponding to the recommended children's mean score at post-tutoring indicated a high degree of variability. Thus, I did not select this task for the current study. Extraneous cognitive requirements of memory demands and comparison skills may explain this variability.

The children's mean scores on the Blending Words task (Wagner et al., 1999) revealed that the Reading Recovery children and the two subgroups responded correctly to less than half of the items at pre- and post-tutoring. According to the criteria that I fashioned relative to the recommended children's performance, the recommended children's mean scores indicated that this task was too difficult for them to complete at pre- and post-tutoring. However, because the ability to blend syllables, onsets and rimes, and phonemes into words is fundamental to reading development (National Reading Panel, 2000), I selected this task for the current study.

The children's mean scores on the Sentence Writing and Spelling task (DeFord, 2000) revealed that the Reading Recovery children and the two subgroups responded correctly to less than half or half of the items at pre-tutoring and more than half of the items at post-tutoring. According to the criteria that I fashioned relative to the recommended children's performance, the recommended children's mean scores indicated this task was not too difficult for them to complete at pre-tutoring, but too easy for them at post-tutoring. Although the post-tutoring mean score hit the too easy mark, the corresponding standard deviation indicated a high degree of variability. I included this task in the current study.

Based on my analyses of the descriptive data from two pilot studies, I selected tasks for the current study, theorized to be neither too easy nor too difficult for recommended Reading Recovery children to complete prior to and following their tutoring. I selected the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) for the current study (see Table 2)

### *Measurement Terms*

To describe the early literacy measures in this study, I used several measurement terms. I define these terms in this section. I developed the definitions of these terms from three different references: (a) *Reading Statistics and Research* (Huck, 2000), (b) *Methods of Educational and Social Science Research: An Integrated Approach* (Krathwohl, 1998) and (c) *Gray Oral Reading Tests* (4<sup>th</sup> ed.) (Wiederholt & Bryant, 2001). First I present the reliability terms. Then, I present the validity terms, followed by the item analysis terms.

*Reliability.* Reliability refers to the consistency of scores produced by a measure.

Reliability types include internal consistency reliability, interrater reliability, parallel split-half reliability, and test-retest reliability.

*Internal consistency reliability.* Internal consistency reliability determines whether all the items in a measure assess a single construct and the interpretability of scores. The internal consistency reliability procedures involve a single testing occasion and a single sample. Internal consistency reliability is assessed by using a split-half reliability procedure that (a) correlates scores on one half of a measure with scores on the other half of the measure from a single sample and (b) applies the Spearman-Brown formula to adjust the reliability coefficient in order to estimate whole test reliability, rather than half test reliability. The Cronbach's alpha procedure and the Kuder-Richardson 20 formula procedure provide estimates of the average split-half reliability of all possible random splits. The Cronbach's alpha procedure is used for a dichotomous response, (e.g., correct/incorrect response); whereas the Kuder-Richardson 20 formula procedure is used for either a non-dichotomous response (e.g., 5-point Likert response) or a dichotomous response (e.g., correct/incorrect response). A measure has high internal consistency reliability when the reliability coefficient approaches 1.0.

*Interrater reliability.* Interrater reliability refers to the extent that two or more raters, who independently score the same measure, agree. A measure has high interrater reliability when the coefficient approaches 1.0.

*Parallel split-half reliability.* Parallel split-half reliability determines whether different forms from the same measure are equivalent. Parallel split-half reliability correlates form A scores of a particular measure with form B scores of the same measure. The measure is given to a single sample on the same testing occasion; half of the sample

is given form A and the other half of the sample is given form B. A measure has high parallel split-half reliability when the reliability coefficient approaches 1.0.

*Test-retest reliability.* Test-retest reliability determines the stability of a construct over time. Test-retest reliability correlates scores from a measure given at one testing occasion with scores from the same measure given at a later testing occasion. The measure is given to a single sample of children on two different testing occasions. A measure has high test-retest reliability when the reliability coefficient approaches 1.0.

*Validity.* Validity refers to an evidence-based judgment that a measure assesses what it purports to assess. Validity types include construct validity (i.e., convergent validity), content validity, criterion-related validity (i.e., concurrent and predictive validity), and face validity.

*Construct validity.* Construct validity refers to the extent to which inferences from a measure's scores accurately reflect the construct that the measure purports to assess.

*Convergent validity.* Convergent validity refers to the extent to which a measure correlates with a related measure(s) of the same construct to provide evidence of construct validity.

*Content validity.* Content validity refers to the comparison of a measure's items with the measure's test specifications or blueprint to determine if the items on the measure assess the behaviors and content that the measure purports to assess.

*Criterion-related validity.* Criterion-related validity refers to the extent to which scores on a measure correlate with scores on a criterion measure or a criterion variable. Criterion-related validity includes concurrent validity and predictive validity.

*Concurrent validity.* Concurrent validity refers to the extent to which scores on a

measure correlate with scores on another measure that contains a relevant criterion. The measures are administered at the same time or within a short interval of time.

*Predictive validity.* Predictive validity refers to the extent to which scores on a measure administered at one point in time accurately forecast scores on another relevant criterion measure administered at a later point in time.

*Face validity.* Face validity refers to the casual, subjective inspection of a measure's items to determine if the items appear to assess the behaviors and content that the measure purports to assess.

*Item analysis.* Item analysis refers to a method of improving a test by correlating the items with either the total score or a criterion measure. Item analysis improves reliability by correlating the items with the total score, retaining only items that correlate highly with the total score, and changing other items to be like those retained. Item analysis improves validity by a similar process, except that a valid criterion measure is used instead of a total score.

*Item difficulty.* Item difficulty is an item analysis technique that calculates a percentage of examinees who correctly answer a given item. Test developers consider percentages between 15% and 85% acceptable for item selection.

*Item discrimination.* Item discrimination is an item analysis technique that determines the extent to which items that are supposed to measure the same characteristic or behavior are related and items that are supposed to measure different characteristics or behaviors are unrelated. Test developers use the resulting correlation coefficient as the criteria for item selection. Some test developers consider discrimination indexes of .20 or higher acceptable for item selection, whereas others consider discrimination indexes of

.35 or higher acceptable.

Researchers and test developers conducted reliability and validity analyses to establish and support the reliability and validity of the measures employed in this study. In the following section, I report the results of these analyses and use the measurement terms that I just defined.

#### *An Observation Survey of Early Literacy Achievement*

In this study, the Reading Recovery teachers administered *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to the children and used the children's pre- and post-tutoring scores to select children for tutoring and to discontinue them from tutoring. *An Observation Survey of Early Literacy Achievement* (Clay, 2002) consists of the following tasks: (a) Letter Identification, (b) Concepts About Print, (c) Word Reading, (d) Writing Vocabulary, (e) Hearing and Recording Sounds in Words, and (f) Text Reading (see Table 3).

Clay designed and developed *An Observation Survey of Early Literacy Achievement* (Clay, 2002) through her own systematic observation of children's early literacy behaviors from 1963 to 1978, beginning with her doctoral dissertation research. She initially published the Letter Identification, Concepts About Print, Word Reading, and Text Reading tasks in 1972, followed by the Writing Vocabulary and Hearing and Recording Sounds in Words tasks in 1979. Clay revised these tasks as she obtained more information from (a) teachers use of the tasks, (b) early literacy research, (c) a predictive validity study, and (d) theoretical discussions with colleagues on early literacy. Then, she published the subsequent revised tasks in 1985, 1993, and 2002 (Clay, 2002).

Clay designed *An Observation Survey of Early Literacy Achievement* (Clay, 2002)

Table 3

*Reading Recovery Teacher Administered Measure, Assessed Early Literacy Component, and Testing Occasions*

Task	Assessed early literacy component	Testing Occasion
<i>An Observation Survey of Early Literacy Achievement</i>		
LI	Letter Knowledge	Pre- and Post-tutoring
CAP	Knowledge of Printed Language Conventions	Pre- and Post-tutoring
WR	Reading Vocabulary	Pre- and Post-tutoring
WV	Writing Vocabulary	Pre- and Post-tutoring
HRSIW	Graphophonemic Awareness	Pre- and Post-tutoring
TR	Oral Reading Accuracy	Pre- and Post-tutoring

*Note.* LI = Letter Identification; CAP = Concepts About Print; WR = Word Reading; WV = Writing Vocabulary; HRSIW = Hearing and Recording Sounds in Words; TR = Text Reading.

for Reading Recovery and classroom teachers to systematically assess children's letter knowledge, knowledge of printed language conventions, reading vocabulary, writing vocabulary, graphophonemic awareness, and oral reading accuracy. Teachers use individual children's performance on this measure to guide and inform their early literacy instruction. Clay also designed this measure as an alternative to standardized measures to

provide teachers with information about individual children, rather than information about groups of children. Teachers use the observational data obtained from this measure to compare a child's performance on different testing occasions or to compare a child's performance to another child's performance (Clay, 2002).

The most recent edition of *An Observation Survey of Early Literacy Achievement* (Clay, 2002) contains reliability and validity information on each of Clay's tasks by referencing various reliability and validity studies. Clay and other researchers conducted the majority of these studies on children in New Zealand during the 1960s, 1970s, and 1980s. Pinnell and colleagues conducted some additional reliability and validity studies on children in the United States during the early 1990s. Recently, Clay conducted a construct validity analysis on children in New Zealand in 2000.

*Letter Identification.* The Letter Identification task (Clay, 2002) assessed a child's letter recognition. The child identified 26 upper- and 28 lower-case letters, including the print forms of *a* and *g*. The child identified each symbol by a name, a sound, or a word that began with the letter or sound. The Reading Recovery teacher recorded the child's responses and scored the responses as correct or incorrect. The teacher also recorded the type of correct response: alphabet response, letter-sound response, or word response. The task did not contain practice items. This task had one form.

As displayed in Table 4, an internal consistency reliability analyses produced high coefficients. Criterion-related validity analyses yielded a high-moderate concurrent coefficient (see Table 5) and moderate and high-moderate predictive coefficients (see Table 6). The construct validity analysis revealed moderate convergent coefficients (see Table 7). Although Clay (2002) reported that this task possessed content and face



Table 4

*Internal Consistency and Test-retest Reliability Information for An Observation Survey of Early Literacy Achievement*

Task	Reliability type	Sample size	Age/Grade	Location	Coefficient	Procedure	Time
LI	Internal Consistency	N = 100	6 years	NZ	.97	Split-half	-
LI	Internal Consistency	N = 107	K – 1 <sup>st</sup>	US	.95	Cronbach's Alpha	-
CAP	Internal Consistency	N = 100	6 years	NZ	.95	Split-half	-
CAP	Internal Consistency	N = 40	5 - 7 years	NZ	.85	Kuder-Richardson 20	-
CAP	Internal Consistency	N = 106	K – 1 <sup>st</sup>	US	.78	Cronbach's Alpha	-
CAP	Internal Consistency	N = 56	K	US	.84 - .88	Split-half	-
CAP	Test-retest	N = 56	K	US	.73 - .89	-	2 Weeks Apart
WR	Internal Consistency	N = 100	6 years	NZ	.90	Kuder-Richardson 20	-
WR	Internal Consistency	N = 107	K – 1 <sup>st</sup>	US	.92	Cronbach's Alpha	-
WV	Test-retest	N = 141	K – 1 <sup>st</sup>	US	.62	-	NA

Table 4 continued

*Internal Consistency and Test-retest Reliability Information for An Observation Survey of Early Literacy Achievement*

Task	Reliability type	Sample size	Age/Grade	Location	Coefficient	Procedure	Time
WV	Test-retest	N = 34	5 years	NZ	.97	-	NA
HRSIW	Internal Consistency	N = 107	K – 1 <sup>st</sup>	US	.96	Cronbach's Alpha	-
HRSIW	Internal Consistency	N = 403	1 <sup>st</sup>	US	.84 - .88	Split-half	-
HRSIW	Test-retest	NA	NA	NZ	.73 - .89	-	NA
TR	Internal Consistency	N = 96	K – 1 <sup>st</sup>	US	.83	Person Separation <sup>a</sup>	-
TR	Internal Consistency	N = 96	K – 1 <sup>st</sup>	US	.98	Item Separation <sup>a</sup>	-

*Note.* LI = Letter Identification; CAP = Concepts About Print; WR = Word Reading; WV = Writing Vocabulary; HRSIW = Hearing and Recording Sounds in Words; TR = Text Reading; NZ = New Zealand; US = United States; NA = Indicates the information was not available; - = Dash indicates the information was not applicable.

<sup>a</sup> Pinnell et al. (1994) reported that person separation and item separation reliability was established using Rasch rating scale analysis.

Table 5

*Criterion-related Validity Information for An Observation Survey of Early Literacy Achievement*

Task	Sample size	Age/Grade	Correlating word reading task or measure	Concurrent coefficient
LI	N = 100	6 years	Word Reading Task: 'Ready to Read' Word Test	.85
CAP	N = 100	6 years	Word Reading Task: 'Ready to Read' Word Test	.79
WR	N = 87	7 years	Schonell R1	.90
WV	N = 50	5 years	Word Reading Task: 'Ready to Read' Word Test	.82

*Note.* These concurrent validity studies were conducted in New Zealand. LI = Letter Identification; CAP = Concepts About Print; WR = Word Reading; WV = Writing Vocabulary.

Table 6

*Criterion-related Validity Information for An Observation Survey of Early Literacy Achievement*

Task	Predictive coefficient (7 years)		Predictive coefficient (8 years)	
	1	2	1	2
LI	.86	.81	.80	.83
CAP	.73	.64	.69	.70
WR	.90	.80	.88	.83
TR	.80	.69	.77	.72

*Note.* This predictive study was conducted on 83 children in New Zealand. LI = Letter Identification; CAP = Concepts About Print; WR = Word Reading; TR = Text Reading; 1 = Refers to the Schonell R1 word reading measure; 2 = Refers to the Fieldhouse word reading measure.

Table 7

*Construct Validity Information for An Observation Survey of Early Literacy Achievement*

Task	Correlating tasks	Convergent coefficients
LI	CAP, WR, WV, HRSIW, TR	.55 - .80
CAP	LI, WR, WV, HRSIW, TR	.50 - .80
WR	LI, CAP, WV, HRSIW, TR	.69 - .89
WV	LI, CAP, WR, HRSIW, TR	.58 - .89
HRSIW	LI, CAP, WR, WV, TR	.79 - .89
TR	LI, CAP, WR, WV, HRSIW	.77 - .89

*Note.* This study was conducted in New Zealand, based on 796 children, aged five to seven years old. LI = Letter Identification; CAP = Concepts About Print; WR = Word Reading; WV = Writing Vocabulary; HRSIW = Hearing and Recording Sounds in Words; TR = Text Reading.

validity, she failed to provide evidence to support the claim.

*Concepts About Print.* The Concepts About Print task (Clay, 2002) assessed a child's knowledge printed language conventions, such as word by word pointing and punctuation. The Reading Recovery teacher and the child actively participated in the task. The teacher read one of four texts, *Sand* (Clay, 1972), *Stones* (Clay, 1979), *Follow Me*, *Moon* (Clay, 2000), or *No Shoes* (Clay, 2000) to the child. As the teacher read the text to

the child, the teacher either asked the child questions or provided the child with directives. The child either answered the questions or pointed to certain text features. The teacher observed the child's responses, recorded them, and scored them as correct or incorrect. The task did not contain practice items. The task contained 24 items and had the same type of questions and directives for all four texts (see Table 8 for sample items).

As shown in Table 4, the internal consistency reliability analyses produced moderate, high-moderate, and high coefficients, and a test-retest reliability analysis yielded moderate and high-moderate coefficients. Criterion-related validity analyses yielded a moderate concurrent coefficient (see Table 5) and moderate predictive coefficients (see Table 6). The construct validity analysis revealed moderate convergent coefficients (see Table 7). Although Clay (2002) reported that this task possessed content and face validity, she failed to provide evidence to support the claim.

*Word Reading.* The Word Reading task (Clay, 2002) assessed a child's reading vocabulary by requiring the child to read from a list of high-frequency words. As the child read the words, the Reading Recovery teacher recorded the child's responses, scoring them as correct or incorrect. The Word Reading task had three versions: (a) 'Ready to Read' Word Test, (b) Ohio Word Test, and (c) Canberra Word Test. The teachers in this study used the Ohio Word Test. This test contained one practice item and three lists of 20 high-frequency words (i.e., list A, list B, and list C). The teachers administered list A at pre-tutoring and list B at post-tutoring (see Table 8 for sample words).

As displayed in Table 4, internal consistency reliability analyses produced high coefficients. Criterion-related validity analysis also yielded a high concurrent coefficient.

Table 8

*Sample Items and Passage Excerpts from An Observation Survey of Early Literacy**Achievement*


---

Task	Sample item and passage excerpt
CAP	<p>Word by Word Pointing: Point to it while I read.</p> <p>Punctuation: What's this for? (The teacher points to or traces a comma.)</p>
WR	and, the, pretty, has, down, where, after, let, here
HRSIW	The bus is coming. It will stop here to let me get on.
TR (Level 7)	<p>Mr. Jumble got a new camera. One day he went to the zoo. He took his new camera with him. Mr. Jumble saw a zebra by a tree. He yelled, "Hold it! I want to take your picture."</p>
TR (Level 12)	<p>Old Man Moss woke up and fell out of bed. He fell on the floor. He fell very, hard, and he landed on his head. Old Man Moss was so cold that he pulled the blanket off the bed. But he pulled so hard that he tore the blanket.</p>

---

*Note.* CAP = Concepts About Print; WR = Word Reading; HRSIW = Hearing and Recording Sounds in Words; TR = Text Reading.

(see Table 5) and high-moderate predictive coefficients (see Table 6). The construct validity analysis revealed moderate and high-moderate convergent coefficients (see Table 7). Although Clay (2002) reported that this task possessed content and face validity, she failed to provide evidence to support the claim.

*Writing Vocabulary.* The Writing Vocabulary task (Clay, 2002) assessed a child's ability to write words with correct spelling. The Reading Recovery teacher prompted the child to write all the words that s/he knew how to write. If the child stopped writing words or failed to write any words, the teacher suggested words to the child. Clay provided teachers with words, such as *is, my, to, see, the, come, and like* to suggest to the child. The task continued for 10 minutes or until the child exhausted his/her writing vocabulary. The teacher observed the child's written attempts, recorded them, and scored them by awarding one point to each word that the child wrote with correct spelling. Clay (2002) outlined specific rules for scoring (a) reversed letters; (b) words written right to left; (c) words written in a series, a rhyming set or spelling pattern group; and (d) capital letters. This task did not contain practice items.

As displayed in Table 4, a test-retest reliability analysis yielded a moderate coefficient, whereas another test-retest analysis yielded a high coefficient. Criterion-related validity analyses yielded a moderate concurrent coefficient (see Table 5). A construct validity analysis revealed moderate and high-moderate convergent coefficients (see Table 7). Although Clay (2002) reported that this task possessed content and face validity, she failed to provide evidence to support the claim.

*Hearing and Recording Sounds in Words.* The Hearing and Recording Sounds in Words task (Clay, 2002) assessed a child's graphophonemic awareness. First, the



Reading Recovery teacher read the complete dictation to the child. Then, the teacher dictated each word to the child, and the child recorded the graphemes that represented the phonemes in each word. The teacher may have prompted the child to attend to the sounds in the dictated words or to write the corresponding graphemes. The teacher scored the child's responses as correct or incorrect, evaluating the child's responses according to phoneme-grapheme appropriateness. Clay (2002) outlined specific rules for scoring (a) additions, omissions, and letters produced in an unusual order; (b) capital letters; (c) substitutions; (d) changes in letter order; (e) and reversed letters. The task contained five separate dictations (Form A-E), each containing 37 phonemes, but no practice items. The teachers administered Form D at pre-tutoring and Form A at post-tutoring (see Table 8 for a sample dictation).

As illustrated in Table 4, internal consistency reliability analyses produced high-moderate and high coefficients, and a test-retest reliability analysis yielded moderate and high-moderate coefficients. A construct validity analysis that intercorrelated the tasks revealed high-moderate convergent coefficients (see Table 7). Although Clay (2002) reported that this task possessed content and face validity, she failed to provide evidence to support the claim.

*Text Reading.* The Text Reading task (Clay, 2002) assessed a child's oral reading processing. After the Reading Recovery teacher provided the child a brief introduction to the text and showed the child some of the pictures in the text, the child read the text. As the child read, the teacher took a Running Record, an assessment of a child's oral reading processing. The teacher employed Clay's (2002) standardized conventions for recording a child's accurate responses, substitutions, omissions, insertions, self-corrections

and repetitions. The teacher recorded a child's performance at three levels of text difficulty: an easy text (.95 - 1.00 accuracy score), an instructional text (.90 - .94 accuracy score), or a hard text (below .89 accuracy score). The child continued to read higher levels of text until the accuracy score fell below .90 accuracy on two consecutive texts. The child read from the Scott Foresman Special Practice Books (1979), a graded set of texts that ranged in difficulty from level B (pre-primer) to level 30 (sixth-grade). If the child failed to read level B at or above .90 accuracy, the teacher placed the child at level A; level A had no corresponding text (see Table 8 for passage excerpts).

The Reading Recovery teacher calculated an accuracy score and a self-correction rate, and conducted a miscue analysis on the cues the child used and neglected to use to make errors and self-corrections. Although the Text Reading task provided the teacher with an accuracy score, a self-correction rate, and miscue analysis, it failed to provide a reading comprehension score. According to Clay (2002), "While the record [Running Record] does not give a measure of something labeled comprehension you can assess this objectively by recording the child's responses to your questions about the story and you can analyze the errors and self-corrections to find out how well the child works for meaning." (p. 11). Although Clay stated that asking a child to answer text-related questions about a story after reading it provides an assessment of reading comprehension, the teachers in this study did not ask the children to respond to text-related questions.

As shown in Table 4, internal consistency reliability analyses produced a moderate coefficient, as well as a high coefficient. An interrater reliability analysis on the recording and scoring of the children's substitutions, tolds, omissions, insertions, and self-corrections produced high ( $r = .98$ ) and moderate ( $r = .68$ ) coefficients, respectively.

As displayed in Table 6, a criterion-related validity analysis yielded moderate predictive coefficients. A construct validity analysis moderate and high-moderate convergent coefficients (see Table 7).

Reading Recovery teachers administered *An Observation Survey of Early Literacy Achievement* (Clay, 2002) at pre- and post-tutoring to assess the children's letter knowledge, knowledge of printed language conventions, reading vocabulary, writing vocabulary, graphophonemic awareness, and oral reading accuracy. Similarly, I administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) at pre- and post-tutoring to assess the children's phonological awareness and orthographic knowledge. I also administered the GORT-4 (Wiederholt & Bryant, 2001) at post-tutoring to assess the children's oral reading processing and reading comprehension processing.

#### *Rhyme Detection Task*

The Rhyme Detection task (Muter et al., 1997) assessed a child's rhyme awareness at pre- and post-tutoring (see Table 9). This task was one of six tasks in Muter et al.'s Phonological Abilities Test. To administer the task, I (a) named the target word and pointed to its corresponding picture and (b) named the three possible choices and pointed to their corresponding pictures. Then, the child selected the word that rhymed with or sounded like the target word. I provided the child with corrective feedback on the practice items and on the first four test items and scored the child's responses as correct or incorrect. The task included pictures that corresponded to the target word and the three choices to reduce the demands placed upon the child's memory. The task had one form,

Table 9

*Researcher Administered Measures, Scoring Procedures, and Assessed Early Literacy Components*

Measure	Scoring procedures	Assessed early literacy component
RD, BW, SWS	Composite	Phonological Awareness: Overall Phonological Awareness Composite <sup>a</sup>
RD	Muter et al., 1997	Phonological Awareness: Rhyme Awareness <sup>a</sup>
BW	Wagner et al., 1999	Phonological Awareness: Combined Syllable, Onset-Rime, and Phonemic Awareness <sup>a</sup>
SWS	Bourassa & Treiman, 2003	Phonological Awareness: Phonological Skeletal Structure Awareness <sup>a</sup>
SWS	DeFord, 2000	Phonological Awareness: Graphophonemic Awareness <sup>a</sup>
SWS	Composite	Orthographic Knowledge: Overall Orthographic Knowledge Composite <sup>a</sup>
SWS	DeFord, 2000	Orthographic Knowledge: Spelling Knowledge <sup>a</sup>
SWS	Researcher	Orthographic Knowledge: Orthographic Pattern Knowledge <sup>a</sup>
SWS	Bourassa & Treiman, 2003	Orthographic Knowledge: Orthographic Acceptability Knowledge <sup>a</sup>
GORT-4	Wiederholt & Bryant, 2001	Oral Reading Processing: GORT-4 Accuracy <sup>b</sup>

Table 9 continued

*Researcher Administered Measures, Scoring Procedures, and Assessed Early Literacy Components*

Measure	Scoring procedures	Assessed early literacy component
GORT-4	Researcher	Oral Reading Processing: Modified GORT-4 Accuracy <sup>b</sup>
GORT-4	Researcher	Oral Reading Processing: Overall Errors, Substitutions, Tolds, Omissions, Insertions <sup>b</sup>
GORT-4	Wiederholt & Bryant, 2001	Oral Reading Processing: GORT-4 Rate, GORT-4 Fluency <sup>b</sup>
GORT-4	Researcher	Oral Reading Processing: Self-Corrections, Repetitions <sup>b</sup>
GORT-4	Wiederholt & Bryant, 2001	Reading Comprehension Processing: GORT-4 Comprehension <sup>b</sup>
GORT-4	Wiederholt & Bryant, 2001	Oral Reading and Reading Comprehension Processing Composite <sup>b</sup>

*Note.* RD = Rhyme Detection; BW = Blending Words; SWS = Sentence Writing and Spelling; GORT-4 = Gray Oral Reading Tests-Fourth Edition.

<sup>a</sup> Refers to measures administered at pre-and post-tutoring. <sup>b</sup> Refers to measures administered at post-tutoring.

10 items, and three practice items (see Table 10 for a sample item). I reported the children's rhyme awareness performance at pre- and post-tutoring as proportions.

Muter et al. (1997) conducted an internal consistency analysis and a test-retest analysis to establish the reliability of their Rhyme Detection task. As displayed in Table 11, an internal consistency analysis produced a high-moderate coefficient and a test-retest analysis yielded a moderate coefficient.

Additionally, Muter et al. (1997) performed various analyses to establish the validity of their task. A criterion-related analysis produced a moderate concurrent coefficient (see Table 12). Additional criterion-related analyses consisted of conducting a series of regression analyses to determine how well the Rhyme Detection task administered to children at four-years old in the United Kingdom predicted their word reading, as measured by a standardized measure of word reading, at five-, six-, and seven-years old. These analyses determined that the Rhyme Detection task predicted word reading at a statistically significant level for five- and seven-year-old children but failed to predict word reading for six-year-old children at a statistically significant level. The construct validity analysis that correlated scores on the Rhyme Detection task with scores on another rhyming task produced a moderate convergent coefficient ( $r = .50$ ), based on 826 children in the United Kingdom between the ages of four and eight years old.

#### *Blending Words Task*

The five- and six-year-old version of the Blending Words task (Wagner et al., 1999) assessed a child's combined syllable, onset and rime, and phonemic awareness at pre- and post-tutoring (see Table 9). This task is one of eight tasks in Wagner et al.'s

Table 10

*Sample Items and Passage Excerpts from the Rhyme Detection Task, the Blending Words Task, the Sentence Writing and Spelling Task, and the GORT-4*

Task	Sample item and passage excerpt
RD	Target Word: boat    Picture Choices: foot, bike, coat
BW	num – ber                      n – ap                      m – oo – n
SWS	A man, a duck, and a dog were in a boat. The boat hit a rock and it filled up. Then the water made it sink.
GORT-4 (Story 2)	Our cat Mimi likes to sit on the roof. Mimi goes up the tall tree by the house. Then she jumps on the roof. She sits and looks at birds.
GORT-4	What does Mimi like best? A. the tree    B. the grass    C. <i>the roof</i> D. the bed
GORT-4 (Story 4)	It was time to get up and go to school. The children made their beds and dressed. One child said, “I can’t find my red shoes.” Mother said, “Then you’ll have to wear the brown ones instead.” The other child said, “I’ve lost my blue book.”
GORT-4	How do you think the family in this story felt? A. <i>hurried</i> B. sorry    C. happy    D. lucky

*Note.* RD = Rhyme Detection, BW = Blending Words, SWS = Sentence Writing and Spelling, GORT-4 = Gray Oral Reading Tests-Fourth Edition.

Table 11

*Internal Consistency and Test-retest Reliability Information for the Rhyme Detection Task and the Blending Words Task*

Task	Reliability type	Sample size	Age/Grade	Location	Coefficient	Procedure	Times
RD	Internal Consistency	N = 60	4 – 7 years	UK	.87	Cronbach's Alpha	-
BW	Internal Consistency	N = 155 <sup>a</sup>	6 years	US	.89	Cronbach's Alpha	-
RD	Test-retest	N = 35	5 years	UK	.80	-	2 Weeks Apart
BW	Test-retest	N = 32	5 – 7 years	US	.88	-	2 Weeks Apart

*Note.* RD = Rhyme Detection; BW = Blending Words; UK = United Kingdom; US = United States; - = Dash indicates the information was not applicable.

<sup>a</sup> Children were drawn from a normative sample (N = 1,656).



Table 12

*Criterion-related Validity Information for the Rhyme Detection Task and the Blending Words Task*

Task	Sample size	Age/Grade	Location	Correlating measure	Concurrent Coefficient
RD	N = 826	4 – 8 years	UK	BAS: Single Word Reading subtest	.51
BW	N = 444 <sup>a</sup>	5 – 7 years	US	TOWRE: Sight Word Efficiency subtest	.52
BW	N = 444 <sup>a</sup>	5 – 7 years	US	TOWRE: Phonetic Decoding Efficiency subtest	.48
BW	N = 444 <sup>a</sup>	5 – 7 years	US	TOWRE: Sight Word & Phonetic Decoding Efficiency Composite	.51
BW	N = 164	K - college	US	WRMT-R: Word Identification subtest	.59 <sup>b</sup>

*Note.* RD = Rhyme Detection; BW = Blending Words; UK = United Kingdom; US = United States; BAS = British Ability Scales test; TOWRE = Test of Word Reading Efficiency; WRMT-R = Woodcock Reading Mastery Tests-Revised.

<sup>a</sup>Children were drawn from a normative sample (N = 1,656). <sup>b</sup>Represents a partial correlation controlling for age.

Comprehensive Test of Phonological Processing. The child (a) listened to an audiotape of isolated sounds that included syllables, onset and rimes, or phonemes and (b) blended them to form words (see Table 10 for sample items). I provided corrective feedback on the practice items and on the first four test items and scored the child's responses as correct or incorrect. This task had one form, 20 items, and three practice items. As with rhyme awareness, I reported the children's combined syllable, onset and rime, and phonemic awareness at pre- and post-tutoring as proportions.

Wagner et al. (1999) conducted two analyses to establish the reliability of the Blending Words task. As shown in Table 11, an internal consistency analysis and a test-retest analysis both produced a high-moderate coefficient.

To establish the validity of the Blending Words task, Wagner et al. (1999) conducted several analyses. A criterion-related validity analysis produced moderate concurrent coefficients (see Table 12). Item analyses established content and construct validity. One item analysis produced an item discrimination coefficient of .50 and a median item difficulty percentage of 20%, based on children (N = 149), aged five years old. The second analysis yielded an item discrimination coefficient of .51 and a median item difficulty percentage of 30%, based on children (N = 155), aged six years old.

Wagner et al. drew these children from a normative sample of children in the United States (N = 1,656). Wagner et al. asserted that item discrimination coefficients of .35 or higher and item difficulty percentages distributed between 15% and 85% are acceptable for item selection. An age differentiation analysis that examined the relationship between the Blending Words task and age found that the means of the five- and six-year-old children increased with age. Because the ability to blend sounds is developmental in

nature, the relationship between the ability to blend sounds and age provides evidence for construct validity.

### *Sentence Writing and Spelling Task*

The Sentence Writing and Spelling task (DeFord, 2000) assessed the children's (a) phonological awareness, specifically phonological skeletal structure awareness and graphophonemic awareness and (b) orthographic knowledge, specifically spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge at pre- and post-tutoring (see Table 9). To administer this task, I read the entire dictation to the child. Then, I read the words in the dictation to the child one at a time, and the child recorded them. This task had three forms: Sentence A, Sentence B, and Sentence C. I administered Sentence B at pre-tutoring and Sentence C at post-tutoring (see Table 10 for a sample dictation). I used Bourassa and Treiman's (2003) scoring procedures to score the task for phonological skeletal structure awareness and orthographic acceptability knowledge. To score for graphophonemic awareness and spelling knowledge, I used DeFord's (2000) scoring procedures. I employed my scoring procedures to score for orthographic pattern knowledge (see Table 9).

