

FINISH-A-RHYME-STORY:
A RHYME CLOZE ASSESSMENT FOR PRESCHOOL CHILDREN

by
Kimberly J. Condie

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Department of Communication Disorders
Brigham Young University

August 2009

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Kimberly J. Condie

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

Barbara Culatta, Chair

Date

Martin Fujiki

Date

Kendra Hall

Date

Barbara Lawrence

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Kimberly J. Condie in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Barbara Culatta
Chair, Graduate Committee

Accepted for the Department

Date

Ron W. Channell
Graduate Coordinator

Accepted for the College

Date

K. Richard Young
Dean, David O. McKay School of Education

ABSTRACT

FINISH-A-RHYME-STORY:

A RHYME CLOZE ASSESSMENT FOR PRESCHOOL CHILDREN

Kimberly J. Condie

Department of Communication Disorders

Master of Science

Educators need measurement tools to determine phonological awareness in young children. This study investigated the appropriateness of rhyme cloze tasks, referred to as Finish-a-Rhyme-Story items, which were designed to measure preschool and kindergarten children's early rhyme development. The rhyme cloze tasks required children to verbally complete a sentence by filling in a final rhyming word that matched a rhyme pattern highlighted in a short story that was read aloud to them. The task required rhyme awareness as well as comprehension of the language in the story.

Twenty-four items were individually administered to preschool ($n = 207$) and kindergarten ($n = 382$) children to determine item performance and discriminative power. Rasch analysis indicated that the difficulty level of the items was well matched for the sample indicating that the items were developmentally appropriate for preschool and

kindergarten children. Several analyses of variance (ANOVA) compared the performance of preschool and kindergarten children as well as the performance of monolingual English speaking (ENG) children and English Language Learners (ELL) to determine if there were group differences on the rhyme cloze measure. Results also indicated that the items have the ability to discriminate between children with high and low level rhyming ability based on the Rasch model; kindergarten children were more aware of the rhyme component than preschool children and ENG children were more aware than ELL children.

ACKNOWLEDGMENTS

This project has been a tremendous growing experience for me, and I appreciate the many people who so willingly gave their time to help me complete it. Thank you to the preschool and kindergarten teachers who gave us the opportunity to work with the wonderful children you see every day; they truly were fun! Thanks to the team of generous and talented people who helped in collecting the data. And thanks especially to the wonderful faculty members who guided me through this process. Dr. Culatta, I have enjoyed working with you and learning from you. Thank you for your continual encouragement and constant support. Dr. Hall, thank you for getting me involved in research and especially for bringing me on to this project. Dr. Fujiki, thank you for your patience and optimism. Dr. Lawrence, thank you for being so involved and accessible during the data collection and analysis; this project could not have happened without you. Also, thanks to Ann Sharp whose work in the beginning of this project helped me get the ball rolling, and Lee Robinson whose listening ear helped me through the bumpy parts.

A thesis is like a roller coaster, and I could not have survived this ride without my loving family. Thanks to Mom and Dad for preparing me for the ride, buying the ticket, and encouraging me to get in line. Thanks to Brodie for reminding me to let go and enjoy the ride even during the plunges and sharp turns. And thanks to Brett for being the seatbelt that held me together through the entire trip.

Table of Contents

	Page
Table of Contents	vii
List of Tables	x
List of Appendixes	xi
Introduction.....	1
Review of Literature	3
Nature of Phonological Awareness.....	3
Size of Phonological Unit.....	4
Sentence level.	4
Syllable level.....	4
Onset-rime level.....	5
Phoneme level.....	5
Demands of Phonological Task	5
Sound play and evoked responses.....	6
Recognition and generation.	6
Metalinguistic manipulation of sounds.....	7
Nature of Rhyme Tasks	8
Sound Play	8
Recognition.....	9
Generation.....	9
Influence of Rhyme on Reading	10
Purposes of Early Assessments of Rhyming	11
Identify Children Who are at Risk.....	11