I scored the Sentence Writing and Spelling task (DeFord, 2000) for phonological skeletal structure awareness. Phonological skeletal structure awareness is a child's capacity to hear phonemes in a spoken word, match the phonemes to either consonants or vowels, and record the consonant-vowel structure. Phonological skeletal structure awareness is the hearing of phonemes and recording of graphemes as either consonants or vowels.

In accordance with Bourassa and Treiman's (2003) scoring procedures, I scored

each child's whole word misspelling as correct or incorrect, depending upon whether it preserved the phonological skeletal structure of a spoken word without inserting or deleting a consonant(s) and/or a vowel(s). For example, if a child wrote *machac* for *magic*, I awarded the child a point because the child's misspelling preserved the phonological skeletal structure (i.e., consonant-vowel-consonant-vowel-consonant) in the word *magic*. The child substituted the *ch* consonants for the *g* consonant, substituted the *a* vowel for the *i* vowel; the child did not insert or delete any consonants or vowels (see Table 13 for this example). In contrast, if a child wrote *mhgisk* for *magic*, I did not award the child a point because the child's misspelling failed to preserve the phonological skeletal structure (i.e., consonant-vowel-consonant-vowel-consonant) in the word *magic*. The child substituted the consonant *h* for the vowel *a* and inserted the *s* consonant (see Table 13 for this example).

When scored for phonological skeletal structure awareness, the Sentence Writing and Spelling task (DeFord, 2000) contained 18 words. I reported the children's phonological skeletal structure awareness at pre- and post-tutoring as proportions because the number of words the children preserved or failed to preserve was a function of the number of words they misspelled. For example, a child who preserved the phonological skeletal structure of three out of three misspellings (1.00) demonstrated a greater understanding of phonological skeletal structure awareness than a child who preserved the phonological skeletal structure of three out of 10 misspellings (.10). A child who preserved the phonological skeletal structure of three out of three misspellings (1.00) demonstrated equal understanding as a child who preserved the phonological skeletal structure of 10 out of 10 misspellings (1.00). In the first example, although the children's

Table 13

*Phonological Skeletal Structure Awareness Scoring Examples*

---

Example 1: Phonological Skeletal Structure Preserved (Score: 1)

Child's Misspelling:	<u>m</u>	<u>a</u>	<u>ch</u>	<u>a</u>	<u>c</u>
Dictated Word:	m	a	g	i	c
Child's Phonological Skeletal Structure:	consonant	vowel	consonant	vowel	consonant
Word's Phonological Skeletal Structure:	consonant	vowel	consonant	vowel	consonant

Example 2: Phonological Skeletal Structure Not Preserved (Score: 0)

Child's Misspelling:	<u>m</u>	<i>h</i>	<u>g</u>	<u>i</u>	<u>s</u>	<u>k</u>
Dictated Word:	m	a	g	i	-	c
Child's Phonological Skeletal Structure:	consonant	<i>consonant</i>	consonant	vowel	<i>consonant</i>	consonant
Word's Phonological Skeletal Structure:	consonant	<i>vowel</i>	consonant	vowel		consonant

---

*Note.* Italic indicates an error in the phonological skeletal structure.

raw scores were equal, the children's phonological skeletal structure awareness differed. In the second example, although the children's raw scores differed, the children's phonological skeletal structure awareness was equal.

In addition to scoring the Sentence Writing and Spelling task (DeFord, 2000) for phonological skeletal structure awareness, I scored it for graphophonemic awareness. Graphophonemic awareness is a child's capacity to hear phonemes in a spoken word, match the phonemes to their corresponding graphemes, and record the graphemes. These graphemes may be exact phoneme matches or may be acceptable phoneme substitutions.

In accordance with DeFord's (2000) scoring procedures, I scored each child's individual phoneme-grapheme response as correct or incorrect, allowing for acceptable phoneme-grapheme substitutions. For example, if a child wrote *sinc* for *sink*, I awarded the child four points because the child accurately recorded four graphemes that matched the four phonemes in the dictated word *sink*. According to DeFord's scoring procedures, substitution of the *c* grapheme for the *k* grapheme is an acceptable substitution (see Table 14 for this example). If a child wrote *cimk* for *sink*, I awarded the child three points because the child accurately recorded three graphemes that matched three of the four phonemes in the dictated word *sink*. DeFord's scoring procedures did not consider the substitution of the *m* grapheme for the *n* grapheme an acceptable phoneme-grapheme substitution, whereas the procedures consider the substitution of the *c* grapheme for the *s* grapheme an acceptable substitution (see Table 14 for this example). When scored for graphophonemic awareness, the Sentence Writing and Spelling task contained 50 phonemes and their corresponding graphemes. I reported the children's graphophonemic awareness at pre- and post-tutoring as proportions.

Table 14

*Graphophonemic Awareness Scoring Examples*

## Example 1

---

Child's Misspelling:	<u>s</u>	<u>i</u>	<u>n</u>	<u>k</u>
Dictated Word:	s	i	n	k
Score:	4			

## Example 2

Child's Misspelling:	<u>c</u>	<u>i</u>	<u>m</u>	<u>c</u>
Dictated Word:	s	i	n	k
Score:	3			

---

*Note.* Italic indicates an incorrect phoneme-grapheme match.

The difference between phonological skeletal structure awareness and graphophonemic awareness was a subtle one, grounded in scoring differences rather than in children's processing differences. Both examined the children's capacity to hear phonemes in spoken words, match the phonemes to the corresponding graphemes, and record the corresponding graphemes. However, phonological skeletal structure awareness focused on whether the children recorded the consonant-vowel structure (i.e., phoneme-consonant or phoneme-vowel match) of spoken words, whereas graphophonemic awareness focused on whether the child recorded the corresponding graphemes for each phoneme (i.e., phoneme-grapheme match) of spoken words. When I administered the

Sentence Writing and Spelling task (DeFord, 2000), the child did not make the cognitive distinction between (a) a phoneme-consonant or a phoneme-vowel match or (b) a phoneme-grapheme match. When I scored the task for either phonological skeletal structure awareness or graphophonemic awareness, I made this distinction.

I also scored the Sentence Writing and Spelling task (DeFord, 2000) for spelling knowledge. Spelling knowledge is a child's capacity to write a word with correct spelling. I scored the children's spelling attempts as correct or incorrect according to DeFord's scoring procedures. When scored for spelling knowledge, this task contained 18 words. Analogous to graphophonemic awareness, I reported the children's spelling knowledge at pre- and post-tutoring as proportions.

I scored the Sentence Writing and Spelling task (DeFord, 2000) for orthographic pattern knowledge. Orthographic pattern knowledge is a child's capacity to record the letter patterns of an English word. I selected six words from Sentence B at pre-tutoring and six words from Sentence C at post-tutoring that contained the following orthographic letter patterns, respectively: (a) vowel digraphs (i.e., *oa* in *boat* and *ee* in *deep*), (b) double consonants (i.e., *ll* in *filled* and *small*), (c) inflectional endings (i.e., *ed* in *filled* and *kicked*), (d) consonant digraphs (i.e., *th* in *then*), (e) r-controlled vowels (i.e., *er* in *water* and *ir* in *dirt*), (f) silent letters (i.e., *e* in *made* and *hole*), and (g) consonant blends (i.e., *nk* in *sink* and *sm* in *small*). I employed my own scoring procedures by awarding a child's misspelling a point if it contained the appropriate orthographic pattern. For example, if a child wrote *waler* for *water*, I awarded the child a point because the child's misspelling contained the r-controlled vowel *er*. In contrast, if a child wrote *wutr* for *water*, I did not award the child a point because the child's misspelling failed to contain



the r-controlled vowel *er*. This task contained 6 words when scored for orthographic pattern knowledge. As with spelling knowledge, I reported the children's orthographic pattern knowledge at pre- and post-tutoring as proportions.

I scored the Sentence Writing and Spelling task (DeFord, 2000) for orthographic acceptability knowledge. Orthographic acceptability knowledge is a child's capacity to record the letter sequence(s) of an English word. I employed Bourassa and Treiman's (2003) scoring procedures by awarding a child's misspellings a point if it contained letter sequences that occur in English words. For example, if the child wrote *part* for *packed*, I awarded the child a point because the child's misspelling contained an acceptable letter sequence found in English words, such as *part*, *partial*, and *participate*. Conversely, if the child wrote *ptk* for *packed*, I did not award the child a point because the child's misspelling failed to contain an acceptable letter sequence. The letter sequence of *ptk* does not occur in English words. When scored for orthographic acceptability knowledge, the task contained 18 words. As with spelling knowledge and orthographic pattern knowledge, I reported the children's orthographic acceptability knowledge at pre- and post-tutoring as proportions.

To establish reliability, I conducted interrater reliability analyses. I conducted interrater reliability analyses because scoring the children's misspellings for graphophonemic awareness, phonological skeletal structure awareness, and orthographic acceptability knowledge involved interpretation. Conversely, the scoring of the children's misspellings for spelling knowledge and orthographic pattern knowledge did not involve interpretation because the children's spellings were either correct or incorrect and orthographic patterns were either present or not present.

The rater was a primary teacher with a master's degree in education. I trained this rater in (a) DeFord's (2000) scoring procedures for graphophonemic awareness and (b) Bourassa and Treiman's (2003) scoring procedures for phonological skeletal structure awareness and orthographic acceptability knowledge, using the writing samples from my pilot work. The rater and I separately scored the children's misspellings for phonological skeletal structure awareness, graphophonemic awareness, and orthographic acceptability knowledge. When scoring disputes occurred, the rater and I (a) rescored the misspellings and reached consensus or (b) rescored the misspellings and resolved any discrepancies via discussion. When we did not reach consensus, I selected my scoring over the rater's scoring. Interrater reliability analyses yielded (a) 100% percent agreement for phonological skeletal structure awareness and graphophonemic awareness at pre- and post-tutoring and (b) 96% and 98% agreement for orthographic acceptability knowledge at pre- and post-tutoring, respectively.

*Gray Oral Reading Tests-Fourth Edition*

The Gray Oral Reading Tests-Fourth Edition (GORT-4) (Wiederholt & Bryant, 2001) assessed the children's oral reading processing and reading comprehension processing at post-tutoring (see Table 9). After I read a brief scripted story introduction to the child, the child read the story out loud. While the child read, I recorded the child's oral reading. The stories did not contain pictures. After the child read the story, I read the text-related questions and their corresponding responses to the child while the child followed along in his/her book. Then, the child responded to these questions, and I recorded the child's responses. Each story possessed five questions. I audiotaped the child's oral reading and responses to the text-related questions for each story. I also

recorded the amount of time in seconds that each child required to read each story. Each child read story 1 and 2 and some children read story 3, 4, and 5 of the GORT-4's 14 stories. The GORT-4 had two forms: Form A and Form B. I administered Form A at post-tutoring (see Table 10 for passage excerpts and sample items). When a child reached their oral reading ceiling, as defined by the GORT-4 scoring procedures, I discontinued the child's oral reading.

To obtain the children's scores for GORT-4 accuracy, rate, fluency, and comprehension, I used the GORT-4 standard scoring procedures (see Table 9). For each GORT-4 story that the child read, I used the GORT-4's conversion table to form converted accuracy and rate scores from (a) the number of substitutions, tolds, omissions, insertions, self-corrections, and repetitions and (b) the number of seconds, respectively. These converted accuracy and rate scores ranged from 0 to 5. Then, I summed these converted accuracy and rate scores across all stories that the child read to form an accuracy total score and a rate total score for each child, respectively. I also summed a child's accuracy total score and rate total score to form a fluency total score for each child. To form a comprehension total score for each child, I summed the number of text-related questions that the child answered correctly across all stories. I used these accuracy, rate, fluency, and comprehension total scores in my statistical analyses of the post-tutoring data.

In addition to using the GORT-4 total scores, I also used the GORT-4 standard scores. Following the GORT-4's conversion guidelines, I converted the recommended and discontinued children's accuracy, rate, fluency, and comprehension total scores to their respective standard scores. The GORT-4 formed these standard scores based on a

distribution with a mean of 10 and a standard deviation of 3. I also followed the GORT-4 interpretation guidelines that provided (a) the rating categories of ‘very superior’, ‘superior’, ‘above average’, ‘average’, ‘below average’, ‘poor’, and ‘very poor’ and (b) corresponding normative sample distribution percentages, centered on the median of this GORT-4 rating category scale. In accordance with these interpretation guidelines, I placed the children’s accuracy, rate, fluency, and comprehension standard scores in these rating categories. Then, I compared the accuracy, rate, fluency, and comprehension standard score performance of the recommended children to the comparable performance of the discontinued children.

In addition, I also summed the recommended and discontinued children’s GORT-4 fluency and comprehension standard scores to form combined oral reading and reading comprehension processing composite standard scores (i.e., GORT-4 Oral Reading Quotient). These composite standard scores provided an indication of the children’s overall oral reading processing and reading comprehension processing. The GORT-4 formed these composite standard scores based on a distribution with a mean of 100 and a standard deviation of 15. In accordance with the interpretation guidelines that provided the rating category scale and the corresponding normative sample distribution percentages, centered on the median of this GORT-4 rating category scale, I placed the children’s composite standard scores in their corresponding rating categories. Then, I compared the composite standard score performance of the recommended children to the comparable performance of the discontinued children. I also compared the recommended children, as well as the discontinued children’s composite standard score performance to the normative sample’s composite standard score performance.

To obtain a child's scores for modified GORT-4 accuracy, overall errors, substitutions, tolds, omissions, insertions, self-corrections, and repetitions, I developed and used a modified version of the GORT-4's scoring procedures (see Table 9). By altering the GORT-4's definition of an error, I modified its scoring procedures. The GORT-4's scoring procedures considered substitutions, tolds, omissions, insertions, self-corrections, and repetitions as errors. I altered this definition by (a) considering substitutions, tolds, omissions, and insertions as errors and (b) self-corrections and repetitions as oral reading behaviors. I further modified the GORT-4's definition of an error by counting the name Mimi in story 2 as an error only one time rather than counting it as an error each time a child substituted another word or required a told for Mimi. I implemented these modifications so that the recording and scoring conventions for the GORT-4 paralleled Clay's (2002) recording and scoring conventions for Running Records, thereby making the findings of this study accessible to teachers familiar with Running Records.

The modified GORT-4 accuracy scores consisted of the number of errors that a child made while reading the GORT-4 stories. Errors included substitutions, tolds, insertions, and omissions. I formed the modified GORT-4 accuracy scores by summing a child's accuracy proportions on story 1, 2, 3, 4, and 5 from the GORT-4 and dividing by 5, the total number of stories. I awarded the children .00 for the stories that they did not read. For example, a child who read story 1 with .93 accuracy, story 2 with .80 accuracy, and story 3 with .75 accuracy did not read story 4 and 5 because the proportions on story 2 and 3 dropped below .90. Thus, I awarded this child .00 for stories 4 and 5. I computed this child's modified GORT-4 accuracy score of .50 by summing .93, .80, .75, .00, and

.00 and dividing by 5. I used the children's modified GORT-4 accuracy scores in my statistical analyses of the post-tutoring data.

The overall error scores included the number of substitutions, tolds, omissions, and insertions that a child made while reading story 1 and 2 of the GORT-4. To obtain detailed information regarding the children's errors, I disaggregated each child's overall errors score to obtain a separate count of the number substitutions, tolds, omissions, and insertions. The substitution scores included the number of times a child provided words or non-words that failed to match the words in story 1 and 2 of the GORT-4. The told scores included the number of times a child (a) appealed for help with words after making attempts to read the words and/or (b) balked at words without making attempts and failed to proceed in reading story 1 or 2 of the GORT-4. When a child made an attempt in the form of a substitution prior to being given a told, I counted this error as a told rather than a substitution. Thus, an error counted as a told remained separate from an error counted as a substitution. The omission scores included the number of times a child omitted a word in story 1 and 2 from the GORT-4, whereas the insertion scores included the number of times a child inserted a word in these two stories. The self-correction scores included the number of times a child fixed a substitution, omission, and/or insertion in story 1 and 2 from the GORT-4. The repetition scores included the number of times a child repeated a word, phrase, and/or sentence in story 1 and 2 from the GORT-4. I used these overall error, substitution, told, omission, insertion, self-correction, and repetition scores in the statistical analyses of the post-tutoring data.

To establish the reliability of the GORT-4, Wiederholt and Bryant (2001) conducted several analyses. As displayed in Table 15, the test-retest analyses yielded

Table 15

*Test-retest Reliability Information for the GORT-4*

GORT-4 subtest	Form A to A 2 weeks apart	Form B to B 2 weeks apart	Form A to B Extended time
Rate	.95	.91	.91
Accuracy	.92	.88	.91
Fluency	.93	.94	.91
Comprehension	.86	.85	.78

*Note.* Children (n = 49) drawn from a normative sample (N = 1,677) in the United States, aged 6-18 years old.

high coefficients for the rate, accuracy, and fluency subtests and high-moderate coefficients for the comprehension subtest. An internal consistency analysis yielded high coefficients for the rate, accuracy, fluency, and comprehension subtests (see Table 16). Similarly, a parallel split-half analysis produced high coefficients for the rate, accuracy, and fluency subtests and a moderate coefficient for the comprehension subtest (see Table 16). The interrater analysis that examined scoring agreement between two raters produced high coefficients ( $r = .96 - .99$ ) for rate, accuracy, fluency, and comprehension subtests.

Table 16

*Internal Consistency and Parallel Split-half Reliability Information for the GORT-4<sup>a</sup>*

GORT-4 subtest	Reliability type	Coefficient	Procedure
Rate	Internal Consistency	.91	Cronbach's Alpha
Rate	Parallel Split-half	.93	Correlated Form A and B
Accuracy	Internal Consistency	.90	Cronbach's Alpha
Accuracy	Parallel Split-half	.91	Correlated Form A and B
Fluency	Internal Consistency	.91	Cronbach's Alpha
Fluency	Parallel Split-half	.93	Correlated Form A and B
Comprehension	Internal Consistency	.95	Cronbach's Alpha
Comprehension	Parallel Split-half	.76	Correlated Form A and B

<sup>a</sup> Children (N = 87), aged six years old, were drawn from a normative sample in the United States (N = 1,677).

Wiederholt and Bryant (2001) conducted new validity studies on the GORT-4 and cited previously conducted validity studies from the Gray Oral Reading Tests-Third Edition (GORT-3) and the Gray Oral Reading Tests-Revised Edition (GORT-R). To justify their inclusion of previously conducted studies, Wiederholt and Bryant asserted that the GORT-3 and GORT-R were essentially the same as the GORT-4 in content and form, except for the story order and the addition of one new story and corresponding



comprehension questions to Form A and Form B.

The criterion-related validity analyses established (a) moderate concurrent coefficients corresponding to the rate and accuracy subtests and (b) low to moderate concurrent coefficients corresponding to the fluency and comprehension subtests. These analyses correlated the children's scores on these four subtests with their concurrent scores on various subtests of reading achievement measures (see Table 17).

Item analyses established content and construct validity by correlating an item score on the GORT-4 rate, accuracy, fluency, and comprehension subtests with the Oral Reading Quotient. These analyses produced acceptable item discrimination coefficients for the GORT-4's rate, accuracy, fluency, and comprehension subtests for Forms A and B. These item analyses also produced median item difficulty percentages for the GORT-4's comprehension subtest for Forms A and B (see Table 18). Wiederholt and Bryant (2001) asserted that item discrimination coefficients of .35 or higher and item difficulty percentages distributed between 15% and 85% are acceptable for item selection.

To establish construct validity further, a correlational analysis intercorrelated the standard scores of the GORT-4's rate, accuracy, fluency, and comprehension subtests. This analysis yielded correlations ranging between  $r = .39$  and  $r = .85$  that reached statistical significance ( $p < .01$ ) for Form A and B, based on the entire normative sample ( $N = 1,656$ ). These intercorrelations suggest that the GORT-4's subtests assess oral reading processing and reading comprehension processing, thereby providing evidence for construct validity. Wiederholt and Bryant (2001) did not report the coefficients for the correlations between rate and fluency and between accuracy and fluency because rate and accuracy scores are part of the fluency score.

Table 17

*Criterion-related Validity Information for the GORT-4*

Test	Sample size	Age/Grade	Correlating measure	Concurrent coefficients			
				R	A	F	C
GORT-4	N = 76	6 – 13 years	Gray Diagnostic Reading Tests	-	-	.41 - .72	.34 - .45
GORT-4	N = 76	6 - 13 years	Gray Silent Reading Tests	-	-	.56	.45
GORT-3	N = 201	K - 5 <sup>th</sup>	Test of Word Reading Efficiency	.75 - .82	.74 - .76	-	.44 - .50

*Note.* Children resided in the United States. R = Rate subtest; A = Accuracy subtest; F = Fluency subtest; C = Comprehension subtest.

Table 18

*Content and Construct Validity Information for the GORT-4*

GORT-4 Subtest	Item discrimination coefficient		Median item difficulty percentage	
	Form A	Form B	Form A	Form B
Rate	.76	.79	-	-
Accuracy	.75	.72	-	-
Fluency	.77	.79	-	-
Comprehension	.62	.56	41%	28%

*Note.* Children (n = 87) were drawn from a normative sample (N = 1,656) in the United States, aged six years old. - = Dash indicates the analysis was not performed.

Because oral reading processing and reading comprehension processing are developmental in nature, performance on the GORT-4's rate, accuracy, fluency, and comprehension subtests should increase with age and correlate strongly with it, thereby providing evidence for construct validity. An age differentiation analysis that inspected raw score mean performance on the GORT-4's rate, accuracy, fluency, and comprehension subtests for Forms A and B found that performance increased with age. This age differentiation analysis that also examined the relationship of performance on the GORT-4's subtests to age yielded moderate coefficients ( $r = .66 - .74$ ) for Forms A and B.

### *Summary*

*An Observation Survey of Early Literacy Achievement* (Clay, 2002) assessed the children's letter knowledge, knowledge of printed language conventions, reading vocabulary, writing vocabulary, graphophonemic awareness, and oral reading accuracy. The children's performance on this measure determined their entry into and exit from Reading Recovery tutoring. Reliability and validity analyses that yielded moderate to high coefficients provided evidence for the reliability and validity of this measure.

Based on the findings from my pilot work, I selected the following three phonological awareness tasks for this study: the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000). Based on criteria that I developed, these tasks were neither too easy nor too difficult for Reading Recovery children to complete prior to and following their tutoring, as evidenced by their mean scores.

The Rhyme Detection task (Muter et al., 1997) assessed the children's rhyme

awareness at pre- and post-tutoring. The Blending Words task (Wagner et al., 1999) assessed the children's combined syllable, onset and rime, and phonemic awareness at pre- and post-tutoring. When I administered the Rhyme Detection and Blending Words tasks, I used Muter et al.'s and Wagner et al.'s standard scoring procedures, respectively. The authors of these measures conducted various reliability and validity analyses and established moderate to strong reliability and validity. The Sentence Writing and Spelling task (DeFord, 2000) assessed the children's phonological skeletal structure awareness, graphophonemic awareness, spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge at pre- and post-tutoring. I used Bourassa and Treiman's (2003) scoring procedures to score the Sentence Writing and Spelling task for phonological skeletal structure awareness and orthographic acceptability. I employed DeFord's scoring procedures to score the Sentence Writing and Spelling task for spelling knowledge and graphophonemic awareness. I used my own scoring procedures to score the Sentence Writing and Spelling task for orthographic pattern knowledge. I conducted interrater reliability analyses and established strong reliability. The children's combined performance on the Rhyme Detection, Blending Words, and Sentence Writing and Spelling tasks also provided insight into their overall phonological awareness and overall orthographic knowledge in the form of composites.

The GORT-4 (Wiederholt & Bryant, 2001) assessed the children's post-tutoring oral reading processing that included oral reading accuracy, oral reading inaccuracy, and oral reading behaviors. Oral reading accuracy consisted of the GORT-4 accuracy scores and modified GORT-4 accuracy scores. Oral reading inaccuracy consisted of the overall error scores as well as the substitution, told, omission, and insertion scores. Oral reading

behaviors consisted of the GORT-4 rate scores, GORT-4 fluency scores, self-correction scores and repetition scores. The GORT-4 also assessed the children's post-tutoring reading comprehension processing that included the GORT-4 comprehension scores. The GORT-4 standard scoring procedures yielded the GORT-4 accuracy, rate, fluency, comprehension scores, whereas the modified scoring procedures yielded the modified accuracy scores and the overall error, substitution, told, omission, insertion, self-correction, and repetition scores. Wiederholt and Bryant conducted various analyses and established moderate to high reliability and validity coefficients corresponding to the GORT-4 accuracy, rate, and fluency subtests. However, they established low to moderate validity coefficients corresponding to GORT-4 comprehension subtest.

### *Procedures*

#### *Pre-tutoring Procedures*

During the first week of September, 2002, the classroom teachers completed alternate rankings to assist the Reading Recovery teachers in the identification of first-grade children to receive Reading Recovery tutoring. Each classroom teacher ranked the children in their classroom based upon their performance on a teacher-selected, early literacy activity. To complete this alternate ranking, the classroom teachers placed (a) the highest-achieving child in the first slot at the top of the ranking, (b) the lowest-achieving child in the last slot at the bottom of the ranking, (c) the second-highest achieving child in the second slot from the top of the ranking, (d) the second lowest-achieving child in the second slot from the bottom of the ranking, and (e) so forth until the teacher placed the two most average-achieving children in the two middle slots on the ranking.

Next, the Reading Recovery teachers administered *An Observation Survey of*

*Early Literacy Achievement* (Clay, 2002) to the bottom 20% of the children on the alternate rankings. For example, in a school with three first-grade classrooms with each classroom containing 30 children, the Reading Recovery teachers administered Clay's tasks to the six children in the three classrooms ranked as the lowest-achieving children by their classroom teachers. The Reading Recovery teachers individually administered these tasks to the children in approximately 30 minutes.

Then, the Reading Recovery teachers selected children for Reading Recovery tutoring. The Reading Recovery teachers selected children with the lowest scores on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) in accordance with the Reading Recovery standards and guidelines (RRCNA, 2001). The school district's teacher leader monitored (a) the classroom teachers' alternate ranking of children, (b) the Reading Recovery teachers' administration of *An Observation Survey of Early Literacy Achievement* and (c) the Reading Recovery teachers' selection of Reading Recovery children.

After the Reading Recovery teachers selected their Reading Recovery children, I distributed informed consent letters to the Reading Recovery children (Appendix B) and received 94% study participation; four children failed to return their informed consent letters. Then, I individually administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) to the Reading Recovery children during the first two weeks of their tutoring. Each child completed the pre-tutoring session in approximately 20 to 30 minutes, spending 5 to 7 minutes on the Rhyme Detection task, 5 to 7 minutes on the Blending Words task, and 10 to 15 minutes on the Sentence Writing and Spelling task.

*Post-tutoring Procedures*

The Reading Recovery teachers re-administered the tasks of *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to the Reading Recovery children upon their completion of tutoring. These teachers did not re-administer these tasks to their own children. Rather, other Reading Recovery teachers in the same school or the teacher leader re-administered them. The Reading Recovery teachers consulted with one another and determined whether the children (a) responded well to tutoring and met the criteria for successful completion or (b) did not respond well to tutoring and failed to meet the criteria for successful completion, as evidenced by their performance on Clay's tasks and the development of self-extending systems. The Reading Recovery teachers also consulted with the children's classroom teachers and determined whether the children were performing satisfactorily in the classroom during literacy instruction. If the children responded well to their tutoring and met the criteria for successful performance, their Reading Recovery teachers assigned them to the discontinued end-of-program status category. If the children did not respond well to their tutoring and failed to meet the criteria for successful performance, their Reading Recovery teachers assigned them to the recommended end-of-program status category. The majority of the Reading Recovery children completed their tutoring in approximately 20 weeks.

In February and March, 2003, I re-administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) to the children upon the completion of their tutoring. I also administered the GORT-4 (Wiederholt & Bryant, 2001). Each child completed the post-tutoring session in approximately 30 to 40 minutes, spending 5 minutes on the



Rhyme Detection task, 5 minutes on the Blending Words task, 10 minutes on the Sentence Writing and Spelling task, and 10 to 20 minutes on the GORT-4, depending upon the number of stories each child read and the rate at which the child read each story.

The teacher leader provided me with the children's pre- and post-tutoring scores on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and (b) the Reading Recovery teachers' decisions to classify children as recommended or discontinued. Furthermore, the teacher leader submitted to me the children's demographic information, including gender, date of birth, race/ethnicity, native language, and lunch program participation.

#### *Data Analyses*

In this section, I describe the various data analysis procedures I employed in this study. First, I explain the formation of the phonological awareness and orthographic knowledge composites and describe the data analysis procedure of repeated measure two-way analyses of variance. Next, I explain the formation of percentages and describe the data analysis procedure of chi-square tests of independence. Then, I describe the data analysis procedure of one-way analyses of variance.

#### *Composite Formation*

To form the overall phonological awareness composite at pre- and post-tutoring, I formed an overall phonological awareness composite proportion for each recommended and discontinued child at pre- and post-tutoring by using their pre- and post-tutoring scores on the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000). First, I scored the Rhyme Detection task for rhyme awareness by taking the number of items

each child answered correctly and dividing by the total number of items to create a proportion. Next, I scored the Sentence Writing and Spelling task for phonological skeletal structure awareness by taking the number of misspellings each child produced that preserved the phonological skeletal structure and dividing by the total number of misspellings each child produced to create a proportion. Then, I scored the Blending Words task for combined syllable, onset and rime, and phonemic awareness by taking the number of items each child answered correctly and dividing by the total number of items to create a proportion. Last, I scored the Sentence Writing and Spelling task for graphophonemic awareness by taking the number of phonemes-graphemes each child produced that matched the dictation and dividing by the total number of phonemes-graphemes in the dictation to create a proportion (see Table 19).

After I obtained the four proportions corresponding to (a) rhyme awareness; (b) phonological skeletal structure awareness; (c) combined syllable, onset and rime, and phonemic awareness; and (d) graphophonemic awareness for each child, I summed them to form an overall phonological awareness proportion for each child. Next, I divided each child's overall phonological awareness proportion by four, the total number of proportions produced by each child, to create an overall phonological awareness composite proportion. I performed this calculation at both pre- and post-tutoring (see Table 19). Then, using each child's pre- and post-tutoring proportion, I calculated a mean proportion across all children at both pre- and post-tutoring.

Similarly, to form the overall orthographic knowledge composite at both pre- and post-tutoring, I formed an overall orthographic knowledge composite proportion for each recommended and discontinued child at both pre- and post-tutoring by using their pre-

Table 19

*Formation of the Overall Phonological Awareness Composite at Pre- and Post-tutoring*

---

Task	Early literacy component	Scoring	Proportion
RD	Rhyme Awareness	# of Items Answered Correctly / Total # of Task Items	= P1
SWS	Phonological Skeletal Structure Awareness	# of Misspellings that Preserved PSS / Total # of Misspellings	= P2
BW	Syllable, Onset-Rime, Phonemic Awareness	# of Items Answered Correctly / Total # of Task Items	= P3
SWS	Graphophonemic Awareness	# of Matching Phoneme-Grapheme / Total # of Task Phoneme-Graphemes	= P4

---

Equations:

$$P1 + P2 + P3 + P4 = \text{Overall Phonological Awareness Proportion}$$

$$\text{Overall Phonological Awareness Proportion} / 4 \text{ (Total \# of Proportions)} = \text{Overall Phonological Awareness Composite}$$


---

*Note.* RD = Rhyme Detection; BW = Blending Words; SWS = Sentence Writing and Spelling; PSS = Phonological Skeletal Structure;

# = Number; P = Proportion.

and post-tutoring scores on the Sentence Writing and Spelling task (DeFord, 2000). First, I scored this task for spelling knowledge by taking the number of words with correct spelling and dividing by the total number of words on the task to create a proportion. Next, I scored the task for orthographic pattern knowledge by taking the number of misspellings that contained orthographic patterns and dividing by the total number of misspellings to create a proportion. Then, I scored this task for orthographic acceptability knowledge by taking the number of misspellings that contained acceptable English letter sequences and dividing by the total number of misspellings to create a proportion (see Table 20).

After I obtained the three proportions corresponding to (a) spelling knowledge, (b) orthographic pattern knowledge, and (c) orthographic acceptability knowledge for each child, I summed them to form an overall orthographic knowledge proportion for each child. Next, I divided each child's overall orthographic knowledge proportion by three, the total number of proportions produced by each child, to create an overall orthographic knowledge composite proportion. I performed this calculation both at pre- and post-tutoring for each child (see Table 20). Then, using each child's pre- and post-tutoring proportion, I calculated a mean proportion across all children at both pre- and post-tutoring.

I formed the phonological awareness composites at pre- and post-tutoring to obtain uniformity in score reporting among (a) rhyme awareness; (b) phonological skeletal structure awareness; (c) combined syllable, onset and rime, and phonemic awareness; and (d) graphophonemic awareness. Similarly, I formed the orthographic knowledge composites at pre- and post-tutoring to obtain uniformity in score reporting

Table 20

*Formation of the Overall Orthographic Knowledge Composite at Pre- and Post-tutoring*

---

Task	Early literacy component	Scoring	Proportion
SWS	Spelling Knowledge	# of Words with Correct Spelling / Total # of Task Words	= P1
SWS	Orthographic Pattern Knowledge	# of Misspellings with Orthographic Patterns / Total # of Misspellings	= P2
SWS	Orthographic Acceptability Knowledge	# of Misspellings with English Letter Sequences / Total # of Misspellings	= P3

---

Equations:

$$P1 + P2 + P3 = \text{Overall Orthographic Knowledge Proportion}$$

$$\text{Overall Orthographic Knowledge Proportion} / 3 (\text{Total \# of Proportions}) = \text{Overall Orthographic Knowledge Composite}$$


---

*Note.* SWS = Sentence Writing and Spelling; # = Number; P = Proportion.

among (a) spelling knowledge (b) orthographic pattern knowledge, and (c) orthographic acceptability knowledge. Because I reported the recommended and discontinued children's performance as proportions, each early literacy component contributed equally to the pre- and post-tutoring composites.

#### *Repeated Measure Two-way Analysis of Variance*

To analyze the data corresponding to the recommended and discontinued children's phonological awareness and orthographic knowledge composites at pre- and post-tutoring, I conducted two 2 (group) X 2 (time) repeated measure analyses of variance (ANOVAs). These two-way ANOVAs had a between-subjects factor or independent variable of group with two levels, recommended and discontinued. These two-way ANOVAS also had a within-subjects factor or independent variable of time with two levels, pre- and post-tutoring. First, I performed this ANOVA test on the dependent variable of overall phonological awareness. Then, I conducted this ANOVA test on the dependent variable of overall orthographic knowledge. Each two-way ANOVA determined whether there was (a) a statistically significant main effect for group, (b) a statistically significant main effect for time, and (c) a statistically significant interaction between group and time for each dependent variable. In an effort to minimize the likelihood of Type I error rates, I selected a priori significance level of .01 where the chance of a Type I error was equal to 1 out of 100 for each statistical test. To estimate the magnitude of the effects of the results, I calculated eta squared ( $\eta^2$ ). To interpret eta squared, I employed Cohen's criteria of .01, .059, and .138, signifying small, medium, and large effects, respectively.

### *Percentages*

I calculated percentages on the early literacy components that formed the phonological awareness and orthographic knowledge composites. To analyze the recommended and discontinued children's rhyme awareness and combined syllable, onset and rime, and phonemic awareness, I computed percentages that represented the number of items they answered correctly on the Rhyme Detection task (Muter et al., 1997) and the Blending Words task (Wagner et al., 1999), respectively. To analyze the children's phonological skeletal structure awareness and graphophonemic awareness, I computed percentages that represented the (a) number of misspellings that preserved the phonological skeletal structure and (b) the number of words that included beginning, middle, and ending phonemes at pre- and post-tutoring on DeFord's (2000) Sentence Writing and Spelling task, respectively. To analyze the children's spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge, I computed percentages that represented the (a) number of words with correct spelling, (b) the number of misspelling with orthographic patterns, and (c) the number of misspellings with acceptable English letter sequences at pre- and post-tutoring on the Sentence Writing and Spelling task, respectively.

### *Chi-square Test of Independence*

I performed several chi-square tests of independence on the components that formed the phonological awareness and the orthographic knowledge composites at pre- and post-tutoring. For each chi-square test the independent variable of group (i.e., recommended and discontinued) remained constant. However, the dependent variable changed for each test. First, I conducted several chi-square tests on the components that

formed the phonological awareness composite: (a) rhyme awareness; (b) phonological skeletal structure awareness; (c) combined syllable, onset and rime, and phonemic awareness; and (d) graphophonemic awareness at pre- and post-tutoring. Then, I conducted additional chi-square tests on the components that formulated the orthographic awareness composite: (a) spelling knowledge, (b) orthographic pattern knowledge, and (c) orthographic acceptability knowledge at pre- and post-tutoring. In an effort to minimize the likelihood of Type I error rates, I selected a priori significance level of .01 for each chi-square test. By setting an alpha level of .01, I established the chance of committing a Type I error to be 1 out of 100 for each test. I calculated a measure of association, phi ( $\Phi$ ), for each chi-square test to estimate the magnitude of the relationship between the independent and dependent variables. To interpret phi, I considered a weak relationship to range between .1 and .3, a medium relationship to range between .4 and .7, and a strong relationship to be at or above .8.

#### *One-way Analysis of Variance*

To analyze the data corresponding to the recommended and discontinued children's oral reading processing and reading comprehension processing at post-tutoring, I conducted several one-way analyses of variance (ANOVAs). Each one-way ANOVA had the same factor or an independent variable of group, recommended and discontinued, and a different dependent variable. I performed these ANOVA tests on the following oral reading dependent variables: GORT-4 accuracy, modified GORT-4 accuracy, overall errors, substitutions, tolds, omissions, insertions, GORT-4 rate, GORT-4 fluency, self-corrections, repetitions, and GORT-4 comprehension.

Each one-way ANOVA determined whether there was a statistically significant



difference between the recommended and discontinued groups for each dependent variable. Because I had selected several dependent variables, I conducted 12 one-way ANOVAs. Comparable to my decisions for the two-way ANOVAs, I selected a priori alpha level of .01 in an attempt to minimize the likelihood of committing Type I errors. Furthermore, I calculated eta squared for each ANOVA to estimate the magnitude of the difference between the two groups. Then, I employed Cohen's criteria to interpret eta squared.

### *Summary*

This study compared recommended and discontinued Reading Recovery children on (a) phonological awareness and orthographic knowledge prior to and following their completion of tutoring and (b) oral reading processing and reading comprehension processing following their completion of tutoring. The sample of Reading Recovery children consisted of 29 recommended children and 26 discontinued children. Fifteen Reading Recovery teachers and one teacher leader tutored these children on a daily basis for approximately 20 weeks. The Reading Recovery teachers administered *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to select children for tutoring. At the onset of the children's tutoring, I administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence and Writing task (DeFord, 2000) to assess the children's phonological awareness and orthographic knowledge. Upon their completion of tutoring, the Reading Recovery teachers re-administered Clay's measure to their Reading Recovery children. These teachers used the children's performance on this measure, along with the children's development of self-extending systems, to determine whether they (a) responded well to

tutoring and met the criteria for successful completion or (b) responded poorly to tutoring and failed to meet the criteria for successful completion. Subsequently, the teachers assigned the children to either the discontinued or recommended end-of-program status category. At this time, I re-administered the Rhyme Detection task, the Blending Words task, and the Sentence Writing and Spelling task to reassess the children's phonological awareness and orthographic knowledge. I also administered the GORT-4 (Wiederholt & Bryant, 2001) to assess the children's oral reading processing and reading comprehension processing. After I gathered the data, I calculated means and standard deviations, conducted repeated measure two-way analyses of variance, calculated percentages, conducted chi-square tests of independence, and conducted one-way analyses of variance. Chapter 4 reports the result of this study and provides a profile of an average-performing recommended child.

## Chapter 4

### *Results*

The current study compared recommended and discontinued Reading Recovery children's (a) phonological awareness and orthographic knowledge prior to and following their completion of tutoring and (b) oral reading processing and reading comprehension processing following their completion of tutoring. This study posed four research questions:

- (a) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on phonological awareness prior to and following their completion of tutoring?
- (b) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring?
- (c) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring?
- (d) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring?

Chapter 4 is organized into six sections. Sections 1 and 2 report the pre- and post-tutoring results corresponding to phonological awareness and orthographic knowledge, respectively. Sections 3 and 4 report the post-tutoring results corresponding to oral reading processing and reading comprehension processing, respectively. Section 5

presents the findings corresponding to the oral reading and reading comprehension processing composite. Section 6 provides a profile of an average-performing recommended Reading Recovery child.

### *Phonological Awareness*

This section answers the following research question: How do recommended Reading Recovery children compare to discontinued Reading Recovery children on phonological awareness prior to and following their completion of tutoring? First, this section reports the results corresponding to the recommended and discontinued children's overall phonological awareness in the form of the pre- and post-tutoring composites. Then, it reports the results corresponding to the following early literacy components that formulated the composites: (a) rhyme awareness; (b) phonological skeletal structure awareness; (c) combined syllable, onset and rime, and phonemic awareness; and (d) graphophonemic awareness.

*Overall phonological awareness.* Overall phonological awareness in the form of composites consisted of the recommended and discontinued children's pre- and post-tutoring scores on the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000). The Rhyme Detection task assessed the children's rhyme awareness. The Blending Words task measured the children's combined syllable, onset and rime, and phonemic awareness. The Sentence Writing and Spelling task assessed the children's phonological skeletal structure awareness and graphophonemic awareness (see Table 9). First, I report the means and standard deviations, followed by the results from a 2 (group) X 2 (time) repeated measure ANOVA with an alpha level of .01.

As displayed in Table 21, the recommended children demonstrated less overall phonological awareness than the discontinued children, collapsed across time. The recommended and discontinued children combined displayed gains in overall phonological awareness from pre- to post-tutoring, collapsed across group. When considering both group and time together, the recommended children showed less overall phonological awareness than the discontinued children at pre-tutoring, and the recommended children continued to show less overall phonological awareness than the discontinued children at post-tutoring. Both groups increased their overall phonological awareness from pre- to post-tutoring.

The ANOVA between-subjects test yielded a statistically significant main effect for group,  $F(1,53) = 18.91, p < .01, \eta^2 = .26$ , suggesting that the recommended children possessed less overall phonological awareness than the discontinued children, collapsed across time. The ANOVA within-subjects test revealed a statistically significant main effect for time,  $F(1,53) = 307.46, p < .01, \eta^2 = .85$ , indicating that the recommended and discontinued children combined made gains in overall phonological awareness from pre- to post-tutoring, collapsed across group. However, the ANOVA within-subjects test failed to yield a statistically significant interaction between group and time,  $F(1, 53) = .31, p = .58, \eta^2 = .01$  (see Table 22). Recommended children demonstrated less overall phonological awareness than discontinued children at a statistically significant level. Recommended and discontinued children combined improved in overall phonological awareness from pre- to post-tutoring at a statistically significant level. Recommended and discontinued improved in parallel in overall phonological awareness with recommended children performing below discontinued children at pre- and post- tutoring.

Table 21

*Means and Standard Deviations for Proportion Correct for Overall Phonological Awareness; Rhyme Awareness; Phonological Skeletal Structure Awareness; Combined Syllable, Onset and Rime, and Phonemic Awareness; and Graphophonemic Awareness for Group, Time, and Group X Time Interaction*

	OPA	RA	PSSA	CSORPA	GA
<b>Group</b>					
Recommended	.48 (.20)	.73 (.28)	.33 (.31)	.24 (.13)	.57 (.28)
Discontinued	.58 (.20)	.83 (.26)	.49 (.33)	.27 (.15)	.71 (.22)
<b>Time</b>					
Pre-tutoring	.36 (.13)	.69 (.31)	.13 (.15)	.16 (.10)	.43 (.16)
Post-tutoring	.69 (.12)	.87 (.21)	.68 (.19)	.34 (.11)	.85 (.13)

Table 21 continued

*Means and Standard Deviations for Proportion Correct for Overall Phonological Awareness; Rhyme Awareness; Phonological Skeletal Structure Awareness; Combined Syllable, Onset and Rime, and Phonemic Awareness; and Graphophonemic Awareness for Group, Time, and Group X Time Interaction*

	OPA	RA	PSSA	CSORPA	GA
<b>Group X Time</b>					
Pre-tutoring / Recommended	.32 (.11)	.62 (.29)	.07 (.10)	.15 (.10)	.35 (.15)
Pre-tutoring / Discontinued	.41 (.12)	.77 (.31)	.19 (.17)	.18 (.11)	.52 (.13)
Post-tutoring / Recommended	.63 (.12)	.84 (.24)	.60 (.21)	.32 (.11)	.80 (.15)
Post-tutoring / Discontinued	.75 (.08)	.90 (.17)	.78 (.11)	.37 (.12)	.90 (.06)

*Note.* OPA = Overall Phonological Awareness; RA = Rhyme Awareness; PSSA = Phonological Skeletal Structure Awareness; CSORPA = Combined Syllable, Onset and Rime, and Phonemic Awareness; GA = Graphophonemic Awareness.

Table 22

2 (Group) X 2 (Time) Repeated Measure ANOVA for Overall Phonological Awareness

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between subjects				
Group	1	18.91*	.26	.00
Error	53	(.02)		
Within subjects				
Time	1	307.46*	.85	.00
Group X Time	1	.31	.01	.58
Error	53	(.01)		

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .01$ .

*Rhyme awareness.* The Rhyme Detection task (Muter et al., 1997) measured the recommended and discontinued children's rhyme awareness at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's rhyme awareness at pre- and post-tutoring. Then, I report the results from chi-square tests of independence for group and rhyme awareness.

The percentages in Table 23 indicate that the recommended children detected



fewer rhymes than the discontinued children at pre- and post-tutoring. These findings raise the possibility that the recommended children possessed less rhyme awareness than the discontinued children at pre- and post-tutoring.

Table 23

*Recommended and Discontinued Children's Percentages for Rhyme Awareness at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	65%	77%
Post-tutoring	84%	90%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 550) = 9.02, p < .01, \Phi = .13$ ; Post-tutoring:  $\chi^2 (1, N = 550) = 4.59, p = .03, \Phi = .09$ .

Chi-square tests of independence yielded a statistically significant difference between the observed and expected frequencies of the recommended and discontinued children's rhyme awareness at pre-tutoring, but not at post-tutoring (see Table 23). Whereas the recommended and discontinued children differed at pre-tutoring, any observed differences at post-tutoring could be due to chance.

*Phonological skeletal structure awareness.* The Sentence Writing and Spelling task (DeFord, 2000) assessed the recommended and discontinued children's phonological

skeletal structure awareness at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's phonological skeletal structure awareness at pre- and post-tutoring. Next, I report the results from chi-square tests of independence for group and phonological skeletal structure awareness. Then, I discuss the means and standard deviations.

The percentages in Table 24 indicate that the recommended children heard fewer phonemes in spoken words, matched fewer phonemes to either consonants or vowels, and recorded fewer consonant-vowel structures than the discontinued children at pre- and post-tutoring, as evidenced by their misspellings. These findings could suggest that the recommended children possessed less phonological skeletal structure awareness than the discontinued children at pre- and post-tutoring.

Table 24

*Recommended and Discontinued Children's Percentages for Phonological Skeletal Structure Awareness at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	7%	18%
Post-tutoring	56%	78%

*Note.* Pre-tutoring:  $\chi^2(1, N = 828) = 24.06, p < .01, \Phi = .17$ ; Post-tutoring:  $\chi^2(1, N = 498) = 24.05, p < .01, \Phi = .22$ .

Chi-square tests of independence yielded statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's phonological skeletal structure awareness at pre- and post-tutoring (see Table 24). Both groups differed at pre- and post-tutoring.

The means and standard deviations displayed in Table 21 suggest that the recommended children demonstrated less phonological skeletal structure awareness than the discontinued children, collapsed across time. The combined recommended and discontinued children made gains in phonological skeletal structure awareness from pre- to post-tutoring, collapsed across group. When considering both group and time together, the recommended children showed less phonological skeletal structure awareness than the discontinued children at pre-tutoring, and the recommended children continued to show less phonological skeletal structure awareness than the discontinued children at post-tutoring. Even though both groups' scores increased from pre- to post-tutoring, the one consistent finding was that the two groups differed at a statistically significant level.

*Combined syllable, onset and rime, and phonemic awareness.* The Blending Words task (Wagner et al., 1999) measured the children's combined syllable, onset and rime, and phonemic awareness pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's combined syllable, onset and rime, and phonemic awareness at pre- and post-tutoring. Then, I report the results from chi-square tests of independence for group and combined syllable, onset and rime, and phonemic awareness.

The percentages in Table 25 indicate that the recommended children blended fewer syllables, onsets and rimes, or phonemes into words than the discontinued children

at pre- and post-tutoring. These findings could suggest that the recommended children possessed less combined syllable, onset and rime, and phonemic awareness than the discontinued children at pre- and post-tutoring.

Table 25

*Recommended and Discontinued Children's Percentages for Combined Syllable, Onset and Rime, and Phonemic Awareness at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	15%	18%
Post-tutoring	32%	37%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 1100) = .93, p = .34, \Phi = .03$ ; Post-tutoring:  $\chi^2 (1, N = 1100) = 3.79, p = .05, \Phi = .06$ .

Chi-square tests of independence failed to yield statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's combined syllable, onset and rime, and phonemic awareness at pre- and post-tutoring (see Table 25). All observed differences can be attributable to chance.

*Graphophonemic awareness.* The Sentence Writing and Spelling task (DeFord, 2000) assessed the recommended and discontinued children's graphophonemic awareness

with respect to beginning, middle, and ending phonemes at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's graphophonemic awareness with respect to beginning, middle, and ending phonemes. Then, I report the results from chi-square tests of independence for group and graphophonemic awareness with respect to beginning, middle, and ending phonemes.

The recommended children heard fewer beginning, middle, and ending phonemes in spoken words, matched fewer beginning, middle, and ending phonemes to the corresponding graphemes, and recorded fewer graphemes than the discontinued children at pre- and post-tutoring (see Tables 26-28). These findings raise the possibility that the recommended children possessed less graphophonemic awareness with respect to beginning, middle, and ending phonemes than the discontinued children at pre- and post-tutoring. These percentages included all the words in Sentence B at pre-tutoring and Sentence C at post-tutoring of DeFord's (2000) task, except for the word 'a'. The percentages corresponding to beginning and ending phonemes came from words with two, three, and four phonemes, whereas the percentages corresponding to middle sounds came from words with three and four phonemes.

Chi-square tests of independence yielded statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's graphophonemic awareness with respect to beginning phonemes at pre- and post-tutoring (see Table 26). Both groups differed at pre- and post-tutoring. Chi-square tests failed to reveal statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's graphophonemic

Table 26

*Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Beginning Sounds at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	40%	62%
Post-tutoring	71%	88%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 802) = 38.95, p < .01, \Phi = .22$ ; Post-tutoring,  $\chi^2 (1, N = 501) = 19.29, p < .01, \Phi = .20$ .

Table 27

*Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Middle Sounds at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	14%	16%
Post-tutoring	67%	73%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 605) = .44, p = .51, \Phi = .03$ ; Post-tutoring:  $\chi^2 (1, N = 475) = 1.80, p = .18, \Phi = .06$ .

Table 28

*Recommended and Discontinued Children's Percentages for Graphophonemic Awareness with Respect to Ending Sounds at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	41%	59%
Post-tutoring	83%	89%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 802) = 25.07, p < .01, \Phi = .18$ ; Post-tutoring:  $\chi^2 (1, N = 501) = 3.13, p = .08, \Phi = .08$ .

awareness with respect to middle phonemes at pre- and post-tutoring (see Table 27). Furthermore, chi-square tests yielded a statistically significant difference between the observed and expected frequencies of the recommended and discontinued children's graphophonemic awareness with respect to ending phonemes at pre-tutoring, but not at post-tutoring (see Table 28). Whereas both groups differed at pre-tutoring, any observed differences at post-tutoring could be due to chance. Overall, recommended and discontinued children differed on graphophonemic awareness with respect to beginning phonemes at pre- and post-tutoring and ending phonemes at pre-tutoring at statistically significant levels. All other differences can be attributable to chance.

*Summary.* A quantitative analysis yielded two notable findings regarding the comparison of recommended to discontinued children on the overall phonological

awareness composites. First, recommended children displayed less overall phonological awareness than discontinued children at a statistically significant level. Second, recommended and discontinued children combined improved in overall phonological awareness from pre- to post-tutoring at a statistically significant level. Additional quantitative analyses revealed several important findings regarding the early literacy components that formulated the overall phonological awareness composites. Recommended and discontinued children differed on (a) rhyme awareness at pre-tutoring, (b) phonological skeletal structure awareness at pre- and post-tutoring, (c) graphophonemic awareness with respect to beginning phonemes at pre- and post-tutoring, and (d) graphophonemic awareness with respect to ending phonemes at pre-tutoring at statistically significant levels.

### *Orthographic Knowledge*

This section answers the following research question: How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring? First, this section reports the results corresponding to the recommended and discontinued children's overall orthographic knowledge in the form of pre- and post-tutoring composites. Then, it reports the results corresponding to the following early literacy components that formulated the composites: (a) spelling knowledge, (b) orthographic pattern knowledge, and (c) orthographic acceptability knowledge.

*Overall orthographic knowledge.* Overall orthographic knowledge in the form of composites consisted of the recommended and discontinued children's pre- and post-tutoring scores on the Sentence Writing and Spelling task (DeFord, 2000). This task



assessed the children's spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge (see Table 9). After I discuss the means and standard deviations, I report the results from a 2 (group) X 2 (time) repeated measure ANOVA with an alpha level of .01.

As displayed in Table 29, the recommended children demonstrated less overall orthographic knowledge than the discontinued children, collapsed across time. The recommended and discontinued children combined displayed gains in overall orthographic knowledge from pre- to post-tutoring, collapsed across group. When considering both group and time together, the recommended children showed less overall orthographic knowledge than the discontinued children at pre-tutoring, and the recommended children continued to show less overall orthographic knowledge than the discontinued children at post-tutoring. Both groups increased their overall orthographic knowledge from pre- to post-tutoring.

The ANOVA between-subjects test yielded a statistically significant main effect for group,  $F(1,53) = 25.37, p < .01, \eta^2 = .32$ , suggesting that the recommended children possessed less overall orthographic knowledge than the discontinued children, collapsed across time. The ANOVA within-subjects test revealed a statistically significant main effect for time,  $F(1,53) = 230.07, p < .01, \eta^2 = .81$ , indicating that the recommended and discontinued children combined made gains in overall orthographic knowledge from pre- to post-tutoring, collapsed across group. However, the ANOVA within-subjects test failed to yield a statistically significant interaction effect between group and time,  $F(1, 53) = .46, p = .50, \eta^2 = .01$  (see Table 30). Recommended children demonstrated less overall orthographic knowledge than discontinued children at a statistically significant

Table 29

*Means and Standard Deviations for Proportion Correct for Overall Orthographic Knowledge, Spelling Knowledge, Orthographic Pattern Knowledge, and Orthographic Acceptability Knowledge for Group, Time, and Group X Time Interaction*

	OOK	SK	OPK	OAK
<b>Group</b>				
Recommended	.26 (.14)	.28 (.20)	.08 (.11)	.44 (.21)
Discontinued	.36 (.15)	.38 (.22)	.11 (.16)	.58 (.20)
<b>Time</b>				
Pre-tutoring	.19 (.09)	.15 (.08)	.01 (.04)	.41 (.21)
Post-tutoring	.43 (.11)	.49 (.16)	.17 (.15)	.61 (.16)

Table 29 continued

*Means and Standard Deviations for Proportion Correct for Overall Orthographic Knowledge, Spelling Knowledge, Orthographic Pattern Knowledge, and Orthographic Acceptability Knowledge for Group, Time, and Group X Time Interaction*

	OOK	SK	OPK	OAK
<b>Group X Time</b>				
Pre-tutoring / Recommended	.15 (.08)	.13 (.08)	.00 (.00)	.33 (.18)
Pre-tutoring / Discontinued	.23 (.09)	.18 (.08)	.02 (.05)	.50 (.22)
Post-tutoring / Recommended	.38 (.10)	.43 (.16)	.15 (.12)	.56 (.17)
Post-tutoring / Discontinued	.48 (.09)	.57 (.12)	.20 (.19)	.67 (.13)

*Note.* OOK = Overall Orthographic Knowledge; SK = Spelling Knowledge; OPK = Orthographic Pattern Knowledge; OAK = Orthographic Acceptability Knowledge.

Table 30

*2 (Group) X 2 (Time) Repeated Measure ANOVA for Overall Orthographic Knowledge*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between subjects				
Group	1	25.37*	.32	.00
Error	53	(.01)		
Within subjects				
Time	1	230.07*	.81	.00
Group X Time	1	.46	.01	.50
Error	53	(.01)		

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .01$ .

level. Recommended and discontinued children combined improved in overall orthographic knowledge from pre- to post-tutoring at a statistically significant level. Recommended and discontinued children improved in parallel in overall orthographic knowledge with recommended children performing below discontinued children at pre- and post-tutoring.

*Spelling knowledge.* The Sentence Writing and Spelling task (DeFord, 2000)

measured the recommended and discontinued children's spelling knowledge at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's spelling knowledge at pre- and post-tutoring. Then, I report the results from chi-square tests of independence for group and spelling knowledge.

The percentages in Table 31 indicate that the recommended children wrote fewer words with correct spelling than the discontinued children at pre- and post-tutoring. These findings raise the possibility that the recommended children possessed less spelling knowledge than the discontinued children at pre- and post-tutoring.

Table 31

*Recommended and Discontinued Children's Percentages for Spelling Knowledge at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	13%	18%
Post-tutoring	47%	57%

*Note.* Pre-tutoring:  $\chi^2(1, N = 990) = 4.99, p = .03, \Phi = .07$ , Post-tutoring:  $\chi^2(1, N = 990) = 10.50, p < .01, \Phi = .10$ .

Chi-square tests of independence yielded a statistically significant difference

between the observed and expected frequencies of the recommended and discontinued children's spelling knowledge at post-tutoring, but not at pre-tutoring (see Table 31). Whereas the recommended and discontinued children differed at post-tutoring, any observed differences at pre-tutoring can be attributable to chance.

*Orthographic pattern knowledge.* The Sentence Writing and Spelling task (DeFord, 2000) measured the recommended and discontinued children's orthographic pattern knowledge of vowel digraphs, double consonants, inflectional endings, consonant digraphs, r-controlled vowels, silent letters, and consonant blends at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's orthographic pattern knowledge at pre- and post-tutoring. Next, I report the results from chi-square tests of independence for group and orthographic pattern knowledge. Then, I report the percentages corresponding to the recommended and discontinued children's orthographic pattern knowledge of a specific double consonant, an inflectional ending, and a consonant digraph at post-tutoring and their corresponding chi-square tests.

As displayed in Table 32, the recommended children failed to include the orthographic patterns of vowel digraphs, double consonants, inflectional endings, consonant digraphs, r-controlled vowels, silent letters, and consonant blends in their misspellings at pre-tutoring. In comparison, the discontinued children included a small number of these same orthographic patterns in their misspellings at pre-tutoring. By post-tutoring, both groups slightly increased the number of orthographic patterns included in their misspellings with the recommended children including less orthographic patterns in their misspellings than the discontinued children (see Table 32). These findings could

Table 32

*Recommended and Discontinued Children's Percentages for Orthographic Pattern Knowledge at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	0%	2%
Post-tutoring	14%	22%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 394) = 4.21, p = .04, \Phi = .10$ , Post-tutoring:  $\chi^2 (1, N = 321) = 3.67, p = .06, \Phi = .11$ .

suggest that the recommended children possessed less orthographic pattern knowledge than the discontinued children.

Chi-square tests failed to yield statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's inclusion of the orthographic patterns in their misspellings at pre- and post-tutoring (see Table 32). All of the observed differences at pre- and post-tutoring can be attributable to chance.

Additional post-tutoring analyses revealed that the recommended children included fewer orthographic patterns, specifically the double consonant *ll* in their misspellings of *small*, the inflectional ending *ed* in their misspellings of *kicked*, and the consonant digraph *th* in their misspellings of *then*, than the discontinued children. These

findings could suggest that the recommended children possessed less knowledge of these specific orthographic patterns than the discontinued children at post-tutoring (see Table 33).

Table 33

*Recommended and Discontinued Children's Percentages for Specific Orthographic Pattern Knowledge at Post-tutoring*

Orthographic pattern	Recommended	Discontinued
Double Consonant ( <i>ll</i> in <i>small</i> ) <sup>a</sup>	30%	50%
Inflectional Ending ( <i>ed</i> in <i>kicked</i> ) <sup>b</sup>	3%	24%
Consonant Digraph ( <i>th</i> in <i>then</i> ) <sup>c</sup>	55%	79%

<sup>a</sup>Double Consonant:  $\chi^2(1, N = 39) = 1.50, p = .22, \Phi = .20$ . <sup>b</sup>Inflectional Ending:  $\chi^2(1, N = 54) = 5.03, p = .03, \Phi = .31$ . <sup>c</sup>Consonant Digraph:  $\chi^2(1, N = 34) = 2.00, p = .16, \Phi = .24$ .

Chi-square tests of independence failed to yield statistically significant differences between the observed and expected frequencies of the recommended and discontinued children's inclusion of the double consonant *ll* in their misspellings of *small*, the inflectional ending *ed* in their misspellings of *kicked*, and the consonant digraph *th* in their misspellings to *then* at post-tutoring (see Table 33). All of the observed



differences at post-tutoring can be attributable to chance.

*Orthographic acceptability knowledge.* The Sentence Writing and Spelling task (DeFord, 2000) assessed the recommended and discontinued children's orthographic acceptability knowledge at pre- and post-tutoring (see Table 9). First, I discuss the percentages corresponding to the recommended and discontinued children's orthographic acceptability knowledge. Then, I report the results from chi-square tests of independence for group and orthographic acceptability knowledge.

The percentages in Table 34 indicate that the recommended children included fewer acceptable English letter sequences in their misspellings than the discontinued children at pre- and post-tutoring. These findings could suggest that the recommended children possessed less orthographic acceptability knowledge than the discontinued children at pre- and post-tutoring.

Chi-square tests of independence revealed a statistically significant difference between the observed and expected frequencies of the recommended and discontinued children's orthographic acceptability knowledge at pre-tutoring, but not at post-tutoring (see Table 34). Whereas the recommended and discontinued children differed at pre-tutoring, any observed differences at post-tutoring could be due to chance.

*Summary.* A quantitative analysis yielded two notable findings regarding the comparison of recommended to discontinued children on the overall orthographic knowledge composites. First, recommended children displayed less overall orthographic knowledge than discontinued children at a statistically significant level. Second, recommended and discontinued children combined improved in overall orthographic knowledge from pre- to post-tutoring at a statistically significant level. Additional

quantitative analyses revealed two important findings regarding two of the early literacy components that formulated the overall orthographic knowledge composites.

Recommended and discontinued children differed on (a) spelling knowledge at post-tutoring and (b) orthographic acceptability knowledge at pre-tutoring at statistically significant levels.

Table 34

*Recommended and Discontinued Children's Percentages for Orthographic Acceptability Knowledge at Pre- and Post-tutoring*

Time	Recommended	Discontinued
Pre-tutoring	33%	49%
Post-tutoring	57%	67%

*Note.* Pre-tutoring:  $\chi^2 (1, N = 845) = 19.97, p < .01, \Phi = .15$ , Post-tutoring:  $\chi^2 (1, N = 845) = 4.94, p = .03, \Phi = .10$ .

*Oral Reading Processing*

This section answers the following research question: How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring? The GORT-4 (Wiederholt & Bryant, 2001) assessed the recommended and discontinued children's oral reading

processing at post-tutoring (see Table 9). First, this section reports the results corresponding to recommended and discontinued children's oral reading accuracy. Next, it reports the results corresponding to the recommended and discontinued children's oral reading inaccuracy. Then, this section reports the results corresponding to the recommended and discontinued children's oral reading behaviors.

*Oral reading accuracy.* Oral reading accuracy consisted of the recommended and discontinued children's accurate, oral reading of the GORT-4 stories at post-tutoring. The children's accurate, oral reading of these stories produced GORT-4 accuracy scores and modified GORT-4 accuracy scores, resulting from the employment of the GORT-4 standard scoring procedures and modified GORT-4 scoring procedures, respectively (see Table 9). First, I discuss the means and standard deviations and report the results from the ANOVAs with alpha levels of .01 for GORT-4 accuracy and modified GORT-4 accuracy scores. Then, I compare the recommended children's GORT-4 accuracy standard scores to the discontinued children's GORT-4 accuracy standard scores.

The means displayed in Table 35 indicated that the recommended children produced lower GORT-4 accuracy scores than the discontinued children. That is, the means suggest that the recommended children read with more substitutions, tells, omissions, insertions, self-corrections, and repetitions than the discontinued children. A one-way ANOVA conducted on the GORT-4 accuracy scores yielded a statistically significant difference between the two groups,  $F(1, 53) = 9.91, p < .01, \eta^2 = .16$  (see Table 36). Recommended children read the GORT-4 stories with less accuracy than the discontinued children.