Provide Relevant Instruction.....	12
Rhyme Cloze as an Assessment Tool.....	13
Nature of the Task.....	14
Previous Uses: Nursery Rhyme Completion	15
Potential Factors Influencing Rhyme Cloze Performance.....	16
Story content.	16
Semantic associations.	17
Vocabulary.....	18
Neighborhood density.....	19
Syntactic complexity.....	19
Factors Identified in the Finish-a-Rhyme-Story Pilot Study and Field Test	20
Semantic associations.	21
Vocabulary.....	22
Syntax.	22
Method	24
Participants.....	24
Stimuli.....	25
Procedures.....	26
Administrators and Training.....	26
Test Administration	27
Demonstration item.....	30
Sample item.	30
Assessment items.....	31

Scoring	31
Data Analysis	33
Results.....	34
Item Performance.....	34
Item Difficulty	34
Infit and Outfit.	35
Discrimination.....	35
Internal consistency	35
Differential Item Function	38
Student Ability	38
Preschool versus Kindergarten Performance	39
Monolingual English Speakers versus English Language Learners Performance.....	41
Discussion.....	46
Item Performance.....	46
Preschool versus Kindergarten Performance	50
Monolingual English Speakers versus English Language Learners Performance.....	51
Possible Factors Confounding Performance.....	52
Item Characteristics	52
Sample Characteristics.....	53
Administration Characteristics.....	56
Future Research	57
References.....	59

List of Tables

Table	Page
1. Pilot Study Item Response Ranking	28
2. Balancing and Ordering of Items Within Sets	29
3. Item Difficulty Ranking based on Rasch Model Analysis.....	36
4. Item Performance Based on Rasch Model Analysis.....	37
5. Differential Item Functioning of Preschool and Kindergarten Performance Measured by Welch (p_w) and Mantel-Haenszel (p_{MH}) Probabilities.....	40
6. Analysis of Variance (ANOVA) Comparing Preschool and Kindergarten Performance and Monolingual English Speakers (ENG) and English Language Learners (ELL) Performance	44
7. Analysis of Variance (ANOVA) Comparing Monolingual English Speakers (ENG) and English Language Learners (ELL) in Preschool and Kindergarten	45

List of Appendixes

Appendix	Page
A. Finish-a-Rhyme-Story Assessment Items.....	64
B. Example of a Finish-a-Rhyme-Story Item Illustration (Item C1).....	66
C. Parent Consent for Student	67
D. Finish-a-Rhyme-Story Administration Protocol.....	69

Introduction

With the current emphasis on stimulating early pre-reading skills, educators need tools for assessing young children's phonological awareness. An assessment tool that provides information about implicit awareness of rhyme could be used to identify preschool children who struggle with phonological awareness. *Implicit awareness of rhyme* is an early stage of phonological awareness in which children are sensitive to rhyme patterns without yet having the ability to talk about rhyme words; this sensitivity develops before the ability to explicitly manipulate sounds (Bradley & Bryant, 1983; Brady & Shankweiler, 1991; Goswami & Bryant, 1990; Muter, Hulme, Snowling, & Stevenson, 2004). The latent nature of implicit phonological awareness creates many challenges in measuring this early phonological skill. However, there is a need for tasks that identify this skill because the sooner phonological struggles are detected the greater the chance that children will be able to receive early intervention and training services.

This study examines a group of original rhyme awareness items using a *cloze* or fill-in-the-blank response format. *Rhyme cloze* items are sentence completion tasks that include rhyme words at the end of each sentence; children are given the opportunity to finish the final sentence with a word that rhymes. The purpose of this study is two-fold. The first purpose is to create rhyme cloze items (referred to as Finish-a-Rhyme-Story items in this study) that evoke implicit rhyme awareness in preschool children. The second purpose is to determine which of the items in the pool are able to discriminate between children with high and low level rhyming performance on the Finish-a-Rhyme-Story items. The study will also examine ability level differences between English language learners (ELL) and children who are monolingual English speakers

(ENG). The results of this study will provide information about the items that may be used to develop a rhyme awareness assessment for young children.