Similarly, the means shown in Table 35 indicate that the recommended children

Table 35

*Means and Standard Deviations for GORT-4 Accuracy and Modified GORT-4 Accuracy at Post-tutoring*

Group	GORT-4 accuracy	Modified GORT-4 accuracy
	Range 0-70	Range 0.00-1.00
	<i>M (SD)</i>	<i>M (SD)</i>
Recommended (n = 29)	1.90 (2.01)	.37 (.20)
Discontinued (n = 26)	4.23 (3.39)	.60 (.25)
Total (N = 55)	3.00 (2.96)	.48 (.25)

Table 36

*One-way ANOVA for GORT-4 Accuracy at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	9.91*	.16	.00
Within Groups	53	(7.53)		

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .01$ .

produced lower modified GORT-4 accuracy scores than the discontinued children. That is, the means suggest that the recommended children read with more substitutions, tolds, insertions, and omissions than the discontinued children. A one-way ANOVA conducted on the modified GORT-4 accuracy scores revealed a statistically significant difference between both groups,  $F(1, 53) = 14.67, p < .01, \eta^2 = .22$  (see Table 37). Recommended children read the GORT-4 stories with less accuracy than the discontinued children.

Table 37

*One-way ANOVA for Modified GORT-4 Accuracy at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	14.67*	.22	.00
Within Groups	53	(.05)		

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .01$ .

I followed the GORT-4's conversion guidelines and converted the recommended and discontinued children's GORT-4 accuracy total scores to standard scores. I also followed the GORT-4 interpretation guidelines that provided (a) the rating categories of 'very superior', 'superior', 'above average', 'average', 'below average', 'poor', and 'very poor' and (b) the corresponding normative sample distribution percentages, centered on

the median of the GORT-4 rating category scale. In accordance with these interpretation guidelines, I placed the recommended and discontinued children's accuracy standard scores in these rating categories. Then, I compared the recommended children's accuracy to the discontinued children's accuracy. Table 38 displays the GORT-4 rating categories, the corresponding normative sample distribution percentages, and the recommended and discontinued children's accuracy standard score performance disaggregated into the GORT-4 rating categories.

The majority of the recommended children's accuracy standard scores placed in the 'below average' category with the remaining scores in the 'poor' category. In comparison, the near majority of the discontinued children's accuracy standard scores placed in the 'below average' category with the remaining scores in either the 'average' or the 'poor' categories (see Table 38). This comparison suggests that the recommended children performed below or equivalent to the discontinued children on GORT-4 accuracy.

*Oral reading inaccuracy.* Oral reading inaccuracy consisted of the recommended and discontinued children's inaccurate, oral reading of story 1 and 2 of the GORT-4 at post-tutoring. The children's inaccurate, oral reading of story 1 and 2 yielded (a) overall error scores that consisted of substitutions, tolds, omissions, and insertions and (b) separate substitution, told, omission, and insertion scores (see Table 9). First, I discuss means and standard deviations and the result from a one-way ANOVA with an alpha level of .01 for the overall error scores. Then, I discuss the means and standard deviations and the results from the one-way ANOVAs with an alpha level of .01 for each error score, including the substitution, told, omission, and insertion scores.

Table 38

*Percentages of Recommended and Discontinued Children's Accuracy, Rate, Fluency, Comprehension, and Composite Standard*

*Scores Placed in the GORT-4 Rating Category Scale*

GORT-4 Rating Category Scale	Normative Sample Percentages	Accuracy		Rate		Fluency		Comprehension		Composite	
		R	D	R	D	R	D	R	D	R	D
Very Superior	2.34%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Superior	6.87%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Above Average	16.12%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Average	49.51%	0%	27%	7%	27%	0%	19%	28%	27%	7%	15%
Below Average	16.12%	59%	42%	69%	58%	62%	46%	48%	38%	14%	23%
Poor	6.87%	41%	31%	24%	15%	34%	35%	24%	35%	59%	50%
Very Poor	2.34%	0%	0%	0%	0%	3%	0%	0%	0%	21%	12%

*Note.* Percentages do not add up to 100% due to round off error. R = Recommended; D = Discontinued.

The means shown in Table 39 indicate that the recommended children read story 1 and 2 from the GORT-4 with more overall errors than the discontinued children. That is, the means suggest that the recommended children read story 1 and 2 from the GORT-4 with more substitution, tolds, omissions, and insertions than the discontinued children. A one-way ANOVA conducted on the overall error scores revealed a statistically significant difference between the groups,  $F(1, 53) = 20.87, p < .01, \eta^2 = .28$  (see Table 40). Recommended children read story 1 and 2 of the GORT-4 with more overall errors than the discontinued children.

The means displayed in Table 39 indicate that the recommended children read story 1 and 2 from the GORT-4 with more substitutions, defined as a child's replacement of words with other words in the text while reading it, than the discontinued children. A one-way ANOVA conducted on the substitution scores revealed a statistically significant difference between the two groups,  $F(1, 53) = 7.21, p = .01, \eta^2 = .12$  (see Table 41). Recommended children read story 1 and 2 from the GORT-4 with more substitutions than the discontinued children. The statistical software produced a p value of .010 for this ANOVA test. Even though the p value is typically less than .01, I rendered this finding of  $p = .010$  statistically significant because of the conservative a priori alpha level of .01.

The means listed in Table 39 indicate that the recommended children read story 1 and 2 from the GORT-4 with more tolds, defined as the test examiner's pronouncing words in the text for a child while the child reads the text, than the discontinued children. A one-way ANOVA conducted on the told scores failed to yield a statistically significant difference between the groups,  $F(1, 51) = 4.80, p = .03, \eta^2 = .09$  (see Table 42).

The means shown in Table 39 indicate that the recommended children read story



Table 39

*Means and Standard Deviations for Overall Errors, Substitutions, Tolds, Omissions, and Insertions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Group	Overall errors	Substitutions	Tolds	Omissions	Insertions
	Range 0-31	Range 0-22	Range 0-23	Range 0-38	Range 0-2
	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Recommended (n = 29)	15.24 (6.54)	8.41 (6.34)	6.83 (5.78) <sup>a</sup>	1.48 (7.05)	.17 (.47)
Discontinued (n = 26)	8.34 (4.27)	4.62 (3.62)	4.04 (2.553) <sup>b</sup>	.08 (.27)	.04 (.20)
Total (N = 55)	11.98 (7.08)	6.62 (5.53)	5.57 (4.77) <sup>c</sup>	.82 (5.13)	.11 (.37)

Table 40

*One-way ANOVA for Overall Errors on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	20.87*	.28	.00
Within Groups	53	(31.23)		

*Note.* Values enclosed in parentheses represent mean square errors.

\* $p < .01$ .

Table 41

*One-way ANOVA for Substitutions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	7.21*	.12	.01
Within Groups	53	(27.42)		

*Note.* Values enclosed in parentheses represent mean square errors. The statistical software yielded an exact *p* value of .010.

\* $p = .01$ .

Table 42

*One-way ANOVA for Tolds on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	4.80	.09	.03
Within Groups	51	(21.24)		

*Note.* Values enclosed in parentheses represent mean square errors.

1 and 2 from the GORT-4 with more omissions, defined as a child's deletion of words from the text while reading it, than the discontinued children. A one-way ANOVA conducted on the omission scores failed to yield a statistically significant difference between the groups,  $F(1, 53) = 1.03$ ,  $p = .32$ ,  $\eta^2 = .02$  (see Table 43).

The means displayed in Table 39 indicate that the recommended children read story 1 and 2 from the GORT-4 with more insertions, defined as a child's addition of words to the text while reading it, than the discontinued children. A one-way ANOVA conducted on the insertion scores failed to yield a statistically significant difference between the groups,  $F(1, 53) = 1.84$ ,  $p = .18$ ,  $\eta^2 = .03$  (see Table 44).

*Oral reading behaviors.* Oral reading behaviors consisted of the recommended and discontinued children's rate and fluency corresponding to the GORT-4 stories at post-tutoring, as well as their self-corrections and repetitions corresponding to story 1 and 2 from the GORT-4 at post-tutoring (see Table 9). First, I discuss the means and standard

Table 43

*One-way ANOVA for Omissions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	1.03	.02	.32
Within Groups	53	(26.29)		

*Note.* Values enclosed in parentheses represent mean square errors.

Table 44

*One-way ANOVA for Insertions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	1.84	.03	.18
Within Groups	53	(.13)		

*Note.* Values enclosed in parentheses represent mean square errors.

deviations and report the results from the one-way ANOVAs with alpha levels of .01 for the GORT-4 rate and fluency scores. Next, I compare the recommended children's GORT-4 rate and fluency standard scores to the discontinued children's GORT-4 rate and

fluency standard scores. Then, I discuss the means and standard deviations and report the results from the one-way ANOVAs with alpha levels of .01 for the self-correction and repetition scores.

The means shown in Table 45 indicate that the recommended children read the GORT-4 stories at a slower rate, defined as the number of seconds a child takes to read a text, than the discontinued children. A one-way ANOVA conducted on the GORT-4 rate scores revealed a statistically significant difference between the groups,  $F(1, 53) = 7.19$ ,  $p = .01$ ,  $\eta^2 = .12$  (see Table 46). Recommended children read the GORT-4 stories at a slower rate than the discontinued children. The statistical software produced a p value of .010 for this ANOVA test. Even though the p value is typically less than .01, I rendered this finding of  $p = .010$  statistically significant because of the conservative a priori alpha level of .01.

Additionally, the means listed in Table 45 indicate that the recommended children read the GORT-4 stories with less fluency, defined as the number of seconds a child takes to read a text and the number of errors a child makes while reading a text, than the discontinued children. A one-way ANOVA conducted on the GORT-4 fluency scores revealed a statistically significant difference between the two groups,  $F(1, 53) = 10.48$ ,  $p < .01$ ,  $\eta^2 = .17$  (see Table 47). Recommended children read the GORT-4 stories with less fluency than the discontinued children.

In a manner analogous to oral reading accuracy, I followed the GORT-4's conversion guidelines and converted the recommended and discontinued children's GORT-4 rate total scores to standard scores and the GORT-4 fluency total scores to standard scores. I also followed the GORT-4 interpretation guidelines that provided the

Table 45

*Means and Standard Deviations for GORT-4 Rate, GORT-4 Fluency, Self-Corrections on Story 1 and 2 from the GORT-4, and Repetitions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Group	GORT-4 rate Range 0-70 <i>M (SD)</i>	GORT-4 fluency Range 0-140 <i>M (SD)</i>	Self- corrections Range 0-6 <i>M (SD)</i>	Repetitions Range 0-25 <i>M (SD)</i>
Recommended (n = 29)	1.76 (1.48)	3.66 (3.18)	1.55 (1.24)	9.52 (6.32)
Discontinued (n = 26)	3.15 (2.33)	7.38 (5.22)	2.12 (1.45)	7.85 (5.17)
Total (N = 55)	2.42 (2.03)	5.42 (4.63)	1.82 (1.36)	8.73 (5.81)

Table 46

*One-way ANOVA for GORT-4 Rate at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	7.19*	.12	.01
Within Groups	53	(3.71)		

*Note.* Values enclosed in parentheses represent mean square errors. The statistical software yielded an exact p value of .010.

\*p = .01.

Table 47

*One-way ANOVA for GORT-4 Fluency at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	10.48*	.17	.00
Within Groups	53	(18.20)		

*Note.* Values enclosed in parentheses represent mean square errors.

\*p < .01.

GORT-4 rating categories and corresponding normative sample distribution percentages, centered on the median of these GORT-4 rating category scale. Then, I compared the recommended children's rate and fluency to the discontinued children's rate and fluency.

The recommended and discontinued children's rate standard scores placed in the 'average', 'below average' and 'poor' categories with the majority of their scores in the 'below average' category. However, the recommended children had more scores in the 'below average' and 'poor' categories and fewer scores in the 'average' category than the discontinued children (see Table 38). This comparison indicates that the recommended children performed below or equivalent to the discontinued children on GORT-4 rate.

The majority of recommended children's fluency standard scores placed in the 'below average' category with the remaining scores placing in the 'poor' and 'very poor' categories. In comparison, the near majority of the discontinued children's fluency standard scores placed in the 'below average' category with the remaining scores placing in the 'average' and 'poor' categories (see Table 38). This comparison indicates that the recommended children performed below or equivalent to the discontinued children on GORT-4 fluency.

In addition to rate and fluency, the oral reading behaviors consisted of self-corrections and repetitions. The means shown in Table 45 indicate that the recommended children read story 1 and 2 from the GORT-4 with fewer self-corrections, defined as a child fixing previously committed errors in a text while reading it, than the discontinued children. A one-way ANOVA conducted on the self-correction scores failed to yield a statistically significant difference between the groups  $F(1, 53) = 2.41, p = .13, \eta^2 = .04$  (see Table 48). The means listed in Table 45 also indicate that the recommended children



Table 48

*One-way ANOVA for Self-corrections on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	2.41	.04	.13
Within Groups	53	(1.81)		

*Note.* Values enclosed in parentheses represent mean square errors.

read story 1 and 2 from the GORT-4 with more repetitions, defined as a child's saying words, phrases, and/or sentences in a text more than once while reading it, than the discontinued children. A one-way ANOVA conducted on the repetition scores failed to reveal a statistically significant difference between the groups  $F(1, 53) = 1.14, p = .29, \eta^2 = .02$  (see Table 49).

*Summary.* Quantitative analyses revealed several important findings regarding the comparison of recommended children to discontinued children on oral reading processing. In regards to oral reading accuracy, recommended children read the GORT-4 stories with less accuracy than discontinued children at statistically significant levels. A standard score comparison supported this finding by revealing that the recommended children performed below or equivalent to the discontinued children on GORT-4 accuracy. In regards to oral reading inaccuracy, recommended children read story 1 and 2 of the GORT-4 with more overall errors and substitutions than discontinued children at

Table 49

*One-way ANOVA for Repetitions on Story 1 and 2 from the GORT-4 at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	1.14	.02	.29
Within Groups	53	(33.67)		

*Note.* Values enclosed in parentheses represent mean square errors.

statistically significant levels. In regards to oral reading behaviors, recommended children read the GORT-4 stories at a slower rate and with less fluency than discontinued children at statistically significant levels. A standard score comparison supported these findings by indicating that the recommended children performed below or equivalent to the discontinued children on GORT-4 rate and fluency.

#### *Reading Comprehension Processing*

This section answers the following research question: How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring? The GORT-4 (Wiederholt & Bryant, 2001) assessed the recommended and discontinued children's reading comprehension processing at post-tutoring (see Table 9). This section reports the results corresponding to the recommended and discontinued children's reading comprehension processing.

As seen in Table 50, the recommended children correctly answered approximately the same number of text-related questions as the discontinued children with the recommended children performing slightly below the discontinued children. A one-way ANOVA conducted on the GORT-4 comprehension scores failed to reveal a statistically significant difference between the groups,  $F(1, 53) = .10, p = .76, \eta^2 = .00$  (see Table 51).

I followed the GORT-4's conversion guidelines and converted the recommended and discontinued children's GORT-4 comprehension total scores to standard scores. I also followed the GORT-4's interpretation guidelines that provided the GORT-4 rating categories and corresponding normative sample distribution percentages, centered on the median of the GORT-4 rating category scale. Then, I compared the recommended children's comprehension to the discontinued children's comprehension.

The near majority of the recommended and discontinued children's comprehension standard scores placed in the 'below average' category. Their remaining scores placed in either the 'average' or 'poor' categories with virtually the same number of scores placing in the 'average' category (see Table 38). This comparison indicates that the recommended children performed comparably to the discontinued children on comprehension.

*Summary.* A quantitative analysis revealed an important finding regarding the comparison of recommended children to discontinued children on reading comprehension processing. Recommended and discontinued children comprehended the GORT-4 stories equivalently. A standard score comparison supported this finding by revealing that recommended and discontinued children comprehended the GORT-4 stories comparably.

Table 50

*Means and Standard Deviations for GORT-4 Comprehension at Post-tutoring*

Group	Comprehension (Range 0-70) <i>M (SD)</i>
Recommended (n = 29)	5.34 (4.57)
Discontinued (n = 26)	5.73 (4.67)
Total (N = 55)	5.53 (4.58)

Table 51

*One-way ANOVA for GORT-4 Comprehension at Post-tutoring*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
Between Groups	1	.10	.00	.76
Within Groups	53	(21.31)		

*Note.* Values enclosed in parentheses represent mean square errors.

*Oral Reading and Reading Comprehension Processing Composite*

This section reports the results corresponding to the recommended and discontinued children's combined oral reading and reading comprehension processing in the form of a composite (see Table 9). The recommended and discontinued children's combined GORT-4 fluency (i.e., combined rate and accuracy) and GORT-4 comprehension performance at post-tutoring provided information on their combined oral reading and reading comprehension processing.

I followed the GORT-4's conversion guidelines and converted the recommended and discontinued children's GORT-4 summed fluency and comprehension standard scores into composite standard scores. I also followed the GORT-4 interpretation guidelines that provided the rating categories and corresponding normative sample distribution percentages, centered on the median of the GORT-4 rating category scale. In accordance with these interpretation guidelines, I placed the recommended and discontinued children's composite standard scores in their respective rating categories. Then, I compared the recommended children's composite standard scores to the discontinued children's composite standard scores. Furthermore, I compared the recommended children's, as well as the discontinued children's, composite standard scores to the composite standard scores of the children in the normative sample distribution.

The recommended and discontinued children's composite standard scores placed in the 'average', 'below average', 'poor', and 'very poor' categories with the majority of their scores in the 'poor' category. Yet the recommended children had more scores in the 'very poor' category and fewer scores in the 'below average' and 'average' categories

than the discontinued children (see Table 38). This comparison suggests that the recommended children performed below or equivalent to the discontinued children on combined oral reading and reading comprehension processing. In addition, the recommended and discontinued children's composite standard score distribution peaked in the 'poor' category (see Table 38). This finding indicates that the vast majority of the recommended and discontinued children performed poorly or very poorly on combined oral reading and reading comprehension processing compared to the normative sample distribution.

*Summary.* A composite standard score comparison revealed that the recommended children performed below or equivalent to the discontinued children on combined oral reading and reading comprehension processing in the form of a composite. An additional standard score composite comparison revealed that the vast majority of the recommended children, as well as the discontinued children, performed poorly or very poorly on combined oral reading and reading comprehension processing relative to the children in the normative sample.

#### *A Profile of an Average-performing Recommended Reading Recovery Child*

This section presents a profile of an average-performing recommended Reading Recovery child. I developed this profile by using the recommended children's mean scores corresponding to their (a) phonological awareness and orthographic knowledge mean scores at pre- and post-tutoring and (b) oral reading processing and reading comprehension processing at post-tutoring. I compared this profile of the average-performing recommended child to the profile of the average-performing discontinued child. This section also links this profile to early literacy development.

*Phonological Awareness*

In this investigation, a child revealed his/her overall phonological awareness through his/her rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness. A child's composite performance on the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000), scored for phonological skeletal structure awareness and graphophonemic awareness, revealed his/her overall phonological awareness. The average-performing recommended child demonstrated less overall phonological awareness than the average-performing discontinued child at pre- and post-tutoring. The average-performing recommended child improved his/her ability to (a) detect rhymes, (b) preserve the phonological skeletal structures in his/her misspellings, (c) blend isolated sound segments into recognizable words, and (d) hear, match, and write phonemes and their corresponding graphemes from pre- to post-tutoring (see Figure 2). This improvement may foster his/her ability to (a) read unfamiliar words by decoding and/or analogizing and (b) write unfamiliar words by segmenting and/or analogizing. Interestingly, parallel improvement characterized the pre- to post-tutoring gains of both the average-performing recommended and discontinued child (see Figure 2).

A child who detected rhyme possessed rhyme awareness, as evidenced by his/her performance on Muter et al.'s (1997) Rhyme Detection task. This task required a child to select the word and its corresponding picture that rhymed with the target word and its corresponding picture. The average-performing recommended child demonstrated less rhyme awareness than the average-performing discontinued child at pre- and post-

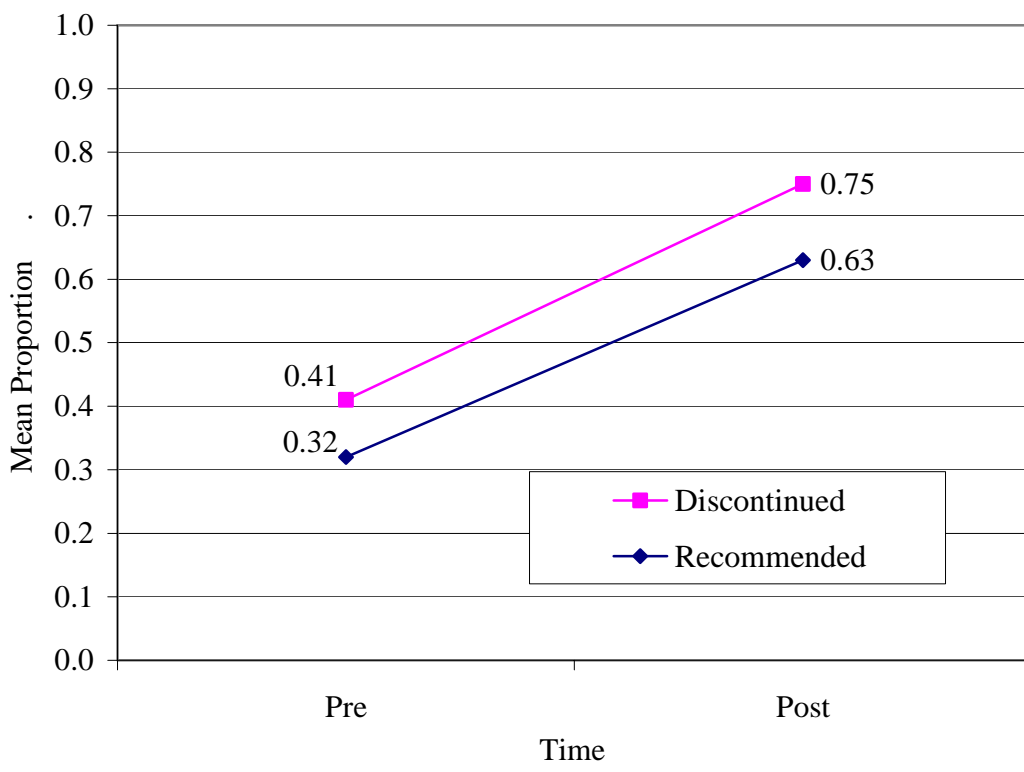


Figure 2. Overall Phonological Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

tutoring. The average-performing recommended child improved his/her ability to detect rhyme from pre- to post-tutoring (see Figure 3). This improvement may advance his/her ability to read and write unfamiliar words by analogizing. As with overall phonological awareness, parallel improvement marked the pre- to post-tutoring gains of both the average-performing recommended and discontinued child (see Figure 3).

A child who heard the phonemes in dictated words, matched these phonemes to consonant or vowel graphemes, and wrote the consonant-vowel structures in his/her misspellings possessed phonological skeletal structure awareness. Likewise, a child who heard the beginning, middle, and ending phonemes in dictated words, matched these



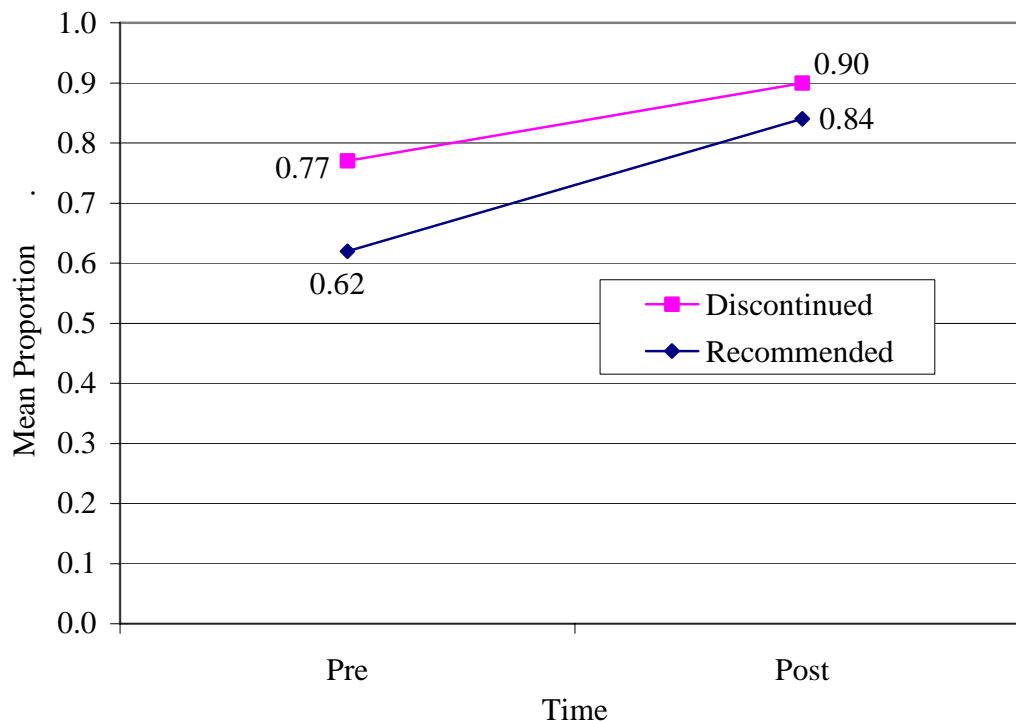


Figure 3. Rhyme Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

phonemes to their corresponding graphemes, and wrote these corresponding graphemes possessed graphophonemic awareness. DeFord's (2000) Sentence Writing and Spelling task assessed phonological skeletal structure awareness and graphophonemic awareness by requiring a child to listen to the words in a dictated story and write them. Bourassa and Treiman's (2003) and DeFord's scoring procedures determined a child's phonological skeletal structure awareness and graphophonemic awareness, respectively. The average-performing recommended child demonstrated less phonological skeletal structure awareness and graphophonemic awareness than the average-performing discontinued child at pre- and post-tutoring. The average-performing recommended child substantially

improved his/her phonological skeletal structure awareness and graphophonemic awareness from pre- to post-tutoring (see Figure 4 and 5). This improvement may promote his/her ability to segment unfamiliar words and write them. Both children improved in parallel from pre- to post-tutoring on phonological skeletal structure awareness and graphophonemic awareness (see Figure 4 and 5).

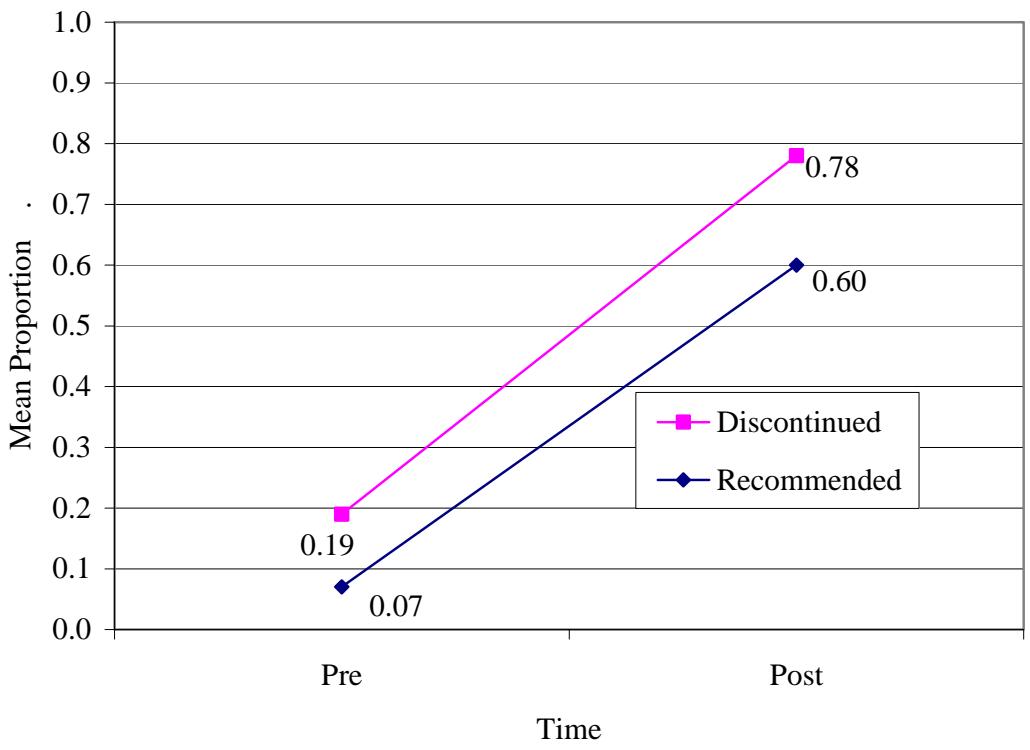


Figure 4. Phonological Skeletal Structure Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

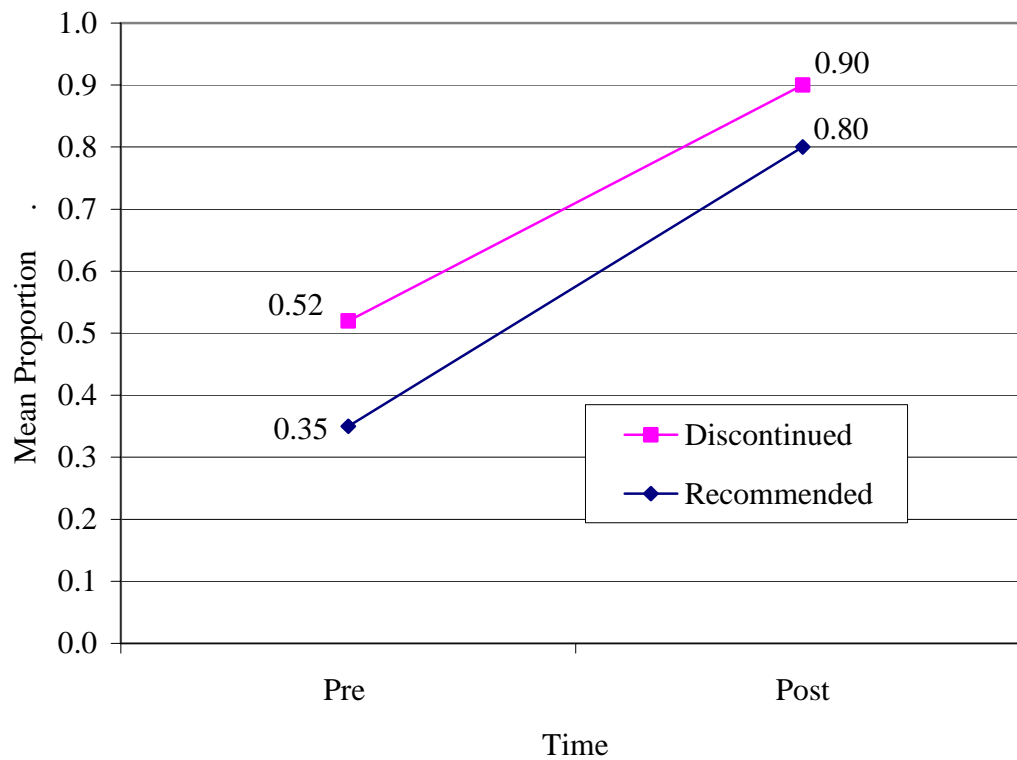


Figure 5. Graphophonemic Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

A child who blended isolated syllables, onset and rimes, or phonemes into recognizable words possessed combined syllable, onset and rime, and phonemic awareness, as evidenced by his/her performance on Wagner et al.'s (1999) Blending Words task. This task required a child to listen to isolated syllables, onset and rimes, or phonemes and blend them. The average-performing recommended child demonstrated less combined syllable, onset and rime, and phonemic awareness than the average-performing discontinued child at pre- and post-tutoring. Also, the average-performing recommended child minimally improved his/her combined syllable, onset and rime, and phonemic awareness from pre- to post-tutoring (see Figure 6). The small improvement may advance his/her ability to read unfamiliar words by decoding and/or analogizing.

However, because the improvement was small, this advancement is unlikely. Parallel improvement characterized the small gains of the average-performing recommended and discontinued child from pre- to post-tutoring (see Figure 6).