Review of Literature

An understanding of phonological awareness is crucial when developing phonological awareness assessments because phonological awareness involves several levels of development. An early stage of phonological awareness is implicit awareness of rhyme. *Implicit awareness* of rhyme is demonstrated when children are sensitive to rhyme patterns but are not yet able to talk about or manipulate rhyme words (Bradley & Bryant, 1983; Brady & Shankweiler, 1991; Goswami & Bryant, 1990; Muter et al., 2004). Because this study addresses implicit awareness of rhyme, this literature review will begin with a discussion of general phonological awareness development and then follow with a more detailed discussion of rhyme. The potential uses for an early rhyme assessment tool will then be addressed. Finally, the possible confounding factors of such an assessment tool will be discussed.

Nature of Phonological Awareness

Phonological awareness is the knowledge of sound structures (Gillon, 2004; Owens, 2005). To understand phonological awareness it is important to understand the size of the phonological unit and the demands of the phonological task. Phonological awareness may be demonstrated through tasks that require implicit or explicit awareness of sounds. That is, phonological awareness tasks may demand simple sensitivity to sound structures or more mature awareness that requires manipulation of sounds. The size of the phonological unit influences children's phonological awareness skills; phonological awareness may occur in relatively large sound units or in small ones. Therefore, children's phonological awareness development may be influenced by two factors: the size of the phonological unit and the demands of the phonological task. Though it is difficult to completely separate the influence of each of these factors, for

the purpose of clarity, the two factors will be discussed separately in this literature review.

Size of Phonological Unit

Within the development of phonological awareness, children recognize larger, holistic sound units before detecting smaller sound structures. Relatively large sound units are composed of several sounds, whereas small sound units are composed of only individual sounds or phonemes (Gillon, 2004; Owens, 2005). As children develop, they will likely acquire phonological awareness beginning at the sentence level and moving to the phoneme level.

Sentence level. In sentences, phonological awareness is demonstrated when children detect the words that compose a sentence. For example, when children are able to clap out or count the number of words in a sentence, they are demonstrating phonological awareness at the sentence level. Once children are aware of the words that compose a sentence, they may develop an awareness of the sounds that compose words. In words, phonological awareness is believed to occur at several levels: the syllable level, the onset-rime level (which is intrasyllabic), and the phoneme level (Gillon, 2004).

Syllable level. Phonological awareness at the syllable level occurs when children recognize the syllables that compose a word. This awareness demonstrates knowledge that words can be divided into different sound units (Gillon, 2004; Owens, 2005). For example, children who have phonological awareness at the syllable level may recognize that *baby* is composed of two syllables: *ba-by*. Once phonological awareness of syllables is established, children may progress to phonological awareness within syllables.

Onset-rime level. Children who have phonological awareness at the onset-rime level may detect sound structures within syllables. *Onset* refers to the beginning sound or cluster in a word (e.g., *m* in *mat*) and *rime* refers to the ending sound (i.e., the vowel and final consonant or coda) in a word (e.g., *at* in *mat*). Words with the same rime are said to rhyme. Children tend to show awareness to rhyme early in development (Gillon, 2004; Goswami, 2001; Paul, 2007). Phonological awareness at the syllable or onset-rime level deals with relatively large sound units (because syllables and rimes are composed of several sounds), whereas phonological awareness at the phoneme level deals with the smallest unit of speech: phonemes (Cunningham, Cunningham, Hoffman, & Yopp, 1998; Snow, Burns, & Griffin, 1998).

Phoneme level. Children who demonstrate phonological awareness at the phoneme level may demonstrate knowledge of the individual sounds that compose a word. Phonemic awareness is more mature than syllable or rhyme awareness because it requires recognition of the smallest sound units; this awareness is needed for some higher level phonological tasks. However, it is important to note that there are many types of phonological tasks.

Demands of Phonological Task

In addition to the size of the unit, children's phonological awareness abilities are influenced by the demands of the phonological task. Various tasks demand a range of phonological skill capacities. Phonological skills do not generally develop at the same time but rather progress from one level of phonological awareness to another more in-depth one. Phonological development progresses from an awareness of sounds which involves sound play to higher metalinguistic manipulation of sounds which includes deletion of one phoneme to make a different word. Some of these task demands may

occur with various sound unit sizes and may relate to both phonological and phonemic awareness.

Sound play and evoked responses. Sound play is an early phonological task when children show awareness of sounds through play. For example, children may play with sounds by making up funny words for objects or playing with nonsense words (Snow et al., 1998). Children's awareness of sounds may also be evoked through tasks that prime sound awareness but do not require children to respond in a particular way. This may occur when children are exposed to rhyming words in a song and, without being asked, begin to sing along with rhyming words of their own. Tasks that evoke responses may tap into children's implicit awareness. Implicit awareness is the earliest form of phonological awareness. When children acquire implicit awareness, they become sensitive to sound structures. They may hear the various sounds in sentences or words and may play with these sounds.

Recognition and generation. Recognition and generation of sounds becomes possible when children develop explicit awareness of sounds. This more mature awareness of sounds is characterized by children's ability to talk about or identify the sound structures they hear. Children who have recognition of sounds are able to identify and sort sounds. One type of recognition task is sorting in which children identify the words in a group that begin or end with the same sound. Children who have developed generation skills would be able to perform tasks that require them to produce certain types of words. That is, when asked to think of a word that starts with /b/, children with generation abilities would be able to produce words such as *ball*, *bike*, or *baby*.

Metalinguistic manipulation of sounds. A higher level metalinguistic understanding of sound structure is required in tasks that demand complex manipulation of sounds. Early manipulation may occur when children change the sound units in a word to produce a different word. This ability leads to higher level metalinguistic manipulation of sound structures. Metalinguistic skills include segmenting and blending. *Segmenting* refers to children's ability to divide a word into smaller sound units. *Blending* refers to children's ability to combine smaller sound units to create words (Goswami, 1986, 1988; Goswami & Bryant, 1992). Depending on developmental level, segmenting and blending may be used with various sizes of phonological units such as syllables or phonemes (Owens, 2005).

When segmenting and blending occur at the phoneme level, they are part of the most mature form of phonological awareness: phonemic awareness. When children acquire phonemic awareness, they have the ability to manipulate phonemes to create and deconstruct words. Differing opinions have been published about the age in which the shift to phonemic awareness occurs, but it is generally expected to take place between the late preschool years and early school years (Charles-Luce & Luce, 1990, 1995; Storkel, 2002; Walley, 1993). Early phonological awareness skills lead to phonemic awareness. For example, rhyme develops before the ability to segment (Cunningham et al., 1998). Very young children may not yet have phonemic awareness skills; however it is likely that they have emerging rhyme skills. Because this study deals with the rhyme awareness of young children, the nature of rhyme development will be addressed in the following section.

Nature of Rhyme Tasks

In the beginning stages of phonological awareness, children tend to have an awareness of rhyme because rhyme is a relatively large sound unit (i.e., groups of sounds as opposed to individual ones). The understanding of rhyming sounds can begin as an implicit awareness and then may mature to a more explicit one. This means that before children are able to explicitly talk about and manipulate rhyme, they have an awareness or sensitivity to rhyme; they can detect or play with rhyme, but cannot label it. Children's rhyme abilities typically progress from rhyme play to recognition and generation.

Sound Play

Implicit awareness or sensitivity to rhyme may be detected when children gravitate toward, play with, and demonstrate enjoyment in rhyme. Many children tend to demonstrate implicit awareness to rhyme before entering school (Dowker, 1989; MacLean, Bryant, & Bradley, 1987). It is not uncommon to hear young children playing with rhyme sounds with both real and nonsense words. Even some 2-year olds have been found playing with sounds such as "pancake, cancake, and canpake" (Snow et al., 1998, p. 51). This type of sound play is considered to be implicit because although children seem to hear or recognize rhyme at this stage, they are not likely to be able to talk about rhyme in a way that explains why these words sound alike. Implicit awareness may also be demonstrated when children produce rhymes during tasks that do not specifically ask them to produce a rhyme but rather evoke that response because of their sensitivity to rhyme. Implicit awareness of rhyme precedes the explicit awareness that enables rhyme generation.