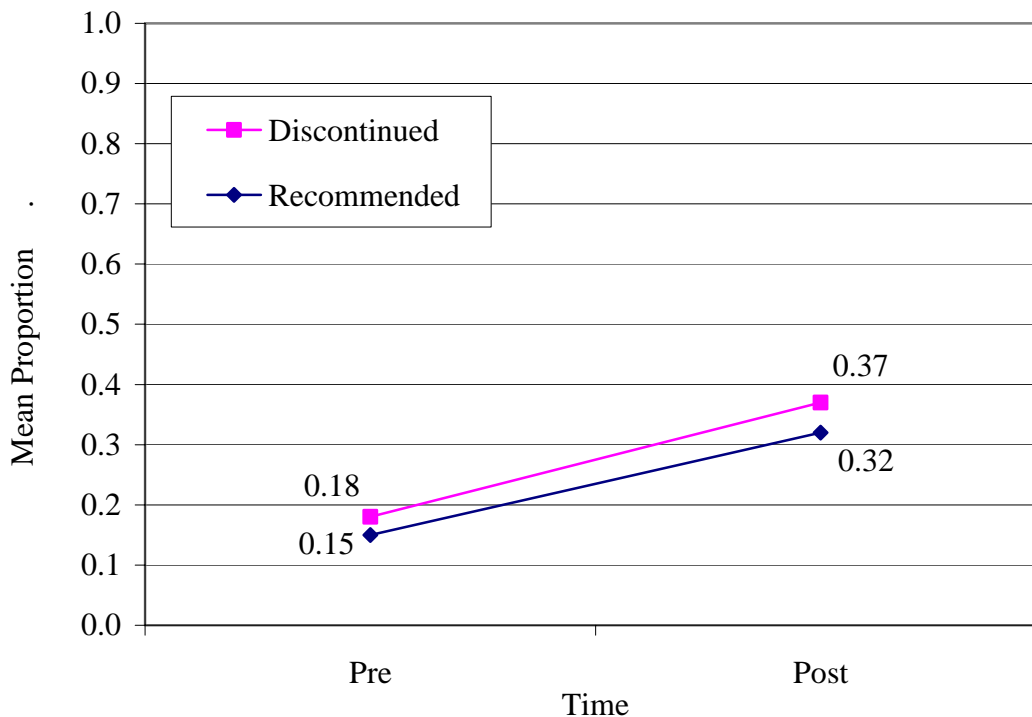


Figure 6. Combined Syllable, Onset and Rime, and Phonemic Awareness of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

### *Orthographic Knowledge*

In this study, a child revealed his/her overall orthographic knowledge by his/her spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge. A child's composite performance on the Sentence Writing and Spelling task (DeFord, 2000), scored for spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge, revealed his/her overall orthographic knowledge.

An average-performing recommended child demonstrated less overall orthographic knowledge than an average-performing discontinued child at pre- and post-tutoring. An average-performing recommended child improved his/her ability to write words with correct spelling, accurate letter patterns, and acceptable English letter sequences from pre- to post-tutoring (see Figure 7). This improvement may facilitate efficient and automatic perception of familiar and unfamiliar words during the reading of texts and/or the generation of words during the writing of texts. However, because an average-performing recommended child made a relatively small gain in overall orthographic knowledge from pre- to post-tutoring, s/he may experience difficulty in reading and writing familiar and unfamiliar words. Interestingly, parallel improvement marked the pre- to post-tutoring gains of both the average-performing recommended and discontinued child (see Figure 7).

A child who wrote words with correct spelling possessed spelling knowledge, as evidenced by his/her performance on DeFord's (2000) Sentence Writing and Spelling task, scored with DeFord's procedures. This task required a child to listen to the words in a dictated story and write them with correct spelling. An average-performing recommended child demonstrated less spelling knowledge than an average-performing discontinued child at pre- and post-tutoring. An average-performing recommended child improved his/her spelling knowledge from pre- to post-tutoring (see Figure 8). This improvement may facilitate his/her efficient and automatic perception of familiar and unfamiliar words during the reading of texts and/or generation of words during the writing of texts. Parallel improvement characterized the pre- to post-tutoring progress of both the average-performing recommended and discontinued child (see Figure 8).

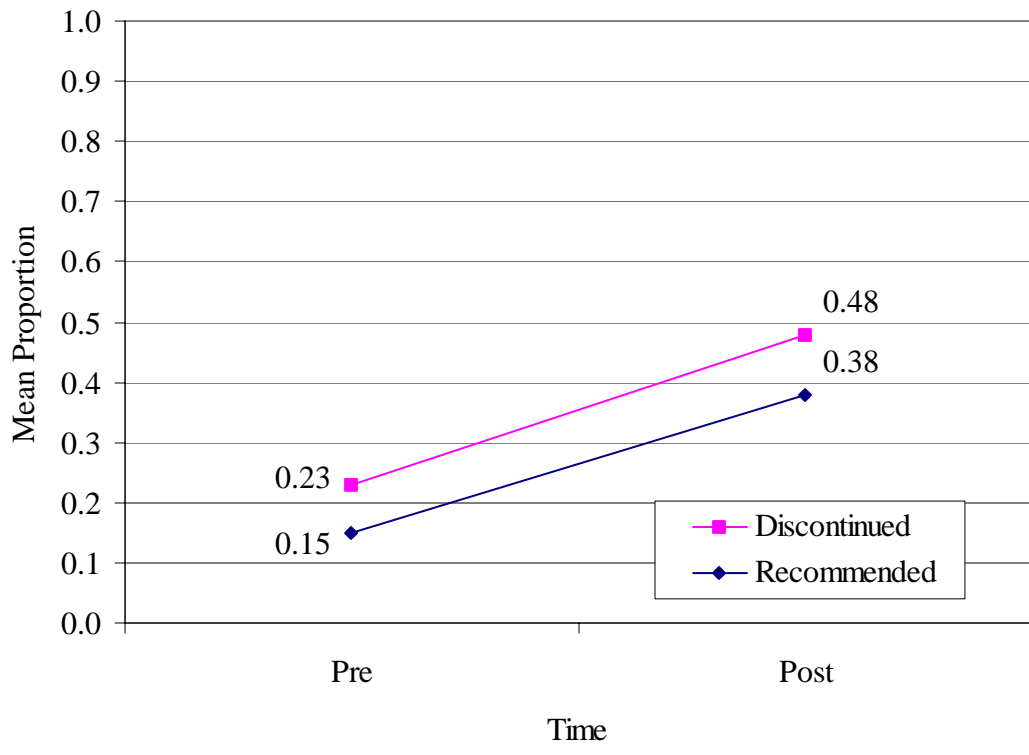


Figure 7. Overall Orthographic Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

A child who wrote words with their correct letter patterns possessed orthographic pattern knowledge. The Sentence Writing and Spelling task (DeFord, 2000) assessed orthographic pattern knowledge by requiring a child to listen to the words in a dictated story and write them with their correct letter patterns. These letter patterns included vowel digraphs, double consonants, inflectional endings, consonant digraphs, r-controlled vowels, silent letters, and consonant blends. I developed and used my own scoring procedures to determine a child’s orthographic pattern knowledge. An average-performing recommended child (a) demonstrated approximately the same level of

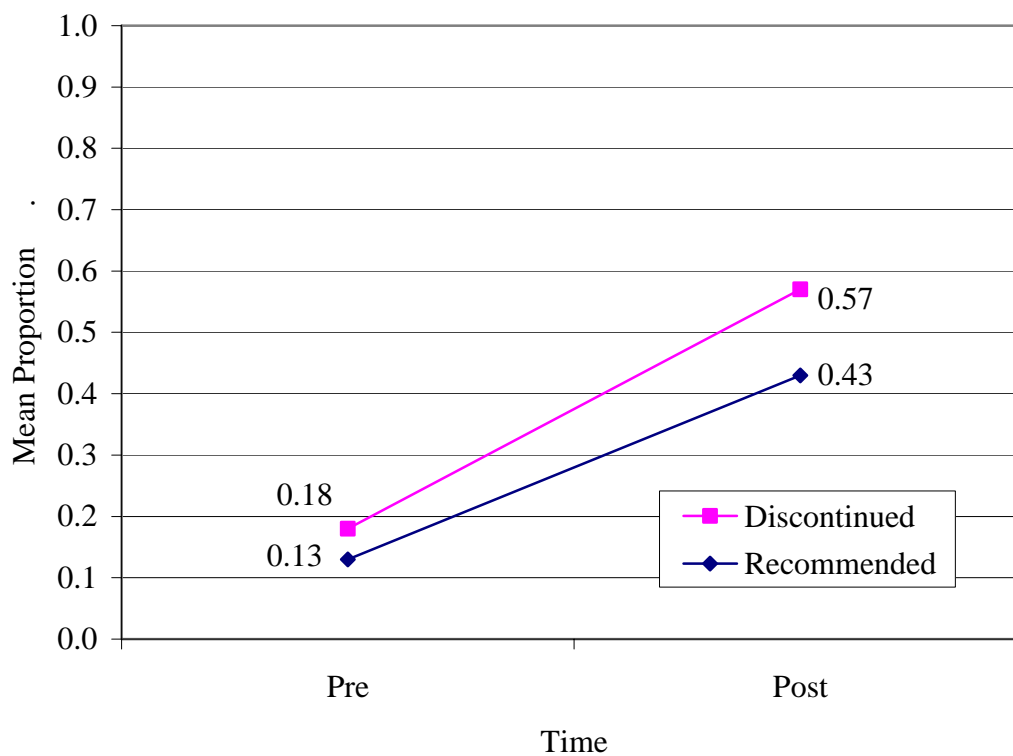


Figure 8. Spelling Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

orthographic pattern knowledge as an average-performing discontinued child at pre- and post-tutoring and (b) made virtually no progress in his/her orthographic pattern knowledge from pre- to post-tutoring (see Figure 9). An average-performing recommended child's marginal improvement in pattern knowledge may hinder his/her perception of words during the reading of texts and/or generation of words during the writing of texts, especially words that contain vowel digraphs, double consonants, inflectional endings, consonant digraphs, r-controlled vowels, silent letters, and consonant blends. Parallel improvement marked the minimal pre- to post-tutoring progress of the average-performing recommended and discontinued child (see Figure 9).

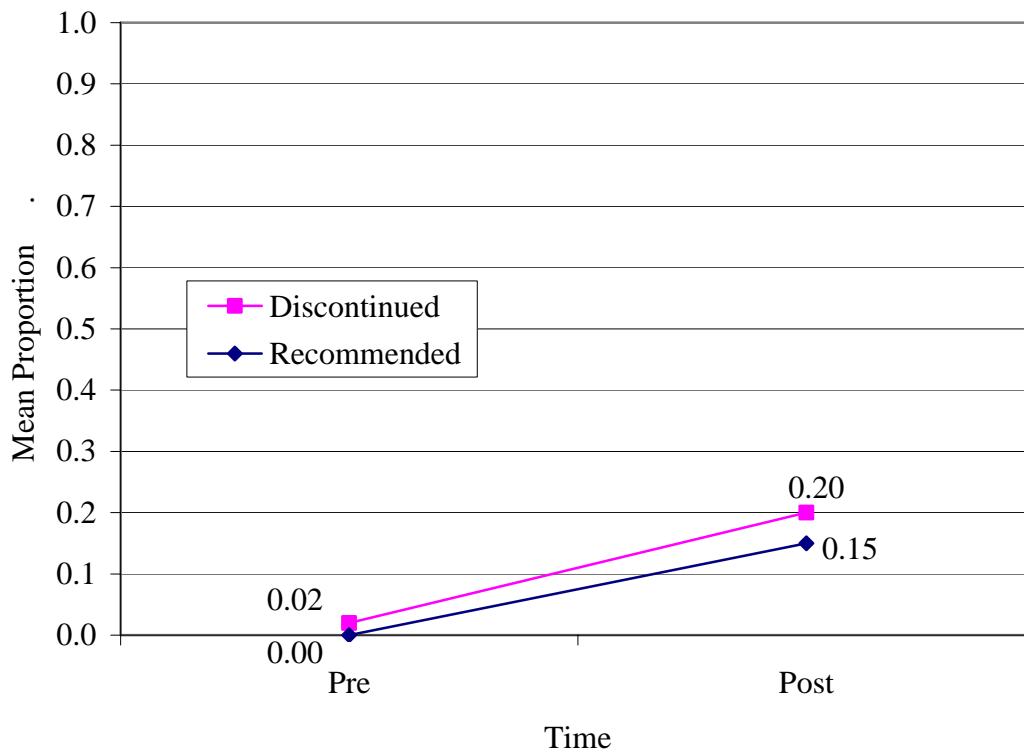


Figure 9. Orthographic Pattern Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

A child who wrote words with acceptable English letter sequences possessed orthographic acceptability knowledge, as evidenced by his/her performance on DeFord's (2000) Sentence Writing and Spelling task, scored with Bourassa and Treiman's (2003) procedures. This task required a child to listen to the words in a dictated story and write them with acceptable English letter sequences. An average-performing recommended child demonstrated less orthographic acceptability knowledge than an average-performing discontinued child at pre- and post-tutoring and increased this knowledge from pre- to post-tutoring (see Figure 10). This improvement may promote his/her perception of words during the reading of texts and/or generation of words during the



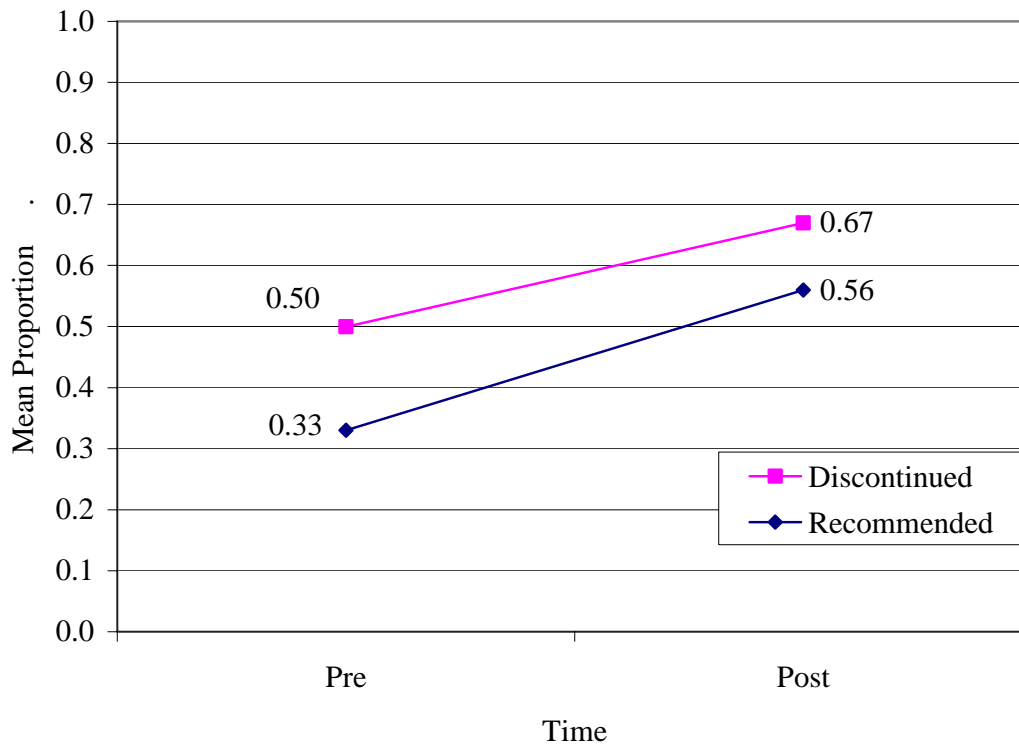


Figure 10. Orthographic Acceptability Knowledge of an Average-performing Recommended and Discontinued Child at Pre- and Post-tutoring

writing of texts. Although an average-performing recommended child possessed virtually no knowledge of specific letter patterns at pre- and post-tutoring (see Figure 9), this child did possess a more general knowledge of acceptable English letter sequences at pre- and post-tutoring (see Figure 10). Interestingly, parallel improvement marked the pre- to post-tutoring gains of both the average-performing recommended and discontinued child (see Figure 10).

#### *Oral Reading Processing*

In this study, a child revealed his/her oral reading processing by reading several GORT-4 stories out loud. A child's oral reading processing consisted of his/her oral reading accuracy, oral reading inaccuracy, and oral reading behaviors. The GORT-4

original and modified scoring procedures determined a child's oral reading processing.

A child's oral reading accuracy consisted of the number of correct oral reading responses that s/he made while reading the GORT-4 stories. When scored with the original GORT-4 scoring procedure, an average-performing recommended child produced an accuracy score more than two times below an average-performing discontinued child (see Figure 11). The original scoring procedure considered substitutions, tolds, omissions, insertions, self-corrections, and repetitions as errors. Similarly, when scored with the modified GORT-4 scoring procedure, an average-performing recommended child produced an accuracy proportion almost two times below an average-performing discontinued child (see Figure 12). The modified procedure considered substitutions, tolds, omissions, and insertions as errors. An average-performing recommended child's limited ability to read texts accurately may indicate difficulty in one or more of the many aspects of early literacy development, including (a) reading unfamiliar words by decoding, analogizing, and/or using context; (b) constructing meaning while reading familiar and unfamiliar words; and/or (c) storing familiar phonemes, graphemes, words, and/or word meanings in the lexicon and retrieving them.

A child's oral reading inaccuracy consisted of the number of incorrect oral reading responses or errors that s/he made while reading story 1 and 2 of the GORT-4, employing the modified GORT-4 scoring procedure. The modified procedure considered substitutions, tolds, omissions, and insertions as errors. An average-performing recommended child read with almost twice as many overall errors (i.e., combined substitutions, tolds, omissions, and insertions) and substitutions as an average-performing discontinued child (see Figure 13). Additionally, an average-performing recommended

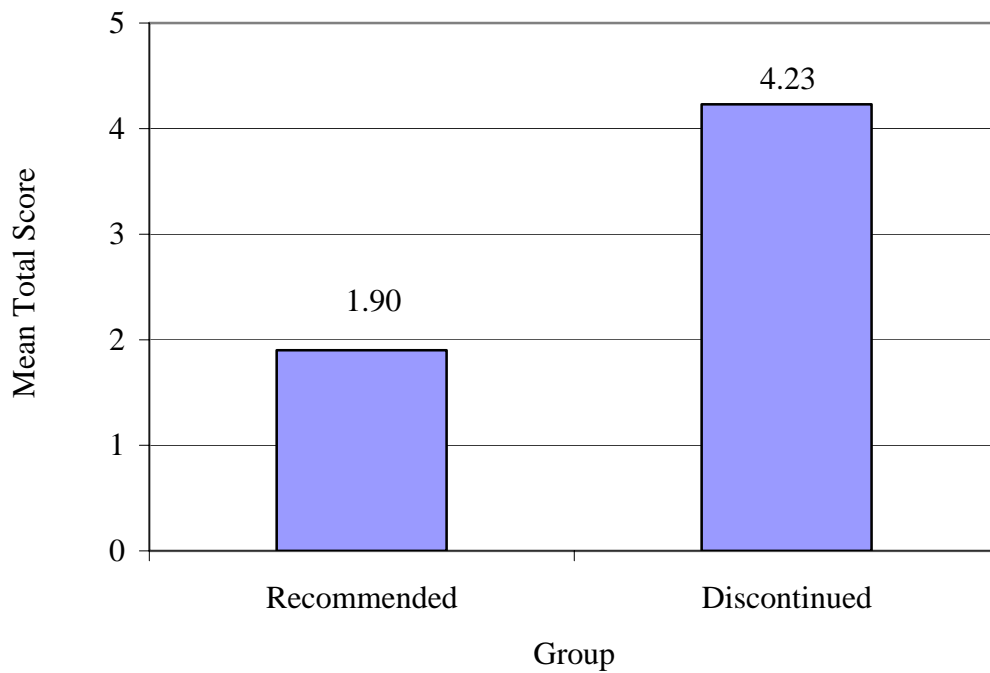


Figure 11. Oral Reading Accuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring

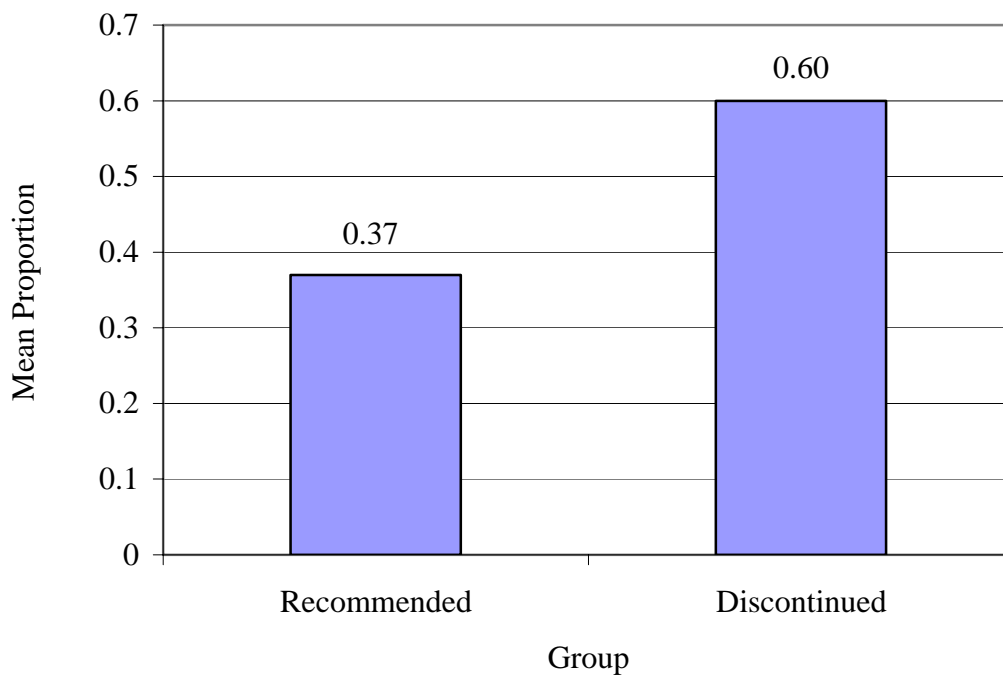


Figure 12. Modified Oral Reading Accuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring

child read with more than one and half times as many tolds as an average-performing discontinued child (see Figure 13). An average-performing recommended child's large number of overall errors, substitutions, and tolds supports the finding that s/he encountered difficulty in oral reading accuracy. Furthermore, an average-performing recommended child made approximately the same small number of omissions and insertions as the average-performing discontinued child (see Figure 13). This finding may signal an average-performing recommended child's (a) ability to match spoken words to printed words during the reading of texts and/or (b) slow, word-by-word reading of texts.

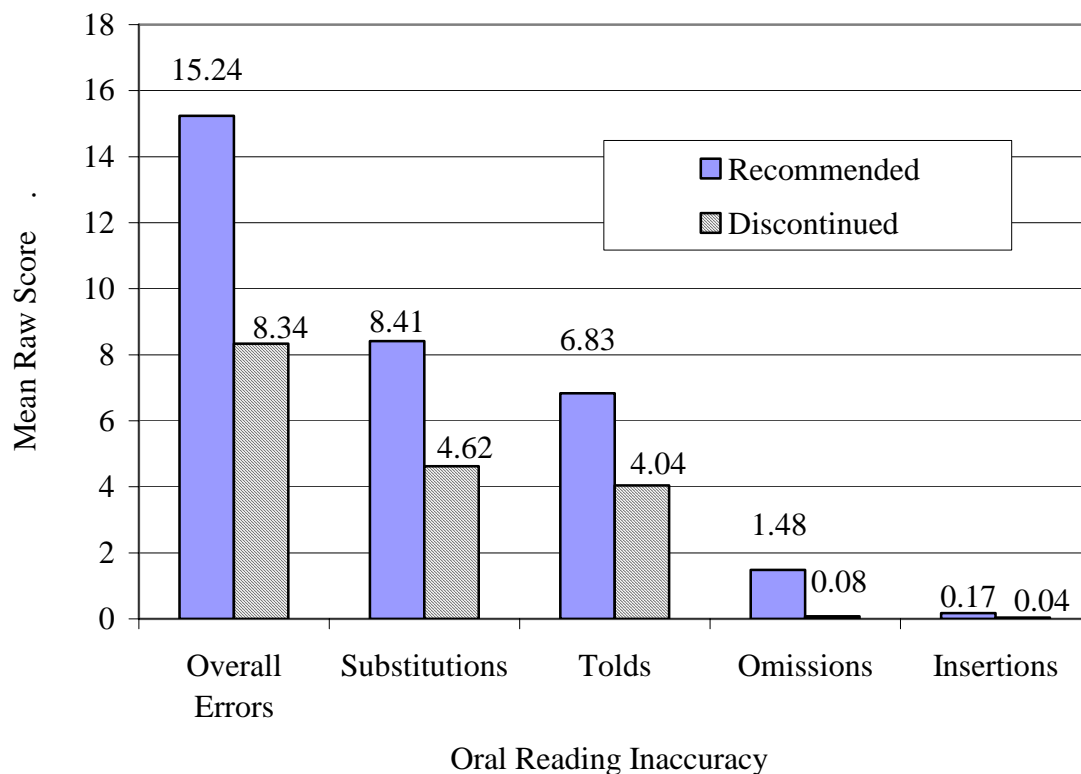


Figure 13. Oral Reading Inaccuracy of an Average-performing Recommended and Discontinued Child at Post-tutoring

A child's oral reading behaviors included (a) the rate and fluency (i.e., combined rate and accuracy) at which s/he read the GORT-4 stories and (b) the number of self-corrections and repetitions s/he produced while reading story 1 and 2 of the GORT-4. An average-performing recommended child read approximately two times slower and with two times less fluency than an average-performing discontinued child, using the original GORT-4 scoring procedures (see Figure 14). This finding may indicate difficulty in one or more of the many aspects of early literacy development, including (a) monitoring the reading process at the word, phrase, sentence, and/or text level; (b) reading unfamiliar graphemes, words, and/or word meanings in the lexicon and retrieving them; and/or (d) visual perceptual problems. Furthermore, an average-performing recommended child made fewer self-corrections and more repetitions than an average-performing discontinued child, using the modified GORT-4 scoring procedures. However, these differences between the two children were small (see Figure 15). The low number of self-corrections may indicate that an average-performing recommended child encountered difficulty in one or more of the many aspects of early literacy development, including (a) monitoring the reading process at the word, phrase, sentence and/or text level; (b) reading unfamiliar words by decoding, analogizing, and/or using context; and/or (c) storing familiar phonemes, graphemes, words, and/or word meanings in the lexicon and retrieving them. The large number of repetitions may indicate that an average-performing recommended child (a) monitored the reading process at the word, phrase, sentence and/or text level; (b) stalled to avoid reading or to allow for processing of an up-coming unfamiliar word in the texts; and/or (c) confirmed accurate reading of words, phrases, and/or sentences.

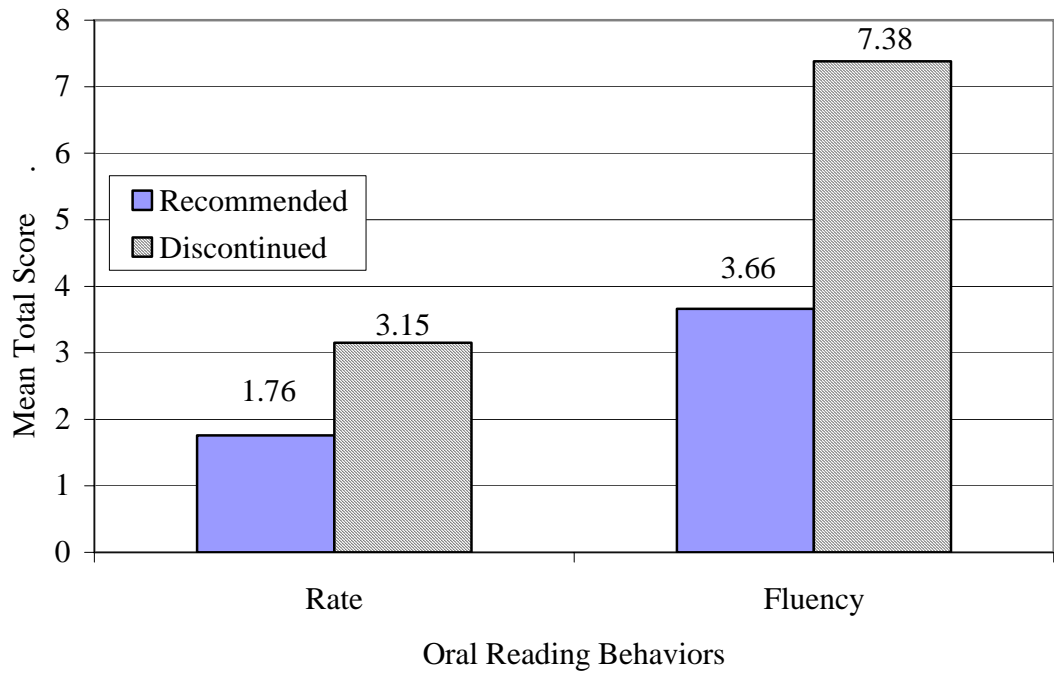


Figure 14. Oral Reading Behaviors of Rate and Fluency of an Average-performing Recommended and Discontinued Child at Post-tutoring

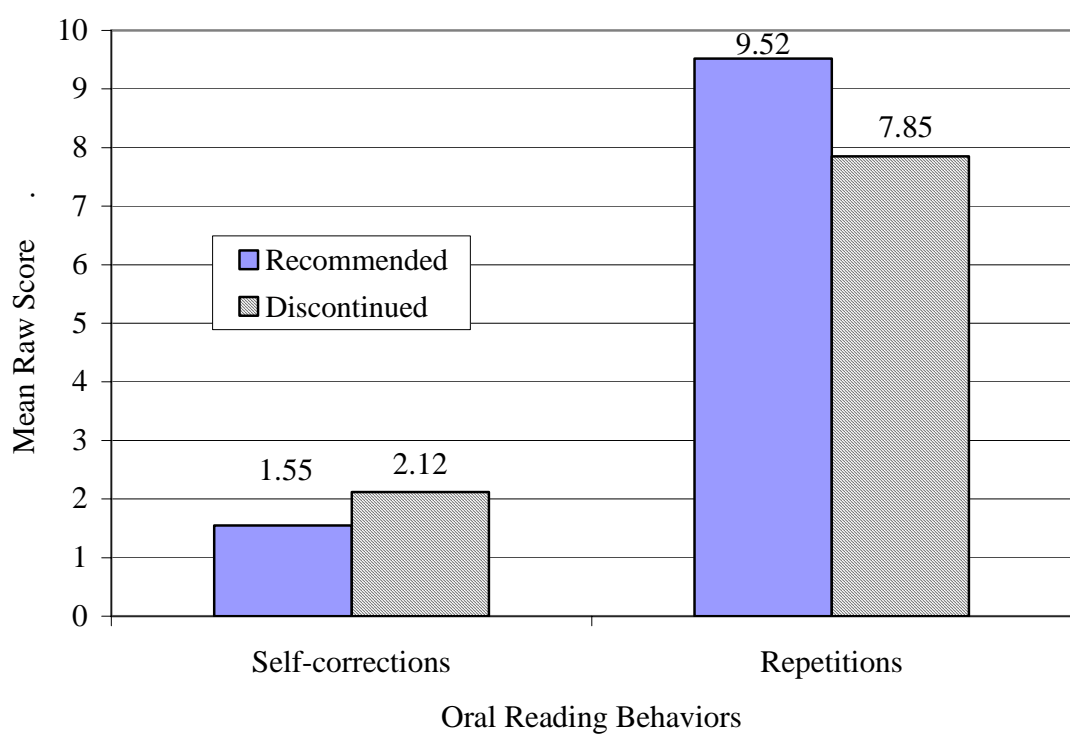


Figure 15. Oral Reading Behaviors of Self-corrections and Repetitions of an Average-performing Recommended and Discontinued Child at Post-tutoring

### *Reading Comprehension Processing*

In this study, a child revealed his/her reading comprehension processing by responding orally to text-related questions after reading several GORT-4 stories out loud. The GORT-4's original scoring procedures determined a child's reading comprehension processing. An average-performing recommended child responded correctly to virtually the same number of text-related questions as an average-performing discontinued child (see Figure 16). An average-performing recommended child's limited ability to respond correctly to text-related questions may indicate difficulty in one or more of the many aspects of early literacy development, including (a) engaging reading comprehension strategies; (b) monitoring reading comprehension; (c) building, storing, and accessing background and/or conceptual knowledge; (d) storing word meanings in the lexicon and retrieving them; (e) oral word reading accuracy; and/or (f) oral word reading behaviors of rate and fluency.