Recognition

Explicit or metalinguistic awareness of rhyme is a more mature level of phonological awareness that is characterized by an ability to label and identify rime patterns. Explicit rhyme knowledge is demonstrated when children become able to recognize rhymes; they may demonstrate knowledge that several words have the same rime and may therefore be sorted into categories of sounds (Goswami & Bryant, 1992). For example, one type of recognition task is an oddity task where children are given a set of words such as *dog*, *cat*, and *hat* and asked to identify the word that doesn't belong. Children with rhyme recognition skills could recognize that *dog* does not fit in the same category as *cat* and *hat* which are rhyming words. In addition, they would be able to analyze or talk about the fact that the words rhyme because they sound the same at the end. Children with explicit awareness to rhyme may also have the ability to generate rhymes.

Generation

Children who have acquired rhyme generation skills are able to produce words that rhyme with a set. For example, when given a set of words such as *car*, *tar*, and *bar* and asked to produce a word that rhymes, children with rhyme generation skills could produce a real or nonsense word, such as *star* or *dar*. Awareness of rhyme, such as rhyme recognition and generation, is generally observed before other, more complex forms of phonological awareness (e.g., phonemic awareness). In fact, young children's awareness of rhyme has been shown to substantially predict later phoneme awareness (Anthony & Lonigan, 2004). Early rhyme awareness facilitates later acquisition of reading skills.

Influence of Rhyme on Reading

The process of learning to read begins long before children participate in formal reading instruction (Justice, Invernizzi, & Meier, 2002; Snow et al., 1998). Rhyme awareness develops before children enter school and plays a key role in the process of reading acquisition (Justice et al., 2002; Neuman, Copple, & Bredekamp, 2000). Decades of research confirm the strong and even predictive relationship between young children's rhyme awareness and their later reading success (Bradley & Bryant, 1983; Goswami & Bryant, 1992; Justice et al., 2002; Snow et al., 1998). Therefore, children's early phonological skills, such as rhyme awareness, influence reading success and may be fundamental to reading instruction. For example, children's familiarity with rhyme may be used as a foundation for early reading instruction. That is, children may be taught to recognize that such words as *bat*, *hat*, *mat*, and *cat* have similar endings; therefore, children learn that if they can read one of these words, they can read the other three. This concept of substitution of phonemes to create multiple words requires higher level skills than simple rhyme awareness; however, the task is based on children's familiarity with rime units.

Because phonological awareness plays a key role in reading acquisition, it is important to develop assessments for various levels of phonological awareness. These assessments would enable educators to monitor children's developing skills. Currently, there are many phonological assessments that measure high level phonological skills; however, there are few assessments that measure beginning phonological awareness abilities such as implicit awareness of sound patterns.

Purposes of Early Assessments of Rhyming

Children who enter school without strong phonological awareness skills are likely to struggle with learning to read (Justice et al., 2002; Snow et al., 1998). Assessments enable parents and teachers to identify obstacles that might impede children's later success (Snow et al., 1998). In the book, *Preventing Reading Difficulties in Young Children*, the authors encourage, "When deciding which factors to use to identify children who are at risk for reading difficulties, the main determinant should be the strength of association" (Snow et al., 1998, p. 102). Since early rhyme awareness has acknowledged strong associations with later reading abilities (Bradley & Bryant, 1983; Bryant, 1998), an assessment that taps into children's implicit awareness to rhyme is an appropriate and needed tool. Unfortunately, some difficulties arise in trying to measure implicit awareness.

The nature of implicit awareness of rhyme makes it difficult to measure. However, an assessment that detects sensitivity to sound patterns may enable discovery of early phonological struggles that could impact emergent reading skills; an assessment that evaluates implicit awareness has the power to identify young children at risk for future reading failure.

Identify Children Who are at Risk

Early identification of children at risk leads to early intervention in the form of phonological awareness training. An assessment tool that identifies early delays and indicates a need for appropriate training could be beneficial for both teachers and students because early intervention can prevent literacy difficulties by providing more comprehensive instruction in areas of weakness (Justice et al., 2002). It has been long-established that children who struggle with rhyme awareness early on are more likely to

struggle with reading later (Anthony & Lonigan, 2004; Bradley & Bryant, 1985; Bradley & Bryant, 1983; Mann & Liberman, 1984; Share, Jorm, Maclean, & Matthews, 1984; Stanovich, Cunningham, & Feeman, 1984; van Kleeck, Gilliam, & McFadden, 1998; Wagner & Torgesen, 1987). Expectations for children's reading preparation and ability are influenced by their grade level. In preschool, implicit awareness of phonological units is both developmentally appropriate and necessary for kindergarten preparation. In kindergarten, children must learn explicit phonological awareness and higher level metalinguistic functioning in order to emerge into higher level skills. It is critical that phonological awareness assessments for children in preschool and kindergarten reflect their developing skills.

Because rhyme is an early developing skill, rhyme assessments must be conducted in preschool, prior to the formal reading instruction that is received in kindergarten and 1st grade. It is necessary for rhyme assessments to occur early because it has been found that once children have begun to develop higher level metalinguistic skills, rhyme ability is not a reliable discriminant (Catts, Fey, Tomblin, & Zhang, 2002; Goswami & Bryant, 1992). However, at the preschool level, an assessment that detects implicit awareness to rhyme would be appropriate. Therefore, with a tool that allows for early identification of deficits, teachers will have the ability to provide relevant instructions to help children improve in their area of weakness.

Provide Relevant Instruction

Children who struggle with phonological awareness have been found to respond to and improve with training (Brady, Fowler, Stone, & Winbury, 1994; Carroll & Snowling, 2001; Snow et al., 1998). For decades, studies have explored and verified the effectiveness of early training (Bradley & Bryant, 1983). Layton, Deeny, Tall, and

Upton (1996) performed a longitudinal study that found 3- and 4-year-olds improved in phonological awareness ability following a phonological awareness program involving several tasks. Similar results are seen when preschool children are trained in various phonological awareness tasks (Bradley & Bryant, 1983; Carroll & Snowling, 2001). Specifically, the rhyming abilities of preschool and kindergarten children improve with specific instruction (Brady et al., 1994; Culatta, Hall, Kovarsky, & Theadore, 2007; Culatta, Setzer, Wilson, & Aslett, 2004; van Kleeck et al., 1998); training in rhyme can enhance children's progress in reading (Bradley & Bryant, 1983; Goswami & Bryant, 1992).

An early rhyme awareness assessment has the potential to provide educators with the ability to evaluate children's skill level and design appropriate instruction (Justice et al., 2002; Neuman et al., 2000; Peverly & Kitzen, 1998). The results of the assessment would allow teachers to establish individualized goals for children. Assessment outcomes would aid teachers in determining which skills need more support and which parts of instruction require manipulation. Using the assessment to gain pre and post data would make it possible for teachers to monitor instruction effectiveness. As teachers have increased opportunities to improve instruction, children are benefited. Therefore, with evidence of success of phonological awareness training, the need for a quality rhyme assessment tool is clear.

Rhyme Cloze as an Assessment Tool

It is important and appropriate to understand the capacity of specific tasks prior to using them as an assessment tool. Various forms of rhyme cloze tasks have been used in numerous settings and for multiple purposes (Christie, Enz, & Vukelich, 2007; Dorsey, 1972; Gillon, 2004; Ziefert, 2000). The nature of a rhyme cloze task makes it a

beneficial tool because it evokes a response that could reveal sensitivity to sound patterns. Specifically, the rhyme cloze task used in this study, the Finish-a-Rhyme-Story task, has the potential to be a useful assessment tool for detecting implicit rhyme awareness in young children.

Nature of the Task

A *rhyme cloze* is a task in which children fill-in or complete a rhyme. Finish-a-Rhyme-Story is considered to be a rhyme cloze task because children have the opportunity to complete a rhyme story. Rhyme cloze tasks do not require children to spontaneously produce rhyming words. Instead, rhyme is primed as the task provides contextual support in the form of a short story that compels use of rhyme words and black-and-white illustrations that keep children's interest and attention (see Appendix B for an example of Finish-a-Rhyme-Cloze illustrations). The task evokes rhyme, thus tapping into children's implicit awareness of rhyme. During rhyme cloze tasks, a short story, in which the last word in each line rhymes with the others but the final word of the story is left blank, is read aloud. The task allows children to produce a word that both rhymes and completes the meaning of the story.