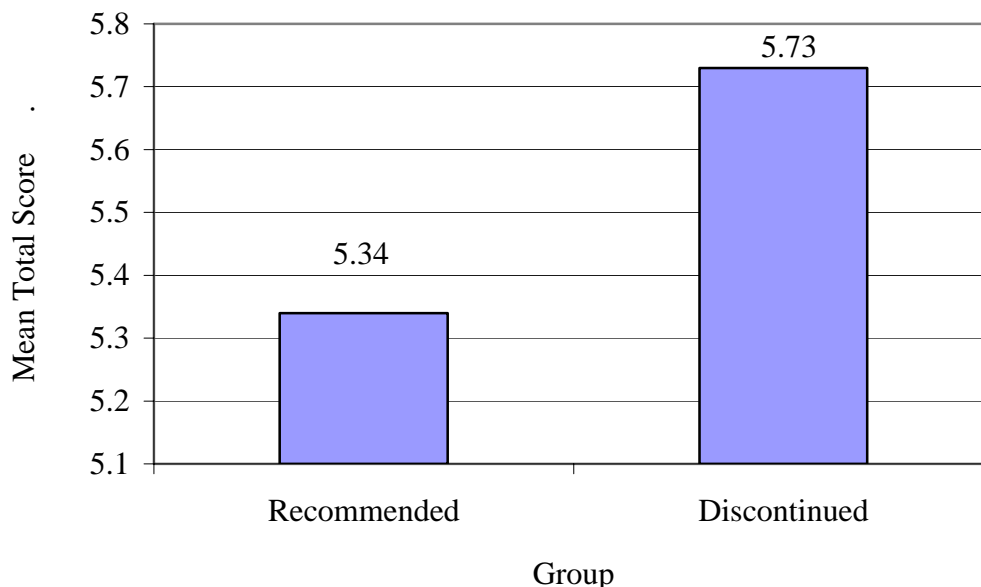


Figure 16. Reading Comprehension Processing of an Average-performing Recommended and Discontinued Child at Post-tutoring

*Summary*

An average-performing recommended child demonstrated variable phonological awareness and orthographic knowledge. For instance, an average-performing recommended child displayed substantial pre- to post-tutoring gains in phonological skeletal structure awareness and virtually no pre- to post-tutoring gains in combined syllable, onset and rime, and phonemic awareness. Although an average-performing recommended child improved in parallel with an average-performing discontinued child from pre- to post-tutoring on phonological awareness and orthographic knowledge, an average-performing recommended child performed at a level consistently below an average-performing discontinued child. Furthermore, upon the completion of tutoring, an average-performing recommended child performed at a level below and virtually equivalently to an average-performing discontinued child on oral reading processing and reading comprehension processing, respectively.

*Summary*

Quantitative analyses revealed two significant findings concerning the composites of overall phonological awareness and overall orthographic knowledge. First, recommended children performed below discontinued children on overall phonological awareness and overall orthographic knowledge. Second, recommended and discontinued children combined displayed gains in overall phonological awareness and overall orthographic knowledge from pre- to post-tutoring. These findings suggest that the two groups improved in parallel over time with the recommended children scoring below the discontinued children at pre- and post-tutoring.

Recommended and discontinued children differed on three of the four early



literacy components that contributed to the overall phonological awareness composites. Recommended children performed below discontinued children on rhyme awareness at pre-tutoring, phonological skeletal structure awareness at pre- and post-tutoring, and graphophonemic awareness with respect to beginning phonemes at pre- and post-tutoring and ending phonemes at pre-tutoring at statistically significant levels. All other differences between the two groups of children on rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness can be attributable to chance.

Recommended and discontinued children also differed on two of the early literacy components that formulated the overall orthographic knowledge composites. Recommended children performed below discontinued children on spelling knowledge at post-tutoring and orthographic acceptability knowledge at pre-tutoring at statistically significant levels. Any other differences between the two groups of children on spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge could be due to chance.

Additional quantitative analyses revealed several important findings concerning oral reading processing. At post-tutoring, recommended children read the GORT-4 stories with less accuracy, at a slower rate, and with less fluency than the discontinued children at statistically significant levels. Additionally, recommended children read story 1 and 2 from the GORT-4 with more overall errors and more substitutions than discontinued children at statistically significant levels.

Standard score comparisons revealed that (a) recommended children performed below or equivalent to discontinued children on GORT-4 accuracy, rate, and fluency and

(b) recommended and discontinued children performed comparably on GORT-4 comprehension. In addition, a composite standard score comparison revealed that the recommended children performed below or equivalent to the discontinued children on combined oral reading and reading comprehension processing. Furthermore, an additional comparison to normative sample distribution revealed that the vast majority of the recommended and discontinued children performed poorly or very poorly on combined oral reading and reading comprehension processing.

A profile, based on mean scores, revealed that an average-performing recommended child demonstrated variable phonological awareness and orthographic knowledge. Although an average-performing recommended child improved in parallel with an average-performing discontinued child from pre- to post-tutoring on phonological awareness and orthographic knowledge, an average-performing recommended child performed consistently below an average-performing discontinued child. This profile also revealed that at post-tutoring an average-performing recommended child performed at a level below and at virtually the same level as an average-performing discontinued child on oral reading processing and reading comprehension processing, respectively.

This study operated under the hypothesis that recommended children perform at a level below discontinued children on (a) phonological awareness and orthographic knowledge prior to and following tutoring and (b) oral reading processing and reading comprehension processing following tutoring. This study produced several statistically significant findings that supported this hypothesis. Furthermore, numerous nonsignificant findings provided information on recommended children's phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing.

Chapter 5 discusses this study's major findings in relation to the statistically significant findings that supported this study's hypothesis and the nonsignificant findings that offered interesting information on recommended children. Chapter 5 also offers recommendations for instruction and future research.

## Chapter 5

### *Discussion*

Chapter 5 is organized into six sections. Sections 1 and 2 summarize the current study and present the major findings, respectively. Section 3 addresses this study's limitations. Sections 4 and 5 offer recommendations for instruction and future research, respectively.

### *Study Summary*

In the United States, first-round Reading Recovery children receive tutoring for approximately the first 20 weeks of the school year from specially trained Reading Recovery teachers. A sizable proportion of these children do not respond well to this tutoring and fail to meet the criteria for successful performance (Gómez-Bellengé et al., 2003). Their Reading Recovery teachers (a) assign them to the end-of-program status category of recommended and (b) refer them for additional assessment and/or consideration for other supplemental instruction. To date, four studies have examined recommended children's phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004).

In response to this emerging program of research, this study compared recommended to discontinued children on (a) phonological awareness and orthographic knowledge prior to and following their tutoring and (b) oral reading processing and reading comprehension processing following their tutoring. The following research questions guided this study:

- (a) How do recommended Reading Recovery children compare to discontinued

Reading Recovery children on phonological awareness prior to and following their completion of tutoring?

(b) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on orthographic knowledge prior to and following their completion of tutoring?

(c) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on oral reading processing following their completion of tutoring?

(d) How do recommended Reading Recovery children compare to discontinued Reading Recovery children on reading comprehension processing following their completion of tutoring?

To answer these research questions, I individually administered several early literacy measures to Reading Recovery children to assess their (a) phonological awareness and orthographic knowledge prior to and following their tutoring and (b) oral reading processing and reading comprehension processing following their tutoring. I conducted this study in a single school district. At the beginning of the study, the sample consisted of 60 Reading Recovery children who were taught by 15 trained Reading Recovery teachers and one Reading Recovery teacher leader. By the end of the study, the sample consisted of 55 Reading Recovery children: 29 recommended children and 26 discontinued children.

The Reading Recovery teachers selected children to receive Reading Recovery tutoring based on the children's scores on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and on the classroom teachers' recommendations. I

administered the Rhyme Detection task (Muter et al., 1997), the Blending Words task (Wagner et al., 1999), and the Sentence Writing and Spelling task (DeFord, 2000) to assess the children's phonological awareness and orthographic knowledge prior to their tutoring. Approximately 20 weeks later, the Reading Recovery teachers (a) re-administered Clay's measure, (b) determined if the children responded well or failed to respond well to tutoring, and (c) assigned the end-of-program status categories of discontinued or recommended, respectively. At this time, I re-administered the pre-tutoring measures to reassess the children's phonological awareness and orthographic knowledge and administered the GORT-4 (Wiederholt & Bryant, 2001) to assess the children's oral reading processing and reading comprehension processing.

To describe the recommended children's phonological awareness and orthographic knowledge at pre- and post-tutoring and oral reading processing and reading comprehension processing at post-tutoring, I compared their performance on the various early literacy measures to the performance of the discontinued children. I reported means, standard deviations, and percentages. I conducted a number of statistical tests on these descriptive statistics, including repeated measure two-way analyses of variance, chi-square tests of independence, and one-way analyses of variance. The findings contribute to an understanding of recommended children's early literacy development.

### *Major Findings*

Previous research has indicated that recommended Reading Recovery children's early literacy development lags behind discontinued Reading Recovery children's early literacy development prior to and following tutoring (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004). Based on these findings, I

hypothesized that recommended children would perform at a level below discontinued children on phonological awareness and orthographic knowledge at pre- and post-tutoring and (b) oral reading processing and reading comprehension processing at post-tutoring. This section reviews the statistically significant results that supported this hypothesis. It also reviews the nonsignificant results that failed to support this hypothesis, but provided interesting information on recommended children. Furthermore, this section links these results to previously conducted research on recommended children.

### *Phonological Awareness*

At the onset of this study, I hypothesized that recommended children would perform at a level below discontinued children on phonological awareness prior to and following their tutoring. The results of this study confirmed this hypothesis in regards to overall phonological awareness in the form of a composite, as well as rhyme awareness, phonological skeletal structure awareness, and graphophonemic awareness. Interestingly, recommended and discontinued children combined improved in overall phonological awareness from the beginning to the end of their tutoring at a statistically significant level. Recommended children possessed less overall phonological awareness than discontinued children at a statistically significant level. Before tutoring, recommended children demonstrated less rhyme awareness, phonological skeletal structure awareness, and graphophonemic awareness with respect to beginning and ending phonemes than discontinued children at statistically significant levels. After tutoring, recommended children displayed less phonological skeletal structure awareness and graphophonemic awareness with respect to beginning phonemes than discontinued children at statistically significant levels. These statistically significant results support this study's hypothesis

that recommended children perform at a level below discontinued children on phonological awareness prior to and following tutoring.

Some of the results corresponding to phonological awareness were nonsignificant, and therefore failed to support this study's hypothesis. Yet these nonsignificant results provided interesting information about recommended children's phonological awareness prior to and following tutoring. Before tutoring, recommended children did not demonstrate less (a) combined syllable, onset and rime, and phonemic awareness and (b) graphophonemic awareness with respect to middle phonemes than discontinued children at statistically significant levels. The two groups produced nearly identical mean scores and percentages. These results are interesting because they suggest that the two groups started their tutoring with the same level of awareness. After tutoring, recommended children did not display less rhyme awareness and graphophonemic awareness with respect to ending phonemes than discontinued children at statistically significant levels. The two groups produced high mean scores and percentages that were nearly equivalent. These results are interesting because they suggest that recommended children (a) developed substantial rhyme awareness and graphophonemic awareness with respect to ending phonemes and (b) possessed relatively the same level of rhyme awareness and graphophonemic awareness with respect to ending phonemes as discontinued children following their tutoring. After tutoring, recommended children did not demonstrate less combined syllable, onset and rime, and phonemic awareness than discontinued children at a statistically significant level. The two groups produced low mean scores and percentages that were nearly equivalent. This turned out to be an important finding. Although Clay (1993) designed the Reading Recovery teaching procedures to develop



children's ability to convert graphemes in text to sound segments, and then to blend these sounds segments into recognizable words, the recommended and discontinued children made minimal pre- to post-tutoring progress in their ability to blend isolated sound segments in speech into recognizable words.

This study added to an emerging program of research on recommended children's phonological awareness. No other studies have compared recommended to discontinued children on (a) rhyme awareness; (b) phonological skeletal structure awareness; (c) combined syllable, onset and rime, and phonemic awareness; and/or (d) graphophonemic awareness with respect to beginning, middle, and ending phonemes prior to and following their tutoring. A few studies have investigated recommended and discontinued children's phonemic awareness and graphophonemic awareness (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004). Spector and Moore found that recommended children demonstrated less phonemic and graphophonemic awareness than discontinued at pre-tutoring. However, the differences in the mean scores failed to reach statistical significance. Center et al. reported recommended and discontinued children's mean scores on a phonemic awareness measure at pre-tutoring. A comparison of the mean scores revealed that recommended children demonstrated less phonemic awareness than discontinued children. Likewise, Chapman et al. reported recommended and discontinued children's means scores on a phonemic awareness measure and a graphophonemic awareness measure at pre- and post-tutoring. A comparison of these mean scores revealed that recommended children demonstrated less phonemic and graphophonemic awareness than discontinued children. Additionally, Clay and Tuck reported the stanines corresponding to recommended and discontinued

children's graphophonemic awareness at pre- and post-tutoring. A comparison of stanines revealed that recommended children generally demonstrated less graphophonemic awareness than discontinued children.

Center et al. (1995), Chapman et al. (2001), and Spector and Moore (2004) administered phonological awareness measures other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to assess the children's phonemic and graphonemic awareness. My study built upon these studies by also administering measures independent of Reading Recovery. The Rhyme Detection task (Muter et al., 1997) assessed the children's rhyme awareness. The Blending Words task (Wagner et al., 1999) assessed the children's combined syllable, onset and rime, and phonemic awareness. The Sentence Writing and Spelling task (DeFord, 2000) assessed the children's phonological skeletal structure awareness and graphophonemic awareness with respect to beginning, middle, and ending phonemes, using two different scoring procedures. These three tasks also assessed the children's overall phonological awareness.

Spector and Moore (2004) conducted statistical analyses to test for statistically significant group differences between recommended and discontinued children on phonemic and graphophonemic awareness at pre- tutoring. My study added to Spector and Moore's research by conducting statistical analyses to test for (a) between- and within-group differences on overall phonological awareness and (b) group differences on rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; and graphophonemic awareness with respect to beginning, middle, and ending phonemes at pre- and post-tutoring.

### *Orthographic Knowledge*

At the start of this study, I hypothesized that recommended children would perform at a level below discontinued children on orthographic knowledge prior to and following their tutoring. The results of this study confirmed this hypothesis in regards to overall orthographic knowledge in the form of a composite, as well as orthographic acceptability knowledge and spelling knowledge. Interestingly, recommended and discontinued children combined improved in overall orthographic knowledge from the beginning to the end of their tutoring at a statistically significant level. Recommended children possessed less overall orthographic knowledge than discontinued children at a statistically significant level. Before tutoring, recommended children demonstrated less orthographic acceptability knowledge than discontinued children at a statistically significant level. After tutoring, recommended children displayed less spelling knowledge than discontinued children at a statistically significant level. These statistically significant results support this study's hypothesis that recommended children perform at a level below discontinued children on orthographic knowledge prior to and following tutoring.

Some of the results corresponding to orthographic knowledge were nonsignificant, and therefore failed to support this study's hypothesis. Yet these nonsignificant results provided interesting information about recommended children's orthographic knowledge prior to and following tutoring. Prior to tutoring, recommended children did not demonstrate less spelling knowledge and orthographic pattern knowledge than discontinued children at statistically significant levels. The two groups produced low mean scores and percentages that were nearly equivalent and virtually identical on

spelling knowledge and orthographic pattern knowledge, respectively. These results are interesting because they indicate that both groups started their tutoring with nearly the same low levels of knowledge. Following tutoring, recommended children did not demonstrate less orthographic pattern knowledge than discontinued children at a statistically significant level. The two groups produced low mean scores and percentages that were nearly equivalent. Although Clay (1993) developed the Reading Recovery teaching procedures to advance children's knowledge of letter patterns, the recommended and discontinued children made minimal pre- to post-tutoring progress in orthographic pattern knowledge.

This study added to an emerging program of research on recommended children's orthographic knowledge. Although no other studies have compared recommended to discontinued children on overall orthographic knowledge, orthographic pattern knowledge, and/or orthographic acceptability knowledge prior to and following their tutoring, Chapman et al.'s (2001) research compared recommended and discontinued children's spelling knowledge. These researchers reported recommended and discontinued children's mean scores on a spelling measure at post-tutoring. A comparison of these mean scores revealed that recommended children demonstrated less phonemic and graphophonemic awareness than discontinued children.

Chapman et al. (2001) administered an orthographic knowledge measure other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to assess the children's spelling knowledge. My study added to Chapman et al.'s study by also administering a measure independent of Reading Recovery. The Sentence Writing and Spelling task (DeFord, 2000) assessed the children's overall orthographic knowledge,

spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge, using various scoring procedures.

This study also added to the emerging program of research on recommended children by conducting statistical analyses. In a manner analogous to phonological awareness, this study tested for (a) between- and within-group differences on overall orthographic knowledge and (b) group differences on spelling knowledge, orthographic pattern knowledge, and orthographic acceptability knowledge at pre- and post-tutoring.

#### *Oral Reading Processing*

At the onset of this study, I hypothesized that recommended children would perform at a level below discontinued children on oral reading processing following their tutoring. The results of this study confirmed this hypothesis in regards to (a) oral reading accuracy; (b) oral reading inaccuracy, as evidenced by the children's overall errors and substitutions; and (c) oral reading behaviors, as evidenced by the children's rate and fluency. After tutoring, recommended children read (a) the GORT-4 stories with less accuracy, (b) story 1 and 2 of the GORT-4 with more overall errors and substitutions, and (c) the GORT-4 stories at a slower rate and with less fluency than discontinued children at statistically significant levels. These statistically significant results support this study's hypothesis that recommended children perform at a level below discontinued children on oral reading processing following tutoring.

This study added to an emerging program of research on recommended children's oral reading processing. Although no other studies have compared recommended to discontinued children on oral reading inaccuracy and oral reading behaviors prior to and following their tutoring, Chapman et al.'s (2001) and Clay and Tuck's (1991) research

examined recommended and discontinued children's oral reading accuracy. Chapman et al. reported recommended and discontinued children's mean scores corresponding to oral reading accuracy at pre- and post-tutoring. A comparison of these mean scores revealed that recommended children read with less accuracy than discontinued children.

Additionally, Clay and Tuck reported the text levels (range 0-23) that recommended and discontinued children read at pre- and post-tutoring. A comparison of these text levels revealed that recommended children generally read lower text levels than discontinued children.

Chapman et al. (2001) administered an oral reading processing measure other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) to assess the children's oral reading accuracy. My study added to Chapman et al.'s study by also administering an oral reading processing measure independent of Reading Recovery. The GORT-4 (Wiederholt & Bryant, 2001) assessed the children's (a) oral reading accuracy; (b) oral reading inaccuracy that included the children's overall errors, substitutions, tolds, omissions, and insertions; and (c) oral reading behaviors that included the children's rate, fluency, self-corrections, and repetitions.

This study also added to the emerging program of research on recommended children by conducting statistical analyses. This study determined whether mean score differences between recommended and discontinued children on oral reading processing were statistically significant differences or differences attributable to chance.

#### *Reading Comprehension Processing*

At the beginning of this study, I hypothesized that recommended children would perform at a level below discontinued children on reading comprehension processing

following their tutoring. The recommended children responded correctly to nearly the same number of text-related questions as the discontinued children following their completion of tutoring. The two groups produced low mean scores that were virtually identical. This nonsignificant result did not support this study's hypothesis that recommended children perform at a level below discontinued children on reading comprehension processing after tutoring.

However, this nonsignificant result provided interesting information about recommended children's reading comprehension processing following tutoring. This result suggests that recommended children, as well as discontinued children, failed to develop adequately reading comprehension processing. This result is interesting because although Clay (1993) developed the Reading Recovery teaching procedures to promote children's reading comprehension monitoring, the recommended and discontinued children demonstrated a limited ability to respond correctly to text-related questions following their tutoring.

This study added to an emerging program of research on recommended children's reading comprehension processing. Chapman et al.'s (2001) research examined recommended and discontinued children's reading comprehension processing. They reported recommended and discontinued children's mean scores corresponding to their ability to respond correctly to text-related questions at pre- and post-tutoring. A comparison of these mean scores revealed that the two groups produced low mean scores that were nearly equivalent.

Chapman et al. (2001) administered a reading comprehension processing measure independent of Reading Recovery because *An Observation Survey of Early*

*Literacy Achievement* (Clay, 2002) does not assess formally children's reading comprehension processing. My study added to Chapman et al.'s study by also administering a measure independent of Reading Recovery. The GORT-4 (Wiederholt & Bryant, 2001) measured the children's reading comprehension processing by assessing their ability to respond correctly to text-related questions following tutoring.

This study also added to the emerging program of research on recommended children by conducting statistical analyses. In a manner analogous to oral reading processing, this study determined whether mean score differences between the recommended and discontinued children on reading comprehension processing were statistically significant differences or differences attributable to chance.

### *Summary*

This study hypothesized that recommended children would perform at a level below discontinued children on (a) phonological awareness and orthographic knowledge at pre- and post-tutoring and (b) oral reading processing and reading comprehension processing at post-tutoring. The results of this study that yielded statistically significant outcomes confirmed this hypothesis in regards to phonological awareness, orthographic knowledge, and oral reading processing. A result that yielded a nonsignificant outcome failed to confirm this hypothesis in regards to reading comprehension processing. The results of the current study built upon an emerging program of research (Center et al., 1995; Chapman et al., 2001; Clay & Tuck, 1991; Spector & Moore, 2004) by examining recommended children's overall phonological awareness; rhyme awareness; phonological skeletal structure awareness; combined syllable, onset and rime, and phonemic awareness; overall orthographic knowledge; orthographic pattern knowledge;



orthographic acceptability knowledge; oral reading inaccuracy; oral reading behaviors, and responding to text-related questions. Furthermore, this study added to these previously conducted studies by employing early literacy measures other than *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and by conducting statistical analyses at pre- and post-tutoring.

### *Limitations*

This study revealed several major findings on recommended children's phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. These findings, however, are characterized by several noteworthy limitations.

Reading Recovery teachers make instructional decisions, such as the selection and employment of the Reading Recovery teaching procedures, in response to their Reading Recovery children's attempts to read texts, write texts, and complete isolated word analysis activities. In this study, variability may have existed among the Reading Recovery teachers' responses to their children's reading and writing attempts. My decision to exclude two Reading Recovery teachers-in-training due to their lack of knowledge of and proficiency in the Reading Recovery teaching procedures may have reduced this variability. This study did not examine the literacy instruction that the Reading Recovery teachers provided to their children. Observations of the teachers' literacy instruction may have provided information about the differences between the recommended and discontinued children's early literacy development. Thus, differences in individually designed diagnostic instruction may have influenced the recommended children's responsiveness to tutoring.

Similarly, first-grade classroom teachers design and deliver literacy instruction based on many factors, such their philosophy of early literacy development and instruction, their pedagogical knowledge of early literacy instruction, the early literacy strengths and needs of the children in their classrooms, and/or the established early literacy programs in their schools or school districts. In this study, variability may have existed among the classroom teachers' literacy instruction. Interestingly, the classroom teachers implemented a new literacy program that featured systematic phonics instruction at the time of data collection. This study did not examine the literacy instruction that classroom teachers provided to the children in their classrooms. Observations of the teachers' literacy instruction may have provided information about the differences between the recommended and discontinued children's early literacy development. Thus, differences in classroom literacy instruction may have influenced the recommended children's responsiveness to tutoring.

#### *Recommendations for Instruction*

This section presents instructional recommendations for recommended children. These recommendations extend to Reading Recovery tutoring, first-grade instruction provided concurrently with tutoring, and kindergarten instruction provided prior to tutoring.

#### *Recommendations for Reading Recovery Tutoring and First-grade Instruction*

Recommended children made considerable gains in phonological skeletal structure awareness and graphophonemic awareness from pre- to post-tutoring (see Table 21). These pre- to post-tutoring gains make a case for the continued employment of the Reading Recovery teaching procedures and first-grade classroom instruction that

supports the instructional objectives of Reading Recovery tutoring.

The recommended children's pre- to post-tutoring gains in phonological skeletal structure awareness and graphophonemic awareness suggest that they may benefit from their Reading Recovery teachers' continued instruction. Clay designed the Reading Recovery teaching procedures to develop children's ability to write unfamiliar words by articulating words slowly, recognizing the phonemes in these words, matching these phonemes to their corresponding graphemes, and writing these graphemes in sound boxes. Sound boxes are teacher drawn boxes that correspond to the number of phonemes in unfamiliar words. Recommended children would also profit from their first-grade classroom teachers providing similar instruction and practice during text writing activities.

Conversely, recommended children made minimal to small gains in rhyme awareness; combined syllable, onset and rime, and phonemic awareness; spelling knowledge; orthographic pattern knowledge; and orthographic acceptability knowledge from pre-to post-tutoring (see Tables 21 and 29). These pre- to post-tutoring gains make a case for the employment of modified and/or new Reading Recovery teaching procedures and underscores the need for first-grade classroom instruction that supports the instructional goals of Reading Recovery tutoring.

Clay (1993) did not design the Reading Recovery teaching procedures to specifically develop children's rhyme awareness. Rather, she designed them to develop children's ability to read and write unfamiliar words by thinking of familiar words that rhyme with these unfamiliar words. The recommended children's small pre- to post-tutoring gains in rhyme awareness suggest that they may profit from their Reading

Recovery teachers modifying the existing teaching procedures and/or employing new ones to develop their rhyme awareness directly. These procedures might focus on children (a) selecting words from texts and subsequently generating other words that rhyme and/or (b) detecting words that rhyme from among several possible word choices. Such procedures might be employed during text reading, text writing, and/or word analysis activities. The recommended children's small pre-to post-tutoring gains also indicate that they may profit from their first-grade classroom teachers providing instruction and practice in producing and detecting rhyme during literacy activities that involve rhyming games, texts, and songs, and the reading or writing of unfamiliar words by analogy.

Clay (1993) did not design the Reading Recovery teaching procedures to develop children's ability to blend the isolated sound segments of syllables, onsets and rimes, and phonemes in speech into recognizable words. Rather, she designed them to develop children's ability to blend these sound segments into recognizable words during the reading of unfamiliar words in texts by decoding and/or analogizing. The recommended children's minimal pre- to post-tutoring gains in combined syllable, onset and rime, and phonemic awareness suggest that they may profit from their Reading Recovery teachers employing new teaching procedures. These procedures might teach recommended children initially to blend syllables in speech into recognizable words, and then teach them to blend onsets and rimes, followed by phonemes. Once recommended children progress in their ability to blend sound segments in speech into recognizable words, these new procedures might include the use of graphemes as markers for sounds to support the blending of sound segments during the reading of unfamiliar words by decoding and

analogizing (National Reading Panel, 2000). The recommended children's minimal pre- to post-tutoring gains also indicate that they may profit from their first-grade classroom teachers providing instruction and practice in blending isolated sound segments in speech into recognizable words during word analysis and/or text reading activities. Additionally, recommended children may benefit from their first-grade teachers providing instruction and practice in other syllable, onset and rime, phonemic awareness skills, such as segmenting spoken words into various sound segments.

Unlike rhyme awareness and combined syllable, onset and rime, and phonemic awareness, Clay (1993) designed the Reading Recovery teaching procedures to develop children's spelling knowledge. That is, she designed the procedures to increase the number of words that children write with correct spelling by having them write words quickly and repeatedly that appear frequently in their writing. The process of writing words quickly and repeatedly allows children to practice producing the sequences of letters so they learn to write these words with correct spelling. The recommended children's small pre- to post-tutoring gains in spelling knowledge suggest that they may benefit from their Reading Recovery teachers modifying the existing teaching procedures and/or employing new ones. These procedures might focus on teaching rimes and that these rimes are spelled the same in different words during the text reading, text writing and/or word analysis activities. Additionally, these procedures might involve generating and writing words with correct spelling that share common rimes during text writing activities. The recommended children's small pre- to post-tutoring gains also indicate that they may profit from their first-grade classroom teachers providing explicit spelling instruction and frequent practice in writing words with correct spelling that occur

frequently in their writing. Recommended children may also profit from their first-grade teachers creating classroom displays of words that occur frequently in their writing for them to refer to during text writing activities. Additionally, recommended children may benefit from completing word sorts in which they group words by rimes.

Similar to spelling knowledge, Clay (1993) designed the Reading Recovery teaching procedures to advance children's orthographic pattern knowledge and orthographic acceptability knowledge. That is, she designed these procedures to develop children's ability to write unfamiliar words with correct letter patterns and acceptable English letter sequences by teaching them to think of familiar words with the same letters patterns and sequences, and then to write these unfamiliar words in letter boxes. Letter boxes are teacher drawn boxes that correspond to the number of graphemes in unfamiliar words. The recommended children's minimal pre- to post-tutoring gains suggest that they may profit from their Reading Recovery teachers modifying the existing teaching procedures and/or employing new ones. These procedures might explicitly teach specific letter patterns and sequences and that these letter patterns and sequences appear in similar-sounding and similar-looking words. Such procedures might be employed during text reading, text writing, and/or word analysis activities. Additionally, these procedures might involve generating and writing words with correct letter patterns and acceptable letter sequences. The recommended children's minimal pre- to post-tutoring gains also indicate that they may profit from their first-grade classroom teachers providing explicit instruction in letter patterns and sequences during text reading, text writing, and/or word analysis activities. Recommended children may also profit from their first-grade teachers creating classroom displays of words that occur frequently in their writing for them to

refer to during text writing activities. Additionally, recommended children may benefit from completing word sorts in which they group words by letter patterns or sequences.

In addition to making minimal to small pre- to post-tutoring gains in phonological awareness and orthographic knowledge, recommended children also performed at a level below discontinued children on oral reading processing at post-tutoring. That is, recommended children read with less accuracy, with more overall errors, more substitutions, at a slower rate, and with less fluency than discontinued children at statistically significant levels (see Tables 36, 37, 40, 41, 46, 47). This poor performance of recommended children relative to discontinued children makes a case for the employment of modified and/or new Reading Recovery teaching procedures and highlights the need for first-grade classroom instruction that supports the instructional goals of Reading Recovery tutoring.

In regards to oral reading accuracy, Clay (1993) designed the Reading Recovery teaching procedures to develop children's ability to read familiar words by sight by having children read words in texts that the teacher isolates, preventing the use of context. Clay also designed the procedures to develop children's ability to read unfamiliar words by decoding, analogizing, and using context. That is, she designed the procedures to provide whole word instruction, as well as unsystematic phonics instruction.

Because the recommended children read with less accuracy, more overall errors, and more substitutions than the discontinued children at post-tutoring, they may benefit from their Reading Recovery teachers modifying the existing teaching procedures. These modified procedures might focus on lexical storage and retrieval of familiar words during

text reading and word analysis activities, as opposed to only during text reading activities. These modified procedures also might explicitly and systematically teach a planned and sequential set of letter-sound correspondences as recommended children learn to read unfamiliar words by decoding, analogizing, and using context. Furthermore, these modified procedures might have recommended children read decodable texts that contain controlled vocabulary to provide practice in using these letter-sound correspondences to read unfamiliar words (National Reading Panel, 2000). The recommended children's low accuracy scores, high overall error scores, and high substitution scores also indicate that they may profit from their first-grade classroom teachers providing opportunities for lexical storage and retrieval of words, as well as explicit and systematic phonics instruction. Recommended children may also benefit from their first-grade teachers providing numerous opportunities to read decodable texts independently and/or with assistance. Additionally, recommended children may profit from their first-grade teachers developing the various knowledge sources that underlie oral reading accuracy, such as phonological awareness, orthographic knowledge, and knowledge of language structures.

In regards to oral reading behaviors, Clay (1993) designed the Reading Recovery teaching procedures to develop children's ability to read texts quickly and fluently. Specifically, Clay designed her procedures to train children's eyes to move quickly across lines of print, develop children's perceptual span of words, model for children fluent reading, teach children to attend to punctuation, develop children's phrasing and expression. Furthermore, Clay designed the procedures to incorporate the use of texts that facilitate fluency, such as texts with repeated phrases.



Because the recommended children read at a slower rate and with less fluency (i.e., combined rate and accuracy) than the discontinued children at post-tutoring, they may benefit from their Reading Recovery teachers employing new teaching procedures that develop their ability to read texts quickly and fluently. These new procedures might include repeated oral readings with teacher feedback and guidance (National Reading Panel, 2000). Although Clay's (1993) teaching procedures call for children to reread novel texts if their first readings of these texts are slow and labored, these procedures do not suggest that children engage in repeated oral reading of familiar texts, characterized by teacher feedback and guidance. The recommended children's low rate scores and fluency scores at post-tutoring also indicate that they may profit from their first-grade classroom teachers providing opportunities to engage in repeated oral reading in which they receive guidance and feedback from their teachers (National Reading Panel, 2000). Recommended children may also profit from their first-grade teachers developing the various knowledge sources (e.g., phonological awareness and orthographic knowledge) that underlie oral reading accuracy because fluency depends upon oral reading accuracy (National Reading Panel, 2000; Snow et al., 1998). Additionally, recommended children may benefit from their first-grade teachers providing instruction in the other components of fluency, such as expression and phrasing. Lastly, recommended children may benefit from their classroom teachers modeling how texts sound when read quickly and fluently.

Although recommended children performed at a level below discontinued children on oral reading processing at post-tutoring, they performed comparably to discontinued children on reading comprehension processing at post-tutoring. That is, recommended children responded correctly to virtually the same number of text-related

questions as discontinued children (see Table 51). The poor performance of recommended children, as well as discontinued children, on reading comprehension processing makes a case for the employment of new Reading Recovery teaching procedures and underscores the need for first-grade classroom instruction that supports the instructional objectives of Reading Recovery tutoring.

Clay (1993) designed the various Reading Recovery teaching procedures to develop children's reading comprehension monitoring, one of several aspects of reading comprehension processing. These procedures develop children's ability to evaluate and regulate their construction of meaning by teaching them to (a) reread words, phrases, and/or sentences in texts to construct, confirm, and/or disconfirm meaning and (b) search for and use semantic cues to self-correct oral reading attempts that do not make sense.

Because the recommended children responded correctly to virtually the same number of text-related questions as the discontinued children at post-tutoring, they may benefit from their Reading Recovery teachers employing new teaching procedures to develop their reading comprehension processing. These new procedures might focus on developing recommended children's ability to construct mental representations of texts by accessing the information contained directly in the texts and integrating this information with their background knowledge. These new procedures may also focus on developing recommended children's ability to access and use conceptual knowledge and knowledge of word meanings. Furthermore, these new procedures may emphasize the explicit instruction of reading comprehension strategies, such as drawing inferences and identifying story structure elements. The recommended children's low comprehension scores at post-tutoring also indicate that they may profit from their first-grade classroom

teachers providing instruction that emphasizes these same aspects of reading comprehension processing. Additionally, first-grade teachers may enhance recommended children's reading comprehension processing by (a) modeling their own thinking processes that they employ to construct meaning of texts, (b) concurrently enhancing their listening comprehension processing, and (c) using direct explanation and collaborative discussion instructional techniques (National Reading Panel, 2000).

#### *Recommendations for Kindergarten Instruction*

Recommended children's pre-tutoring performance compared to discontinued children's pre-tutoring performance on phonological awareness and orthographic knowledge makes a case for kindergarten instruction to emphasize phonological awareness and orthographic knowledge. Prior to tutoring, recommended children demonstrated less rhyme awareness, graphophonemic awareness with respect to beginning and ending sounds, and orthographic acceptability knowledge than discontinued children at statistically significant levels (see Tables 23, 26, 28, 34). However, although the recommended children performed at a level below discontinued children, they displayed some rhyme awareness, graphophonemic awareness with respect to beginning and ending sounds, and orthographic acceptability knowledge.

Based on these results, recommended children may benefit from their kindergarten teachers continuing to foster their ability to detect rhyme by providing instruction and practice in producing and detecting rhyme during activities that involve rhyming games, texts, and songs, and the reading or writing of unfamiliar words by analogy. Recommended children may also profit from their kindergarten teachers continuing to foster their graphophonemic awareness by developing their ability to

articulate words, recognize beginning and ending phonemes in these words, match these phonemes to graphemes, and write the corresponding graphemes during text writing activities. Furthermore, recommended children may benefit from their kindergarten teachers continuing to enhance their ability to write words with acceptable English letter sequences by (a) creating classroom displays of words that they refer to during text writing activities and (b) providing opportunities to complete word sorts in which they group words by letter sequences.

Recommended children's low pre-tutoring performance that paralleled discontinued children's low pre-tutoring performance on phonological awareness and orthographic knowledge also makes a case for kindergarten instruction to emphasize phonological awareness and orthographic knowledge. Prior to tutoring, recommended and discontinued children produced low and nearly equivalent scores on (a) combined syllable, onset and rime, and phonemic awareness; (b) graphophonemic awareness with respect to middle sounds, (c) spelling knowledge; and (d) orthographic pattern knowledge (see Tables 25, 27, 31, 32).

Based on these results, recommended children may benefit from their kindergarten teachers providing instruction that fosters their ability to blend isolated sound segments in speech into recognizable words during word analysis activities and/or text reading activities. Recommended children may also profit from their kindergarten teachers developing their ability to articulate words, recognize middle phonemes in these words, match these phonemes to graphemes, and write the corresponding graphemes during text writing activities. Furthermore, recommended children may profit from their kindergarten teachers providing explicit instruction in letter patterns during text reading,

writing, and word analysis activities. These children may also benefit from their kindergarten teachers creating classroom displays of words that they refer to during text writing activities and providing opportunities to complete word sorts in which they group words by rimes and letter patterns.

### *Summary*

The results of this study indicate that continued instruction is in order for recommended children in phonological skeletal structure awareness and graphophonemic awareness due to their substantial gains from pre- to post-tutoring. This continued instruction applies to Reading Recovery tutoring and first-grade classroom instruction. Yet the results of this study also suggest that modified instruction is in order for recommended children in rhyme awareness; combined syllable, onset and rime, and phonemic awareness; spelling knowledge; orthographic pattern knowledge; and orthographic acceptability knowledge due to their small to minimal gains from pre- to post-tutoring. Furthermore, the results of this study indicate that modified instruction in oral reading processing and reading comprehension processing is in order for recommended children due to their low post-tutoring performance. This modified instruction extends to Reading Recovery tutoring and first-grade classroom instruction. The recommended children's pre-tutoring performance makes a case for kindergarten classroom instruction to continue to emphasize rhyme awareness, graphophonemic awareness with respect to beginning and ending sounds, and orthographic acceptability knowledge. However, the recommended children's pre-tutoring performance also makes a case for kindergarten instruction to address combined syllable, onset and rime, and phonemic awareness; graphophonemic awareness with respect to middle sounds, spelling

knowledge; and orthographic pattern knowledge. Overall, recommended children may profit from their Reading Recovery and classroom teachers assessing the various aspects of their early literacy development and providing appropriate and targeted instruction that prioritizes their instructional needs.

### *Recommendations for Research*

Reading Recovery teachers use the Reading Recovery teaching procedures to guide their individually designed diagnostic instruction. Research designed to determine if modifications to Clay's existing teaching procedures, such as making phonics instruction more explicit and systematic, improve recommended children's responsiveness to tutoring may prove beneficial. Also, research designed to determine if the development of new procedures, such as adding teaching procedures that enhance reading comprehension processing, improve recommended children's responsiveness to tutoring may prove worthwhile.

Instructional interactions between Reading Recovery teachers and their children, characterized by teachers' prompts and demonstrations and children's responses, influence children's pre- to post-tutoring progress in phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing. Research directed toward examining if certain instructional interactions, such a teacher-student dialogue, are more effective than others in improving and/or altering recommended children's responsiveness to tutoring may prove fruitful.

Recommended children continued to struggle in their early literacy development when they complete their tutoring approximately mid-way through first-grade. These children may profit from additional supplementary tutoring in first-grade that employs

different lesson components and teaching procedures from those used in Reading Recovery. Other supplementary tutoring programs with different lesson components and teaching procedures have demonstrated their effectiveness. For example, Reading Rescue, a tutoring program that resembles Reading Recovery but provides more systematic phonics instruction, accelerates the early literacy development of first-grade children (Dreyer, Ehri, & Flugman, 2005). Research designed to investigate if additional supplemental tutoring that employs different lesson components and teaching procedures from those used in Reading Recovery advances recommended children's early literacy development in first-grade may prove worthwhile.

Recommended children receive daily, 30-minute tutoring lessons for approximately 20 weeks. Although these children do not respond well to their tutoring, they make gains in their early literacy development from pre- to post-tutoring. Perhaps recommended children require more tutoring to respond well to it and meet the criteria for successful performance. Research designed to investigate if extended tutoring time, such as an increase in the number of weeks, lessons per day, and/or minutes per lesson, improves recommended children's tutoring responsiveness may prove beneficial.

Phillips and Smith (1997) found that the majority of Reading Recovery children who failed to meet the criteria for successful performance continued to make early literacy progress when provided extended tutoring in the weeks following their official completion of Reading Recovery tutoring. However, these researchers (a) conducted their study on small sample of children in New Zealand and (b) investigated extended tutoring only in terms of the number of weeks. Future investigations should (a) include a larger sample of recommended children in the United States and (b) examine extended tutoring

in relation to an increase in the number of lessons per day and/or the number of minutes per lesson, in addition to an increase in the number of weeks.

Beyond instructional factors, questions remain concerning the many other child-based factors, as well as family- and/or community-based factors, that may adversely influence recommended children's tutoring responsiveness. Research designed to investigate the many other child-based factors may contribute to understanding the early literacy development of recommended children. For example, limited background knowledge, lack of proficiency in the forms and functions of language, limited English proficiency, lexical storage and retrieval problems, and/or weak memory capacity may hinder the recommended children's tutoring responsiveness. Furthermore, research designed to examine family- and/or community-based factors may also contribute to understanding the early literacy development of recommended children. For example, the presence and use of books in the children's homes, as well as the availability and use of literacy-related resources in the children's communities, may promote or hinder the recommended children's tutoring responsiveness.

### *Conclusion*

The current study compared recommended Reading Recovery children to discontinued Reading Recovery children on (a) phonological awareness and orthographic knowledge prior to and following tutoring and (b) oral reading processing and reading comprehension processing following tutoring. Results corresponding to overall phonological awareness and overall orthographic knowledge in the form of composites indicated that recommended and discontinued children improved in parallel over time with the recommended children performing at a level below the discontinued children at



pre- and post-tutoring. Investigations into the components that formulated the overall phonological awareness composite revealed that recommended children performed at a level below discontinued children on (a) rhyme awareness at pre-tutoring, (b) phonological skeletal structure awareness at pre- and post-tutoring, and (c) graphophonemic awareness with respect to beginning sounds at pre- and post-tutoring and ending sounds at pre-tutoring. Similarly, investigations into the components that formulated the orthographic knowledge composite revealed that recommended children performed at a level below discontinued children on spelling knowledge at post-tutoring and orthographic acceptability knowledge at pre-tutoring. Inquiry into oral reading processing indicated that recommended children read with less accuracy, with more overall errors, with more substitutions, at a slower rate, and with less fluency than discontinued children at post-tutoring. Interestingly, recommended children performed comparable to discontinued children on reading comprehension processing at post-tutoring.

The results from this study contributes to the understanding of recommended children's early literacy development by explaining how their phonological awareness, orthographic knowledge, oral reading processing, and reading comprehension processing relate to an interactive model of early literacy development. The results make apparent recommendations for Reading Recovery tutoring and classroom instruction. Furthermore, the results illuminate the need for future research in the early literacy development of recommended children.

## Appendix A

### Pilot Study 1 and 2

#### *Phonological Awareness Tasks: Task Difficulty and Administration Issues*

##### *Introduction*

Two pilot studies examined the difficulty of a plethora of phonological awareness tasks. The findings informed the selection of phonological awareness tasks in a dissertation study on recommended Reading Recovery children's early literacy development. The first pilot study assessed kindergarten children's performance on modified phonological awareness tasks at the end of their school year. The second pilot study assessed Reading Recovery children's performance on various phonological awareness tasks upon their entry into and exit from tutoring. In addition to task difficulty, the two pilot studies examined task administration issues, including the length of the assessment sessions and the utility of practice items with corrective feedback.

##### *Rationale*

Researchers have employed numerous phonological awareness tasks, such as blending or segmenting tasks, to assess the many dimensions of phonological awareness (e.g., National Reading Panel, 2000). However, researchers have often failed to examine the difficulty of their selected phonological awareness tasks relative to the children participating in their studies (Lewkowicz, 1980). That is, researchers have often overlooked task difficulty.

Because of their limited awareness of the phonological structure of spoken words, children who struggle in their early literacy development do not perform as well as children who progress with ease in their early literacy development on various

phonological awareness tasks (Whitehurst & Lonigan, 2001). Thus, the identification of tasks that are sensitive to children's varying levels of phonological awareness becomes a critical avenue to explore prior to launching a study on children who struggle in their early literacy development.

### *Purpose*

The primary purpose of these two pilot studies was to investigate the task difficulty of several phonological awareness tasks to ensure appropriate task selection for a dissertation study that compared recommended and discontinued Reading Recovery children's phonological awareness. A secondary purpose of these pilot studies was to examine task administration issues, specifically the length of the assessment sessions and the utility of practice items with corrective feedback. The following research questions guided the two pilot studies:

### *Pilot Study 1 Research Questions*

- (a) How difficult are modified phonological awareness tasks for kindergarten children to complete at the end of their school year?
- (b) How many minutes do kindergarten children remain engaged and focused during an assessment session?
- (c) How well do practice items with corrective feedback communicate task understanding to kindergarten children?

### *Pilot Study 2 Research Questions*

- (a) How difficult are phonological awareness tasks for Reading Recovery children to complete at pre- and post-tutoring?
- (b) How many minutes do Reading Recovery children require to complete

five phonological tasks at pre- and post-tutoring?

*Literature Review*

In this section, I review studies that examined factors that may increase the difficulty of phonological awareness tasks. Specifically, I examined the factors of task hierarchy, linguistic complexity, and cognitive requirements.

Troia (1999) argued that knowledge of the difficulty levels of phonological awareness tasks and sensitivity to children's developing phonological awareness are critical factors to consider when investigating children's phonological awareness. According to Troia, when researchers consider the difficulty of phonological awareness tasks relative to their sample of children, they avoid ceiling and floor effects, which restrict the range of variance and limit the effectiveness of their statistical analyses. Troia asserted the importance of selecting tasks that are neither too easy nor too difficult for a given sample of children.

Adams (1990) presented the following hierarchy of phonological awareness tasks, ranging from most to least difficult: phoneme manipulation tasks, phoneme segmentation tasks, blending and syllable-splitting tasks, oddity tasks, and nursery rhyming tasks. Likewise, Stanovich (1992) coined the terms 'deep phonological sensitivity' and 'shallow phonological sensitivity' to describe this phonological awareness hierarchy.

Several studies have also investigated the hierarchy of phonological awareness tasks in kindergarten (e.g., Yopp, 1988) and first-grade children (e.g., Stahl & Murray, 1994). In an analysis of kindergarten children's performance on various phonological awareness tasks, Stanovich, Cunningham, and Crammer (1984) found that the strip-initial-consonant task was the most difficult task to complete with the supply-initial-

consonant, the final-consonant-same, and the final-consonant-different tasks also being relatively difficult to complete. The researchers also found that the initial consonant, initial-consonant-same, initial-consonant-different, and initial-consonant-not-same tasks less challenging for the children to complete. Furthermore, Stanovich et al. found that rhyme choice, rhyme supply, and substitute-initial-consonant tasks were the easiest to complete.

Yopp (1988) also researched this notion of task difficulty and formulated a hierarchy of phonological awareness tasks. Based on the performance of kindergarten children, she ranked the following phonological awareness tasks from most to least difficult: phoneme deletion, phoneme segmentation, phoneme counting, sound isolation, word-to-word matching, phoneme blending, auditory discrimination, and rhyme discrimination tasks.

More recently, Stahl and Murray (1994) assessed the difficulty of phonological awareness tasks for kindergarten and first-grade children. These researchers ranked the following tasks from most to least difficult: phonological segmentation, phoneme deletion, phoneme blending, and phoneme isolation. The phoneme segmentation task required the children to segment words into onsets and rimes and to segment words into individual phonemes tasks. However, Stahl and Murray used the phoneme segmentation scores, not the onset and rime segmentation scores, to determine the task's position in the hierarchy.

Most recently, Chafouleas, Lewandowski, Smith, and Blachman (1997) conducted a study on task difficulty with kindergarten, first-, and second-grade children. Chafouleas et al.'s findings revealed a hierarchy in phonological awareness tasks similar

to the findings of Stanovich et al. (1984), Yopp (1988), and Stahl and Murray (1994). Chafouleas et al.'s hierarchy, arranged from most to least difficult, included a phoneme manipulation task, a phoneme segmentation task, a phoneme blending task, an alliteration task, and a rhyme task.

In addition to task hierarchy, linguistic complexity and cognitive requirements contribute to the difficulty of phonological awareness tasks. Based on her research on phonological awareness, Backman (1983) concluded that "Tasks which on the surface appear to be measuring the same phenomenon may in fact require different degrees of linguistic awareness [complexity], or may differ in their cognitive requirements" (p. 476).

Stahl and Murray (1994) assessed the linguistic complexity of various phonological awareness tasks, specifically onset and rime within a syllable. They found that onsets and rimes were the easiest for children to manipulate, followed by vowels and consonants in rimes, followed by phonemes in cluster onsets, followed by phonemes in cluster consonants in rimes. Additionally, Stahl and Murray researched the effects of linguistic components, specifically nasal blends (e.g., *nk*) and liquid blends (e.g., *ld*). They found that tasks that contained these linguistic components were more challenging than tasks that did not contain these components because the children tended to treat the blends as individual phonemes rather than separate phonemes.

Treiman (1992) also examined the importance of linguistic complexity in phonological awareness tasks. She discovered that tasks containing syllabic units were easier for children to complete than tasks containing phonemic units. Furthermore, Treiman discovered that intrasyllabic units, such as onsets and rimes, were as central to linguistic complexity as syllabic and phonemic units.

Tunmer and Hoover (1992) posited that various cognitive requirements impact children's performance on phonological awareness tasks. Such cognitive requirements or "spinoff skills" include the ability "to maintain and operate on verbal material in working memory, to generate orthographic images, and to apply phoneme-grapheme correspondence rules" (p. 192). Stanovich et al. (1984) also acknowledged the employment of cognitive requirements, such as "short-term memory, stimulus comparison, and processing of task instructions", are necessary for the successful completion of particular phonological awareness tasks (p. 176).

In summary, these studies have investigated the difficulty of various phonological awareness tasks. That is, these studies have highlighted the importance of considering task hierarchy, linguistic complexity, and extraneous cognitive requirements when selecting phonological awareness tasks to assess children's phonological awareness.

#### *Methodology: Pilot Study 1*

In this section, I first describe the sample of children who participated in the first pilot study. Next, I describe the phonological awareness tasks and detail the task modifications. Then, I outline the procedures.

#### *Kindergarten Children*

This first pilot study consisted of 40 kindergarten children who attended two suburban schools with half-day kindergarten programs in a small school district. I selected these children from a larger pool of 83 children who returned their informed consent forms (see Appendix A). All of the children were Caucasian and came from lower-middle class families residing in a Mid-Atlantic state. The sample consisted of 20 females and 20 males. I assessed the children in May of their kindergarten year and

considered them to be a representative sample of first-grade children at the beginning of first-grade.

### *Modified Phonological Awareness Tasks*

In this section, I describe the modified rhyme awareness and phonemic awareness tasks that I administered to the children. For each task, I describe the original task, followed by the modifications I made to each task. I modified these tasks by (a) selecting a reduced number of task and/or practice items, (b) developing practice items, and/or (c) adding picture support. Each modified rhyme and phonemic awareness task contained five or seven task items, respectively with three or four practice items.

*Modified rhyme detection task.* Calfee, Chapman, and Venezky's (1972) Rhyme Detection assessed the children's ability to detect rhyme in two word pairs. This task contained 20 items and three practice items with corrective feedback. Calfee et al. scored the items as correct or incorrect. They failed to report reliability, validity, and normative information. I modified this task by selecting five of the 20 items. I also include the three practice items (see Table 1).

*Modified rhyme oddity task.* Bradley and Bryant's (1983) Rhyme Oddity task assessed the children's ability to select the 'odd one out' from among four word choices. This task consisted of first, middle, and end sound categories, based on the location of the 'odd one out' within each word. Bradley and Bryant randomized the position of the 'odd one out' across each word set. The task contained 30 items with 10 items in each sound category. The task also contained six practice items with corrective feedback, two practice items corresponding to each sound category. Bradley and Bryant scored the items as correct or incorrect. They failed to provide reliability, validity, and normative



Table 1

*Pilot Study: Phonological Awareness Tasks Employed in Pilot Study 1 and 2*

Pilot Study 1 Modified Task	Citation	Pilot Study 2 Complete Task	Citation
Modified Rhyme Detection	Calfee et al., 1972	Rhyme Detection <sup>ab</sup>	Muter et al., 1997
Modified Rhyme Oddity	Bradley & Bryant, 1983	Rhyme Oddity	Dodd et al., 2000
Modified Rhyme Production	MacLean et al., 1987 Stanovich et al., 1984		
Modified Auditory Blending	Roswell & Chall, 1997	Blending Words <sup>ab</sup>	Wagner et al., 1999
Modified Sound Matching <sup>a</sup>	Wagner et al., 1999	Sound Matching <sup>a</sup>	Wagner et al., 1999
Modified Phoneme Segmentation	Yopp, 1988	Sentence Writing and Spelling <sup>b</sup>	DeFord, 2000
Modified Phoneme Deletion	Bruce, 1964		

<sup>a</sup> Indicates a norm-referenced task. <sup>b</sup> Indicates the tasks selected for the dissertation study.

information. I modified this task by selecting five of the 30 items, two items from the middle sound category and three items from the end sound category. I omitted items from the first sound category because they focused on the alliteration or the onset, rather than the rhyme or the rime. I also modified this task by including four of the six practice items, two practice items corresponding to the middle category and two practice items corresponding to the end category. I further modified this task by providing picture support to reduce the short-term memory demands placed on the children (see Table 1).

*Modified rhyme production task.* MacLean, Bryant, and Bradley's (1987) Rhyme Production task measured the children's ability to produce rhyming words when provided a starting word. They scored the items as correct or incorrect. MacLean et al. failed to provide reliability, validity, and normative information. I modified this task by selecting five of the 20 items. I also modified this task by adding practice items from Stanovich, Cunningham and Crammer's (1984) Rhyme Production task. Stanovich et al.'s task mirrored MacLean et al.'s task. Because Stanovich et al. included only one practice item in their task, I selected this practice item and the first two task items as practice items (see Table 1).

*Modified auditory blending task.* The Roswell-Chall Auditory Blending task (Roswell & Chall, 1997) assessed the children's ability to blend two phonemes into a word, an onset and rime into a word, or three or more phonemes into a word. The task contained 30 items with 10 items in each blending skill and three practice items with corrective feedback. Roswell and Chall scored the items as correct or incorrect. Roswell and Chall, as well as Yopp (1988), reported reliability and validity information. I modified this task by selecting seven of the 30 items with three items in the two phoneme

category, two items in the onset and rime category, and two items in the three or more phoneme category. I included the three practice items (see Table 1).

*Modified sound matching task.* Wagner, Torgesen, and Rashotte's (1999) Sound Matching task assessed the children's ability to match sounds by having them point to the picture that started or ended with the same sound as the word spoken by the examiner. This task contained 20 items, 10 beginning sound matching items and 10 ending sound matching items. It also contained six practice items with corrective feedback, three beginning sound matching items and three ending sound matching items. Because this task required children to use their verbal memory, Wagner et al. provided pictures for each word set to reduce memory demands. They scored the items as correct or incorrect. Wagner et al. reported reliability, validity, and normative information. I modified Wagner et al.'s Sound Matching task by selecting seven of the 20 items, four beginning sound matching items and three ending sound matching items. I also modified this task by selecting three of the six practice items, two beginning sound matching items and one ending sound matching item (see Table 1).

*Modified phoneme segmentation task.* The Yopp-Singer Phoneme Segmentation task (Yopp, 1988) assessed the children's ability to sequentially recognize and articulate individual phonemes in a spoken word. This task contained 22 items and four practice items. Yopp provided corrective feedback on the task and practice items. She scored the items as correct or incorrect. She also reported reliability and validity information. I modified this task by including seven of the 22 items. I included the four practice items (see Table 1).

*Modified phoneme deletion task.* The Bruce Phoneme Deletion task (Bruce, 1964)

assessed children's ability to delete a phoneme from the beginning, middle, or end of a word. This test contained 30 items and no practice items. Bruce scored the items as correct or incorrect. Yopp reported reliability and validity information. I modified this task by including seven of the 30 items. I further modified this task by including three practice items with corrective feedback from the Yopp-Singer Phoneme Segmentation task (Yopp, 1988) to serve as practice items (see Table 1).

### *Procedures*

Two weeks prior to the end of their kindergarten school year, I individually administered the following modified rhyme awareness tasks to the kindergarten children: (a) a modified version of Calfee et al.'s (1972) Rhyme Detection task, (b) a modified version of Bradley and Bryant's (1983) Rhyme Oddity task, and (c) a modified version of MacLean et al.'s (1987) and Stanovich et al.'s (1984) Rhyme Production task (see Table 1). I also administered the following modified phonemic awareness tasks: (a) a modified version of Wagner et al.'s (1999) Sound Matching task, (b) a modified version of Yopp's (1988) Yopp-Singer Phoneme Segmentation task, (c) a modified version of Bruce's (1964) Phoneme Deletion task, and (d) a modified version of Roswell and Chall's (1997) Auditory Blending task (see Table 1). First, I administered the modified rhyme awareness tasks, followed by the modified phonemic awareness tasks. Each child completed these tasks in approximately 15 minutes.

Due to scheduling conflicts and time constraints, I did not administer the modified Rhyme Oddity task and the modified Phoneme Segmentation task to some of the children. Thus, only 24 of the 40 kindergarten children completed these modified tasks.

## *Results*

In this section, I present the results of the first pilot study. First, I report the descriptive statistics. Next, I report the results corresponding to the number of minutes the children required to complete the phonological awareness tasks. Then, I report the informal observations of children completing practice items with corrective feedback.

### *Phonological Awareness Tasks*

Table 2 displays the kindergarten children's mean scores and standard deviations. The mean scores corresponding to the modified Rhyme Detection, Rhyme Oddity, and the Rhyme Production tasks indicated that on average the children answered three to four of the five items correctly. The mean scores corresponding to the modified Sound Matching and Auditory Blending tasks suggested that on average the children answered four of the seven items accurately. Furthermore, the mean scores corresponding to the modified Phoneme Segmentation and Phoneme Deletion tasks indicated that on average the children answered one to two of the seven items correctly.

### *Administration Issues*

Informal observations revealed that the kindergarten children completed the modified phonological awareness tasks in approximately 20 minutes. Informal observations also indicated that the children increased their understanding of the modified phonological awareness tasks after providing incorrect responses on practice items and receiving corrective feedback.

## *Discussion*

In this section, I first discuss task difficulty in light of the results. Next, I discuss task selection for a second pilot study. Then, I discuss the task administration issues of

assessment session length and practice items with corrective feedback.

Table 2

*Pilot Study: Means and Standard Deviations for Kindergarten Children on the Modified Phonological Awareness Tasks in Pilot Study 1*

Task	Range	Sample Size	M	(SD)
Modified Rhyme Detection	0–5	40	3.75	(1.50)
Modified Rhyme Oddity	0-5	24	3.00	(1.41)
Modified Rhyme Production	0-5	40	3.50	(1.92)
Modified Auditory Blending	0-7	40	4.10	(2.47)
Modified Sound Matching	0-7	40	4.45	(1.93)
Modified Phoneme Segmentation	0-7	24	1.54	(1.61)
Modified Phoneme Deletion	0-7	40	1.48	(1.45)

*Note.* M = Mean; SD = Standard Deviation.

### *Task Difficulty*

I established criteria to determine whether the modified phonological awareness tasks were too easy or too difficult for the kindergarten children to complete. I considered the modified rhyme awareness tasks (range 0-5) with mean scores of (a) 4.0 or above too easy and (b) 1.0 or below too difficult. Additionally, I considered the modified phonemic awareness tasks (range 0-7) with mean scores of (a) 5.5 and above too easy and (b) 1.5 or

below too difficult. According to my criteria and the children's mean scores, the modified Rhyme Detection, Rhyme Oddity, Rhyme Production, Auditory Blending, and Sound Matching tasks were neither too easy nor too difficult. That is, these tasks were appropriate and sensitive to the children's developing phonological awareness. Conversely, the modified Phoneme Segmentation and Phoneme Deletion tasks were too difficult.

#### *Task Selection for Pilot Study 2*

Although the children's mean scores indicated that the modified Rhyme Detection, Rhyme Oddity, and Rhyme Production tasks are sensitive and appropriate tasks, I did not select them because of the researchers failure to report reliability and validity information corresponding to the tasks and the tasks' lack of normative information. I conducted a second pilot study in which I administered replacement phonological awareness tasks that were reliable and valid and possessed normative information.

I replaced the modified version of Calfee et al.'s (1972) Rhyme Detection task with Muter, Hulme, and Snowling's (1997) Rhyme Detection task (see Table 1). I also replaced the modified version of Bradley and Bryant's (1983) Rhyme Oddity task with Dodd, Crosbie, McIntosh, Teitzel, and Ozanne's (2000) Rhyme Oddity task (see Table 1). Muter et al.'s Rhyme Detection task and Dodd et al.'s Rhyme Oddity task assessed the same underlying constructs as Calfee et al.'s Rhyme Detection task and Bradley and Bryant's Rhyme Oddity tasks, respectively. Finally, I omitted the modified version of MacLean, Bryant, and Bradley's (1987) and Stanovich et al.'s (1984) Rhyme Production task from the second pilot study without supplying a replacement task (see Table 1). I did

not provide a replacement task because the administration of Muter et al.'s Rhyme Detection task and Dodd et al.'s Rhyme Oddity task seemed adequate to assess the children's rhyme awareness.

Although the children's mean scores indicated that the modified Auditory Blending task was a sensitive and appropriate task, I did not include the original version of this task in the second pilot because Roswell and Chall did not provide a standardized way of articulating the isolated sound segments that the children blended into words. I replaced this task with Wagner et al.'s (1999) Blending Words task, a reliable and valid task with normative information that provided an audiotape of the isolated sound segments. This audiotape ensured that the sound segments were presented to the children in a consistent manner. Wagner et al.'s Blending Words task assessed the same construct as the modified Auditory Blending task (see Table 1).

The children's mean scores indicated that the modified Sound Matching task was a sensitive and appropriate task. Wagner et al.'s (1999) original version of this task was reliable and valid with normative information. I included the original version of this task in my second pilot study (see Table 1).

The children's mean scores suggested that the modified Phoneme Segmentation task was too difficult for the children. Stahl and Murray (1994) and van Bon and Duighuise (1995) found oral segmentation tasks more difficult for children to complete than written segmentation tasks due to the demands placed upon memory. Based on the finding from this pilot study and on Stahl and Murray's and van Bon and Duighuise's findings, I did not include the modified Phoneme Segmentation task or the original version of this task in my second pilot study. I replaced the modified Phoneme



Segmentation test with DeFord's (2000) Sentence Writing and Spelling task (see Table 1). DeFord's task assesses the same underlying construct as the modified Phoneme Segmentation task. However, the children wrote the graphemes corresponding to the phonemes, rather than articulating them.

Finally, the children's mean scores suggested that the modified Phoneme Deletion task was too difficult for the children. According to Goswami and Bryant (1990), Bruce (1964) also found that the original version of this task was too difficult for kindergarten and first-grade children, as evidenced by their mean scores. Based on the finding from this pilot study and on Bruce's finding, I did not include the modified Phoneme Deletion task or the original version of this task in my second pilot study (see Table 1).

#### *Task Administration*

Informal observations indicated that the children seemed to focus their attention on the phonological awareness tasks during the 15 minutes assessment sessions. They did not appear tired, distracted, or unable to concentrate. I used this information to justify my decision to extend the assessment sessions in the second pilot study and the dissertation study to 30 minutes.

Informal observation also indicated that the children appeared to listen to the verbal directions. The children stated that they understood how to complete the various phonological awareness tasks. Yet some children restated their understanding of the tasks once they answered practice items incorrectly and subsequently received corrective feedback from the examiner. That is, the children seemed to produce incorrect responses because they failed to understand the directions. The practice items with correct feedback appeared to clarify their misunderstanding. I used this information to justify my decision

to include practice items with corrective feedback on the phonological awareness tasks in my second pilot study and my dissertation study.

### *Methodology: Pilot Study 2*

In this section, I first describe the sample of children who participated in this second pilot study. Next, I describe the phonological awareness tasks. Then, I outline the procedures.

#### *Reading Recovery Children*

This second pilot study consisted of 29 Reading Recovery children in first-grade who attended five schools in a large school district in a Mid-Atlantic state. These Reading Recovery children received daily tutoring from trained Reading Recovery teachers for approximately 20 weeks. The Reading Recovery teachers in their respective schools selected the first-grade children to receive Reading Recovery tutoring. I selected the Reading Recovery children to participate in this second pilot study from a larger pool of 61 Reading Recovery children who returned their informed consent forms (see Appendix B). The mean age of the Reading Recovery children at pre-and post-tutoring was 6 years 3 months and 6 years 8 months, respectively. The sample of Reading Recovery children consisted of 13 girls and 20 boys. All of the children spoke English as their first language with the exception of one child. The sample consisted of 19 (58%) African-American children, 11 (33%) Caucasian children, two (6%) Asian children, and one (3%) Hispanic child.