Rhyme cloze stories should have semantic and phonological components that make the task suitable for young children; the semantic constituent or context given in the task makes it more concrete for young children because it does not require the high level metalinguistic processing that is needed for generative tasks that lack context. These characteristics make rhyme cloze tasks suitable for preschool-aged children. For this reason, similar tasks have been used in the past to evoke rhyme awareness.

Previous Uses: Nursery Rhyme Completion

The nature of rhyme cloze stories may be easy for young children because the stories have many similarities to nursery rhymes. Not surprisingly, nursery rhyme completion tasks have previously been used to monitor children's rhyme ability (Dorsey, 1972). Because children are commonly exposed to nursery rhymes, many studies have been performed to examine the effect of nursery rhymes on children. It has been shown that children's nursery rhyme knowledge is related to their awareness of rhyme up to 2 years later; this relationship remains to be true even when children's IQ and their mother's level of education are considered. Additionally, studies show that 3-year-old's familiarity with nursery rhymes has a positive effect on their phonological awareness and on their later success in reading at age 5 and 6 (Bryant, MacLean, Bradley, & Crossland, 1990; Goswami & Bryant, 1990; MacLean et al., 1987). Nursery rhyme knowledge in 5- and 6-year-olds led to significant success in word identification skills in second grade (Fernandez-Fein & Baker, 1997). These findings on nursery rhyme knowledge may link general rhyme knowledge to later reading achievement.

Another common factor between nursery rhymes and rhyme cloze tasks is the child-friendly nature. It is important to have child-friendly tasks when working with young children (Kieff & Casbergue, 2000). Tasks that are fun and friendly make children more comfortable and therefore more likely to produce responses that accurately represent their ability. Rhyme cloze stories can be delivered in a friendly, playful manner which allows children to become engaged in the task. Therefore, although the task is an assessment, children may not view it as a test. With the developmentally appropriate, familiar, and appealing nature of Finish-a-Rhyme-Story,

this assessment has the potential to be an effective tool for detecting implicit rhyme awareness in young children.

Potential Factors Influencing Rhyme Cloze Performance

Though a rhyme cloze task is likely to be child-friendly and developmentally appropriate to use when assessing implicit rhyme awareness in young children, the task may be more complex than it appears. Because rhyme cloze tasks embed rhyming words into a story context, there are linguistic as well as phonologic factors operating in the task. This section will identify the variables that could impact student performance.

Story content. The story's content or the linguistic context of rhyme cloze tasks plays a key role in the nature of the task. The content and theme of the story has the ability to make the task more salient and more enjoyable for young children. However, the story content may also contribute to some of the confounding issues of the task. For example, young children are initially more focused on the meanings of words than the sounds that compose words because the meaning of words is more significant to their ability to communicate (Goswami & Bryant, 1990). As children process sentences or phrases they are continually updating their understanding of what they are reading or hearing. They apply their knowledge of word and sentence meaning to their knowledge of the world and to their memory of the preceding information in order to make sense of what they are processing (Ehri & Snowling, 2004). A semantic response which completes the story of a rhyme cloze item is a natural outcome of this process. For this reason, a semantic response, regardless of its phonological form, may be more compelling to young children performing a rhyme cloze task than a response that both semantically and phonologically fits the story (i.e., completes the meaning and matches the phonological structure of the modeled words).

Several studies have addressed young children's affinity for semantic responses in rhyming tasks; these studies provide semantic and phonologic distracters in assessments where children are asked to choose a word that matches a target. It was hypothesized that young children would initially choose semantic distracters and then progress to choosing phonological distracters as they mature and acquire phonological awareness (Carroll & Snowling, 2001); for example, when given the words *house*, *horse*, and *mouse* and asked to identify the two words that go together children who choose *mouse* and *horse* would focus on the fact that both are animals and would ignore the fact that *mouse* and *house* share the same rime ending. Children who choose a semantic distracter would likely not yet be using phonological awareness to complete the task (Bryant, 1998). However, these studies have found that preschool children have a more difficult time overlooking phonologic rather than semantic distracters and therefore are demonstrating phonological awareness (Byrne & Fielding-Barnsley, 1993; Cardoso-Martins, 1994; Carroll & Snowling, 2001). This suggests that although some very young children may be distracted by semantic responses, phonological awareness assessments are appropriate for preschool children (Carroll & Snowling, 2001).