#### *Phonological Awareness Tasks*

The results from the first pilot study, assisted in my selection of the phonological awareness tasks for this second pilot study. In this section, I describe the phonological

awareness tasks that I administered to the Reading Recovery children. These phonological awareness tasks were independent of the Reading Recovery program. However, the Sentence Writing and Spelling task was similar to Clay's (2002) Hearing and Recording Sounds in Words, one of six subtests of *An Observation Survey of Early Literacy Achievement* (Clay, 2002).

*Rhyme Detection task.* The Rhyme Detection task (Muter et al., 1997), one of six subtests in the Phonological Abilities Test, assessed the children's ability to select the word that rhymes with or sounds like the target word from among three choices. Picture support reduced the demands placed upon short-term memory. This task contained 10 items and three practice items. Muter et al. provided corrective feedback on the practice items and on the first four task items. They scored the items as correct or incorrect. They also provided reliability, validity, and normative information (see Table 1).

*Rhyme Oddity task.* The Rhyme Oddity task (Dodd et al., 2000), one of six subtests in the Preschool and Primary Inventory of Phonological Awareness, assessed the children's ability to select the non-rhyming word or the 'odd one out' from among three word choices. Picture support reduced the demands on short-term memory. This task contained 12 items, an explanation of rhyme, and two practice items with corrective feedback. Dodd et al. scored the items as either correct or incorrect. They also provided reliability, validity, and normative information (see Table 1).

*Sound Matching task.* The Sound Matching task (Wagner et al., 1999), one of eight subtests in Comprehensive Test of Phonological Processing (CTOPP), measured the children's ability to match sounds by having them point to the picture that started or ended with the same sound as the word spoken by the examiner. This task contained 20

items, 10 beginning sound matching items and 10 ending sound matching items. It also contained six practice items, three beginning sound matching items and three ending sound matching items. Wagner et al. provided corrective feedback on the practice items and on the first three task items. Picture support reduced the demands placed upon short-term memory. Wagner et al. scored the items as correct or incorrect. They also provided reliability, validity, and normative information (see Table 1).

*Blending Words task.* The Blending Words task (Wagner et al., 1999), one of eight subtests in the CTOPP, assessed the children's ability to blend isolated syllables, onset and rimes, or phonemes in speech into a word. Wagner et al. provided an audiotape of the isolated sound segments. This task contained 20 items and three practice items. Wagner et al. provided corrective feedback on the practice items and on the first four task items. They scored the items as correct or incorrect. Furthermore, they provided reliability, validity, and normative information (see Table 1).

*Sentence Writing and Spelling task.* The Sentence Writing and Spelling task (DeFord, 2000) assessed the children's ability to hear phonemes in a dictated word, match the phonemes to the corresponding graphemes, and record the graphemes. First, the examiner read the complete dictation to the child. Then, the examiner read each word of the dictation, one at a time, while the child transcribed each word. DeFord scored this 50-point task by (a) evaluating the child's attempts in accordance with acceptable phoneme substitutions and (b) marking each recorded grapheme as correct or incorrect. She failed to provide reliability, validity, and normative information (see Table 1).

### *Procedures*

I individually administered five phonological awareness tasks to the Reading

Recovery children at the beginning of their tutoring in September and upon the completion of their tutoring in February. First, I administered the rhyme awareness tasks in the following order: the Rhyme Detection task (Muter et al., 1997) and the Rhyme Oddity task (Dodd et al., 2000). Then, I administered the phonemic awareness tasks in the following order: the Sentence Writing and Spelling task (DeFord, 2000), the Sound Matching task (Wagner et al., 1999), and the Blending Words task (Wagner et al., 1999). I administered the tasks in this sequence to break up the children's verbal responding by incorporating writing after the children completed the rhyme awareness tasks.

Upon the Reading Recovery children's completion of tutoring, the Reading Recovery teachers decided whether the children met the Reading Recovery program's criteria for successful completion. The teachers based their decisions on several factors, including the children's performance on *An Observation Survey of Early Literacy Achievement* (Clay, 2002) and the children's development of a self-extending system. If the teachers decided that the children met criterion performance, they assigned the children to the end-of-program status category of discontinued. Conversely, if the teachers decided that the children failed to meet criterion performance, they assigned the children to the end-of-program status category of recommended. From among the 29 Reading Recovery children, the teachers classified 15 of them as discontinued and 14 of them as recommended. I examined the discontinued and the recommended children's performance on these phonological awareness tasks to inform my selection of tasks for my dissertation study.

### *Results*

In this section, I present the results of the second pilot study corresponding to all

the Reading Recovery children, the recommended Reading Recovery children, and the discontinued Reading Recovery children. First, I report descriptive statistics corresponding to the five phonological awareness tasks. Then, I report the number of minutes the children required to complete the phonological awareness tasks.

#### *Phonological Awareness Tasks*

Table 3 displays the children's mean scores and standard deviations corresponding to the rhyme awareness tasks. The pre- and post-tutoring mean scores on the Rhyme Detection task (Muter et al., 1997) indicated that all the Reading Recovery children, as well as the subgroups of recommended and discontinued children, possessed a comparable ability to detect rhyme. On average, the children answered seven of the 10 items correctly at pre-tutoring and eight or nine of the ten items correctly at post-tutoring (see Table 3). Furthermore, the children's pre- and post-tutoring mean scores on the Rhyme Oddity task (Dodd et al., 2000) indicated that all three groups possessed a comparable ability to select the non-rhyming word (i.e., the 'odd one out'). On average, the children answered six or seven of the 12 items accurately (see Table 3).

Table 3 also displays the children's mean scores and standard deviations corresponding to the phonemic awareness tasks. The children's pre- and post-tutoring mean scores on the Sound Matching task (Wagner et al., 1997) suggested that all the Reading Recovery children, as well as the two subgroups of children, possessed a comparable ability to match sounds by selecting pictures that started or ended with the same sound as the words spoken by the task examiner. On average, the children answered seven to 10 out of 20 items correctly at pre-tutoring and 14 to 15 out of 20 items correctly at post-tutoring (see Table 3).

Table 3

*Pilot Study: Means and Standard Deviations for All Reading Recovery Children, Recommended Children, and Discontinued Children on the Phonological Awareness Tasks in Pilot Study 2 at Pre- and Post-tutoring*

Task	Range	All Reading Recovery Children <sup>a</sup>				Recommended Children <sup>b</sup>				Discontinued Children <sup>c</sup>			
		Pre		Post		Pre		Post		Pre		Post	
		M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
RD	0-10	7.14	(2.98)	8.83	(1.51)	7.29	(3.20)	8.36	(1.91)	7.00	(2.88)	9.27	(.88)
RO	0-12	6.59	(3.03)	6.55	(2.82)	7.00	(3.11)	6.21	(3.14)	6.20	(3.00)	6.87	(2.56)
SM	0-20	8.62	(2.86)	14.48	(3.24)	7.43	(2.74)	13.57	(3.74)	9.73	(2.58)	15.33	(2.53)

Table 3 continued

*Pilot Study: Means and Standard Deviations for All Reading Recovery Children, Recommended Children, and Discontinued Children on the Phonological Awareness Tasks in Pilot Study 2 at Pre- and Post-tutoring*

Task	Range	All Reading Recovery Children <sup>a</sup>				Recommended Children <sup>b</sup>				Discontinued Children <sup>c</sup>			
		Pre		Post		Pre		Post		Pre		Post	
		M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
BW	0-20	3.48	(2.54)	5.79	(2.08)	2.57	(2.21)	5.29	(1.59)	4.33	(2.61)	6.27	(2.40)
SWS	0-50	22.10	(8.74)	41.55	(5.95)	18.57	(9.60)	37.93	(6.11)	25.40	(6.57)	44.93	(3.28)

*Note.* RD = Rhyme Detection; RO = Rhyme Oddity; SWS = Sentence Writing and Spelling; SM = Sound Matching; BW = Blending Words; Pre = Pre-tutoring; Post = Post-tutoring; M = Mean; SD = Standard Deviation. <sup>a</sup>Indicates sample size (n = 29).

<sup>b</sup>Indicates sample size (n = 14). <sup>c</sup>Indicates sample size (n = 15).



Likewise, the children's pre- and post-tutoring mean scores on the Blending Words task (Wagner et al., 1997) indicated that all the Reading Recovery children, as well as the two subgroups, possessed a similar ability to blend isolated syllables, onsets and rimes, and phonemes in speech into words. On average, the children answered three to four out of 20 items correctly at pre-tutoring and five to six out of 20 items correctly at post-tutoring (see Table 3).

Similarly, the children's pre- and post-tutoring mean scores on the Sentence Writing and Spelling task (DeFord, 2000) indicated that all the Reading Recovery children, as well as the two subgroups, possessed a comparable ability to hear phonemes in dictated words, match the phonemes to the corresponding graphemes, and record the graphemes. On average, the children answered 19 to 25 out of 50 items correctly at pre-tutoring and 38 to 45 out of 50 items correctly at post-tutoring (see Table 3).

#### *Administration Issue of Assessment Session Length*

Informal observations revealed that the Reading Recovery children completed the phonological awareness tasks in approximately 20 minutes during the pre-tutoring assessment sessions and in approximately 15 minutes during the post-tutoring assessment sessions.

### *Discussion*

In this section, I discuss task difficulty and task selection for my dissertation study. Then, I discuss the task administration issue of assessment session length.

#### *Task Difficulty and Task Selection for the Dissertation Study*

Analogous to the first pilot study, I established criteria to determine whether the phonological tasks were too easy or too difficult for the recommended Reading Recovery

children. I fashioned my criteria based on the recommended children's mean score performance. I considered the recommended children's means on Muter et al.'s (1997) Rhyme Detection task (range 0-10) of (a) 8.0 or above too easy and (b) 2.0 or below too difficult. In addition, I considered the children's mean scores on Dodd et al.'s (2000) Rhyme Oddity task (range 0-12) of (a) 9.0 or above too easy and (b) 3.0 or below too difficult. I also considered the children's mean scores on Wagner et al.'s (1999) Sound Matching and Blending Words tasks (range 0-20) of (a) 15.0 or above too easy and (b) 5.0 or below too difficult. Finally, I decided that the children's mean scores on DeFord's (2000) Sentence Writing and Spelling task (range 0-50) of (a) 38.0 or above too easy and (b) 12.0 or below too difficult. In addition to these criteria, I also considered extraneous cognitive requirements, including memory demands and comparison skills.

According to my criteria and the recommended children's pre-tutoring mean score, the Rhyme Detection task (Muter et al., 1997) was neither too easy nor too difficult for the children. Although the recommended children's post-tutoring mean score hit the too easy mark, the corresponding standard deviation indicated a high degree of variability (see Table 3). Thus, I included this task in my dissertation study (see Table 1).

The recommended children's pre- and post-tutoring mean scores on the Rhyme Oddity task (Dodd et al., 2000) indicated that it was neither too easy nor too difficult. However, the children's mean scores did not increase from pre- to post-tutoring, and the corresponding standard deviations indicated a high degree of variability in the distribution of scores at pre- and post-tutoring (see Table 3). Memory demands and comparison skills may serve as a possible explanation for the mean scores to remain approximately the same from pre- to post-tutoring and the large standard deviations.

Because memory capacity and comparison skills may be cognitive requirements for successful completion of this task, I did not select it for my dissertation study.

The recommended children's pre- and post-tutoring mean scores on the Sound Matching task (Wagner et al., 1999) indicated that it was neither too easy nor too difficult. However, the standard deviation corresponding to the post-tutoring mean score indicated a high degree of variability (see Table 3). As with the Rhyme Oddity task (Dodd et al., 2000), memory capacity and comparison skills may have contributed to this large standard deviation. Because memory capacity and comparison skills appeared to be cognitive requirements for the successful completion of this task, I did not select it for my dissertation study.

According to my criteria and the recommended children's pre- and post-tutoring mean scores, the Blending Words task (Wagner et al., 1999) was too difficult. However, the ability to blend isolated syllables, onsets and rimes, and phonemes in speech into words is a fundamental early reading skill (National Reading Panel, 2000). Although this task proved to be difficult for the recommended children, I selected it for my dissertation study because of its importance to early reading.

The recommended children's pre-tutoring mean score on the Sentence Writing and Spelling task (DeFord, 2000) indicated that it was neither too easy nor too difficult, whereas the children's post-tutoring mean score hit the too easy mark. Although the standard deviation corresponding to this post-tutoring mean score indicated a high degree of variability (see Table 3), I included this task in my dissertation study.

#### *Assessment Session Length*

Informal observations indicated that the majority of the Reading Recovery

children completed the five phonological awareness tasks in approximately 12-15 minutes during the 30-minute pre-tutoring sessions and in approximately 10-12 minutes during the 30-minute post-tutoring sessions. In my dissertation study, I will reduce the number of phonological awareness tasks that the children complete from five to three tasks. I plan to use remaining minutes in the 30-minute assessment session to assess the children's oral reading processing and reading comprehension processing.

### *Conclusion*

Based on my analyses of the descriptive statistics from these pilot studies, I selected tasks for my dissertation study, theorized to be neither too easy nor too difficult for recommended Reading Recovery children to complete prior to and following their tutoring. Furthermore, based on my informal observations from these pilot studies, I gained insight into the tasks administration issues of assessment session length and utility of practice items with corrective feedback to inform my dissertation study.

## Appendix A

## Pilot Study 1: Informed Consent

May, 2001

Dear Parents or Guardians,

I am a doctoral student at the University of Maryland who is interested in learning more about first-grade reading and writing. I am conducting a research study that examines children's awareness of sounds in words, as well as their spelling knowledge in relation to their reading and writing development.

In order to learn more about children's reading and writing development, I would like to meet and work with your child during the last week of May for one 15-minute session. Word games and a writing activity will be completed during this session.

Your child's participation in this study is voluntary. You are welcome to ask questions and to withdraw your child from participation. These reading and writing activities will not be used to evaluate your child's progress in school, nor are these activities designed to advance your child's literacy development. There is no risk surrounding your child's participation. Further, all gathered information will remain confidential and your child's name will not be released, kept in a record, or documented. However, with your permission, your child's performance on these activities will be released to your child's kindergarten teacher.

I am required to obtain your permission before your child can participate in these reading and writing activities. I am also obligated to obtain your permission in order to release information regarding your child's performance to his/her classroom teacher. Please complete the form below and return it to your child's classroom teacher. If you have any questions, please feel free to contact me at (202)-544-1768 or my advisor, Dr. Marilyn Chambliss, at (301)-405-7410. Thank you for your cooperation.

Sincerely,

Judith S. Concha

---

Please check, sign, and return this form to your child's classroom teacher.

I grant permission for my child to participate in the reading and writing activities.

I grant permission for my child's performance to be shared with his/her teacher.

I do not grant permission for my child to participate in reading/writing activities.

I do not grant permission for my child's performance to be shared with his/her teacher.

---

Name of Child

---

Signature of Parent/Guardian

---

Date

## Appendix B

## Pilot Study 2: Informed Consent

September, 2001

Dear Parents or Guardians,

I am a doctoral student at the University of Maryland who is interested in learning more about first-grade reading and writing. I am conducting a research study that examines children's awareness of sounds in words, as well as their spelling knowledge in relation to their reading and writing development.

In order to learn more about children's reading and writing development, I would like to meet with your child for two 30-minute sessions. Word games and reading/writing activities will be completed during these sessions. The first session will be held in September and the second session in February.

Your child's participation in this study is voluntary. You are welcome to ask questions and to withdraw your child from participation. These reading and writing activities will not be used to evaluate your child's progress in school, nor are these activities designed to advance your child's literacy development. There is no risk surrounding your child's participation. Further, all gathered information will remain confidential and your child's name will not be released, kept in a record, or documented. However, with your permission, your child's performance on these activities will be released to your child's first-grade classroom teacher and Reading Recovery teacher.

I am required to obtain your permission before your child can participate in these reading and writing activities. I am also obligated to obtain your permission in order to release information regarding your child's performance to his/her classroom teacher and Reading Recovery teacher. Please complete the form below and return it to your child's classroom teacher. If you have any questions, please contact me at (202)-544-1768 or my advisor, Dr. Marilyn Chambliss, at (301)-405-7410. Thank you for your cooperation.

Sincerely,

Judith S. Concha

---

Please check, sign, and return this form to your child's classroom teacher.

I grant permission for my child to participate in the reading and writing activities.

I grant permission for my child's performance to be shared with his/her teachers.

I do not grant permission for my child to participate in the reading/writing activities.

I do not grant permission for my child's performance to be shared with his/her teachers.

---

Name of Child

---

Signature of Parent/Guardian

---

Date

## References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Backman, J. (1983). The role of psycholinguistic skills in reading acquisition. *Reading Research Quarterly, 18*, 466-479.
- Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 301*, 419-421.
- Bruce, P. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology, 34*, 158-170.
- Calfee, R. C., Chapman, R., & Venezky, R. (1972). How a child needs to think to learn to read. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp 139-182). New York: John Wiley & Sons.
- Chafouleas, S. M., Lewandowski, L. J., Smith, C. R., & Blachman, B. A. (1997). Phonological awareness skills in children: Examining performance across tasks and ages. *Journal of Psychoeducational Assessment, 15*, 334-347.
- Clay, M. M. (2002). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- DeFord, D. E. (2000). *Reading and writing assessment portfolio*. Carlsbad, CA: Dominic Press.
- Dodd, B., Crosbie, S., McIntosh, B., Teitzel, T., & Ozanne, A. (2000). *The preschool and primary inventory of phonological awareness*. London: Psychological Corporation.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale,

- NJ: Lawrence Erlbaum Associates.
- Lewkowicz, N. K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72*, 686-700.
- MacLean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33*, 255-281.
- Muter, V., Hulme, C., & Snowling, M. (1997). *The phonological abilities test*. London: Psychological Corporation.
- National Reading Panel. (2000). *Teaching children to read: An evidenced-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development.
- Roswell, F. G., & Chall, J. S. (1997). *Roswell-Chall auditory blending test*. Cambridge, MA: Educators Publishing Service.
- Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.
- Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 307-342). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task complexity. *Journal of Experimental Child Psychology, 38*, 175-190.
- Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P.



- Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 65-106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34*, 28-52.
- Tumer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 175-211). Hillsdale, NJ: Lawrence Erlbaum Associates.
- van Bon, W. H. J., & Duighuisen, H. C. M. (1995). Sometimes spelling is easier than phonemic segmentation. *Scandinavian Journal of Psychology, 46*, 82-94.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-ed.
- Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 11-29). New York: Guilford Press.
- Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 26*, 159-177.

## Appendix B

## Dissertation Study: Informed Consent

September, 2002

Dear Parents or Guardians,

I am a doctoral student at the University of Maryland and a former first-grade teacher and Reading Recovery teacher. I am conducting a research study that examines children's awareness of sounds, spelling knowledge, oral reading, and reading comprehension. I am requesting permission for your child to participate in my study.

If you grant permission, I will meet individually with your child for two 30-minute sessions. I will hold the first session in September and the second session in March. During these sessions, I will ask your child to complete reading and writing activities that involve looking at pictures, listening to sounds, reading stories, and writing a story. I will tape-record your child's responses to clarify discrepancies in the data.

Your child's participation in my study is voluntary. You are welcome to ask questions and free to withdraw your child from participation. The data I collect will remain confidential. I will not identify your child's name. I will destroy the audiotape upon the completion of my data collection process. The data I collect will be grouped with data that other first-grade children provide for reporting and presentation. However, with your permission, I will share your child's performance with your child's first-grade classroom teacher and Reading Recovery teacher.

Please complete the form below and return it to school with your child. If you have any questions, please contact me at (202)-544-1768 or my advisor, Dr. Marilyn Chambliss, at (301)-405-7410. Thank you for your cooperation

Sincerely,

Judith S. Concha

---

Please complete this form and return it to your child's classroom teacher.

- I grant permission for my child to participate in the reading/writing activities.  
 I grant permission for my child's performance to be shared with his/her teachers.  
 I do not grant permission for my child to participate in the reading/writing activities.  
 I do not grant permission for my child's performance to be shared with his/her teachers.

---

 Name of Child

---

 Signature of Parent/Guardian

---

 Date

## References

- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Adams, M. J., & Huggins, A. W. (1985). The growth of children's sight vocabulary: A quick test with education and theoretical implication. *Reading Research Quarterly, 20*, 262-281.
- Baker, L. (1984). Children's effective use of multiple standards for evaluating comprehension. *Journal of Educational Psychology, 76*, 588-597.
- Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. J. Pearson, M. Kamil, M. R. Barr, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 353-394). New York: Longman.
- Baumann, J. F., & Bergeron, B. S. (1993). Story map instruction using children's literature: Effects on first graders' comprehension of central narrative elements. *Journal of Reading Behavior, 25*, 407-437.
- Biemiller, A. (1970). The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly, 6*, 75-96.
- Blachman, B. A. (2000). Phonological awareness. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 483-502). Mahwah, NJ: Erlbaum.
- Bowey, J. A., & Hansen, J. (1994). The development of orthographic rimes as units of word recognition. *Journal of Experimental Child Psychology, 58*, 465-488.
- Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 301*, 419-421.

- Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. International Academy for Research in Learning Disabilities Monograph Series, No. 1. Ann Arbor, MI: The University of Michigan Press.
- Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, 88, 18-37.
- Bruce, P. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology*, 34, 158-170.
- Bourassa, D., & Treiman, R. (2003). Spelling in children with dyslexia: Analyses from the Treiman-Bourassa early spelling test. *Scientific Studies of Reading*, 7, 309-333.
- Bryant, P. E., Maclean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme, alliteration, phoneme deletion, and learning to read. *Developmental Psychology*, 26, 429-438.
- Byrne, B., & Fielding-Barnsley, R. (1991). Evaluation of a program to teach phonemic awareness to young children. *Journal of Educational Psychology*, 83, 451-455.
- Byrne, B., & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology*, 85, 104-111.
- Byrne, B., & Fielding-Barnsley, R. (1995). Evaluation of a program to teach phonemic awareness to young children: A 2- and 3-year follow-up and a new preschool trial. *Journal of Educational Psychology*, 87, 488-503.
- Calfee, R. C., Chapman, R., & Venezky, R. (1972). How a child needs to think to learn to read. In L. W. Gregg (Ed.), *Cognition in learning and memory* (pp. 139-182).

New York: John Wiley and Sons.

- Calfee, R. C., Lindamood, P., & Lindamood, C. (1973). Acoustic-phonetic skills and reading-kindergarten through twelfth-grade. *Journal of Educational Psychology*, 64, 293-298.
- Cassar, M., & Treiman, R. (1997). The beginning of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology*, 89, 631-644.
- Center, Y., Wheldall, K., Freeman, L., & McNaught, M. (1995). An evaluation of Reading Recovery. *Reading Research Quarterly*, 30, 240-263.
- Chall, J. S. (1996). *Stages of reading development* (2<sup>nd</sup> ed.). Fort Worth, TX: Harcourt-Brace.
- Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2001). Does success in the Reading Recovery Program depend on developing proficiency in phonological-processing skills?: A longitudinal study in a whole language instructional context. *Scientific Studies of Reading*, 5, 141-176.
- Clay, M. M. (1968). A syntactic analysis of reading errors. *Journal of Verbal Learning and Verbal Behavior*, 7, 434-438.
- Clay, M. M. (1972). *Sand*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1979). *Stones*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.
- Clay, M. M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.

- Clay, M. M. (2000). *Follow me, moon*. Portsmouth, NH: Heinemann.
- Clay, M. M. (2000). *No shoes*. Portsmouth, NH: Heinemann.
- Clay, M. M. (2001). *Change over time in children's literacy development*. Portsmouth, NH: Heinemann.
- Clay, M. M. (2002). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Clay, M. M., & Tuck, B. (1991). *A study of Reading Recovery subgroups: Including outcomes for children who did not satisfy discontinuing criteria*. Auckland, New Zealand: University of Auckland.
- Cohen, A. S. (1975). Oral reading errors of first-grade children taught by a code emphasis approach. *Reading Research Quarterly, 10*, 616-650.
- Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly, 3*, 499-545.
- DeFord, D. E. (2000). *Reading and writing assessment portfolio: Phonics, writing, and spelling assessment*. Carlsbad, CA: Dominic Press.
- Dodd, B., Crosbie, S., McIntosh, B., Teitzel, T., & Ozanne, A. (2000). *The preschool and primary inventory of phonological awareness*. London: Psychological Corporation.
- Dreyer, L. G., Ehri, L. C., & Flugman, B. (2005, June). *Reading rescue: First-grade tutoring facilitates reading acquisition in struggling readers*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Toronto, Ontario, Canada.
- Ehri, L. C. (1986). Sources of difficulty in learning to spell and read. In M. L. Wolraich

- & D. Routh (Eds.), *Advances in developmental and behavioral pediatrics, Vol. 7* (pp. 121-195). Greenwich, CT: JAI Press.
- Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning reading* (pp. 3-40). Mahwah, NJ: Erlbaum.
- Ehri, L. C. (2004). Development of the ability to read word: Update. In R. B. Ruddell & M. R. Ruddell (Eds.), *Theoretical models and processes of reading*, (5<sup>th</sup> ed., supplementary articles, pp. 1-35). Newark, DE: International Reading Association.
- Ehir, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly*, 27, 12-26.
- Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology*, 75, 3-18.
- Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure?: A meta-analysis of the intervention research. *Journal of Educational Psychology*, 92, 605-619.
- Fox, B., & Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research*, 4, 331-342.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Colheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301-330). Hillsdale, NJ: Erlbaum.

- Gentry, J. R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The Reading Teacher, 36*, 192-200.
- Gómez-Bellengé, F. X., Rodgers, E., & Fullerton, S. K. (2003). *Reading Recovery and Descubriendo la Lectura national report, 2001-2002*. Columbus, OH: National Data Evaluation Center.
- Gómez-Bellengé, F. X., & Rodgers, E. M. (2004). *Reading Recovery and Descubriendo la Lectura national report, 2002-2003*. Columbus, OH: National Data Evaluation Center.
- Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42*, 73-83.
- Goswami, U. (1988). Orthographic analogies and reading development. *Quarterly Journal of Experimental Psychology, 40*, 239-268.
- Goswami, U. (1990). Phonological priming and orthographic analogies in reading. *Journal of Experimental Child Psychology, 49*, 323-340.
- Goswami, U. (1998). The role of analogies in the development of word recognition. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning reading* (pp. 41-63). Mahwah, NJ: Erlbaum.
- Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 111-125). New York: Guildford Press.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Erlbaum.
- Guttentag, R., & Haith, M. (1978). Automatic processing as a function of age and reading



- ability. *Child Development*, 49, 707-716.
- Hatcher, P., Hulme, C., & Ellis, A. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development*, 65, 41-57.
- Henderson, E. H. (1985). *Teaching spelling*. Boston: Houghton Mifflin.
- Henderson, E. H., & Templeton, S. (1986). A developmental perspective of formal spelling instruction through alphabet, pattern, and meaning. *Elementary School Journal*, 86, 292-316.
- Hiebert, E. H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher*, 23, 15-25.
- Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal*, 7, 171-188.
- Huck, S. W. (2000). *Reading statistics and research* (3<sup>rd</sup> ed.). New York: Longman.
- Iversen, S., & Tunmer, W. E. (1993). Phonological processing skills and the Reading Recovery program. *Journal of Educational Psychology*, 85, 112-126.
- Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437-447.
- Juliebo, M., Malicky, G. V., & Norman, C. (1998). Metacognition of young readers in an early intervention programme. *Journal of Research in Reading*, 21, 24-35.
- Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction*, 5, 143-165.

- Kinnunen, R., Vauras, M., & Niemi, P. (1998). Comprehension monitoring in beginning readers. *Scientific Studies of Reading, 2*, 353-375.
- Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Krathwohl, D. R. (1998). *Methods of educational and social science research: An integrated approach* (2<sup>nd</sup> ed.). New York: Longman.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.
- Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18*, 201-212.
- Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263-284.
- Lundberg, I., Olofsson, A., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology, 21*, 159-173.
- MacLean, M., Bryant, P. E., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly, 33*, 255-282.
- Mann, V. A. (1993). Phoneme awareness and future reading ability. *Journal of Learning Disabilities, 26*, 259-269.
- Mann, V. A., Tobin, P., & Wilson, R. (1987). Measuring phonological awareness through invented spellings of kindergarten children. *Merrill-Palmer Quarterly, 33*, 365-

391.

- Morrow, L. M. (1984). Reading stories to young children. Effects of story structure and traditional questioning strategies on comprehension. *Journal of Reading Behavior, 16*, 273-288.
- Muter, V., Hulme, C., & Snowling, M. (1997). *The phonological abilities test*. London: Psychological Corporation.
- Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 71*, 3-27.
- Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*, 320-337.
- Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*, 154-167.
- National Reading Panel. (2000). *Teaching children to read: An evidenced-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development.
- Nicholson, T. (1991). Do children read words better in context or in lists: A classic study revisited. *Journal of Educational Psychology, 83*, 444-450.
- Pearson, P. D., & Duke, N. K. (2002). Comprehension instruction in the primary grades. In M. Pressley & C. C. Block (Eds.), *Comprehension instruction: Research-based*

- best practices* (pp. 247-258). New York: Guilford Press.
- Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first-grade children. *Merrill-Palmer Quarterly*, *33*, 283-319.
- Phillips, G., & Smith, P. (1997). *A third chance to learn: The development and evaluation of specialised interventions for young children experiencing the greatest difficulty in learning to read*. Wellington, New Zealand: New Zealand Council for Educational Research.
- Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly*, *29*, 8-39.
- Reading Recovery Council of North America (2001). *Standards and guidelines of the Reading Recovery Council of North America* (3<sup>rd</sup> ed.). Columbus, OH: The Reading Recovery Council of North America.
- Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, *36*, 321-339.
- Rittle-Johnson, B. & Siegler, R. (1999). Learning to spell: Variability, choice, and change in children's strategy use. *Child Development*, *70*, 332-348.
- Roswell, F. G., & Chall, J. S. (1997). *Roswell-Chall auditory blending test*. Cambridge, MA: Educators Publishing Service.
- Rumelhart, D. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 573-603). New York: Academic Press.
- Rumelhart, D. E. (1994). Toward an interactive model of reading. In R. B. Ruddell, M. R.

- Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading*, (4<sup>th</sup> ed., pp. 864-894). Newark, DE: International Reading Association.
- Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.
- Scott Foresman special practice books* (1979). Columbus, OH: Ohio State University.
- Shanahan, T. (1998). On the effectiveness and limitations of tutoring in reading. In P. D. Pearson & A. Iran-Nejad (Eds.), *Review of Research in Education: Vol. 23* (pp. 217-234). Washington DC: American Educational Research Association.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly, 30*, 958-996.
- Share, D. L., Jorm, A. F., MacLean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76*, 1309-1324.
- Snow, C. E., Burns, S. M., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Spector, J. E., & Moore, P. (2004). Does phonological processing distinguish between students who are more or less responsive to Reading Recovery? *Literacy Teaching and Learning, 8*, 1-25.
- Stahl, S. A., & Murray, B. (1998). Issues involved in defining phonological awareness and its relation to early reading. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning reading* (pp. 65-87). Mahwah, NJ: Erlbaum.

- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*, 3-71.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-406.
- Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task complexity. *Journal of Experimental Child Psychology, 38*, 175-190.
- Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30*, 139-181.
- Sulbzy, E., & Teale, W. (1991). Emergent literacy. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 727-757). Mahwah, NJ: Erlbaum.
- Templeton, S., & Morris, D. (2000). Spelling. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 525-543). Mahwah, NJ: Erlbaum.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York: Oxford University Press.
- Treiman, R., Berch, D., & Weatherston, S. (1993). Children's use of phoneme-grapheme correspondences in spelling: Roles of position and stress. *Journal of Educational Psychology, 85*, 466-477.
- Treiman, R., & Bourassa, D. (2000a). Children's written and oral spelling. *Applied*

- Psycholinguistics*, 21, 183-204.
- Treiman, R., & Bourassa, D. (2000b). The development of spelling skill. *Topics in Language Disorders*, 20, 1-18.
- Treiman, R., & Zukowski, A. (1996). Children's sensitivity to syllables, onsets, rimes, and phonemes. *Journal of Experimental Child Psychology*, 61, 193-215.
- Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly*, 34, 28-52.
- Vandervelden, M. C., & Siegel, L. S. (1995). Phonological recoding and phoneme awareness in early literacy: A developmental approach. *Reading Research Quarterly*, 30, 854-875.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vellutino, F. R., & Scanlon, D. (2001). Emergent literacy skills, early instruction, and individual differences as determinants of difficulties in learning to read: The case for early intervention. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 295-321). New York: Guilford Press.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-ed.
- Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly*, 28, 179-200.
- Weber, R. (1970). A linguistic analysis of first-grade reading errors. *Reading Research Quarterly*, 5, 427-451.
- Wiederholt, J. L., & Bryant, B. R. (2001). *Gray oral reading tests* (4<sup>th</sup> ed.). Austin, TX:

Pro-ed.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, 26, 159-177.