Semantic associations. Story content is not the only issue confounding rhyme cloze responses. The specific target words used in an item also play a key role in children's performance. It is possible for certain items to be more difficult than others because of the degree of semantic relatedness between key words in the story. For example, *bread* and *butter* have high semantic associations because they often go together whereas *bread* and *rolls* do not often go together but are in the same generic category (Carroll & Snowling, 2001; Moss, Hare, Day, & Tyler, 1994). The semantic

relatedness of the key words can impact the rhyme cloze response in two ways. First, the final words in the story may have a common semantic association that leads children to give the associated word and overlook a response that may fit the story both semantically and phonologically. Using the words above, this may occur when a story such as *On this day that is so sunny, I think I will use my money, to buy myself some bread and ___* evokes the response *butter* instead of *honey* which would fit semantically and phonologically. Second, semantic relatedness may impact children's responses if the response word has a semantic association that is more common to children. For example, in a story such as *Mix some dough in a bowl, bake it over a hot coal, and soon you'll have a tasty ___* children may be more inclined to respond with *bread* rather than *roll* because bread is more common to them. This second example may also be related to children's vocabulary.

Vocabulary. Without adequate vocabulary, children may lack the knowledge necessary to effectively understand key words in a rhyme cloze task. A substantial amount of vocabulary growth occurs during exposure to oral and written language (Hayes & Ahrens, 1988; Nagy & Anderson, 1984; Nagy, Herman, & Anderson, 1985; Snow et al., 1998). The background experiences that children bring to a rhyme assessment—such as joint book reading and language play—influence familiarity with certain vocabulary (Carroll & Snowling, 2001; Ninio & Bruner, 1978; West, Stanovich, & Mitchell, 1993; Whitehurst, Falco, Lonigan, & Fischel, 1988). Studies have shown that words that are acquired early are easier for children to understand because these words are better represented in children's memory (Carroll & Snowling, 2001; De Cara & Goswami, 2003; Metsala, 1999). Furthermore, vocabulary level is shown to be an

independent predictor of children's phonological awareness (Carroll & Snowling, 2001). Therefore, in a rhyme cloze task, the complexity of the vocabulary that is used will influence performance. Thus it is important that age-appropriate vocabulary is used when creating rhyme cloze items. Otherwise, words that are unfamiliar, abstract, or too complex for young children could confound the assessment results.

Neighborhood density. *Neighborhood density* refers to the amount of exposure an individual receives to similar sounding words. The number of words that have similar sounds increases within children's mental lexicon as their vocabulary grows. Neighborhood density has consequences in children's ability to handle phonological awareness tasks such as detecting onset and rime because words with more common rimes may be more familiar to children. For example, children are likely to be familiar with more words that contain the rime *-at* than words that contain the rime *-ig*. Studies show that words used in an assessment affect children's performance when the amount of exposure to similar sounding words is dense (Carroll & Snowling, 2001; De Cara & Goswami, 2003). However, other research shows that neighborhood density does not influence children's rhyming ability (Stadler, Watson, & Skahan, 2007). The disagreement on the relationship between neighborhood density and rhyme requires more research. However, to ensure that neighborhood density does not cause confounding issues in a rhyme cloze task, it should be taken into account and controlled when items are created.

Syntactic complexity. Syntactic complexity of the sentences may also play a role in children's ability to correctly respond to rhyme cloze items. Research has shown that syntactic awareness is necessary for making predictions about the next sequence of

words (Siegel, 1993). Therefore, children's understanding of syntax has a great impact on their ability to predict the word required to complete a sentence. Unfortunately, natural syntax may be disrupted if designers of a rhyme cloze measure take poetic license and create awkward sentences to make rhyme words fit at the end of the story lines. For example, instead of saying *In order to taste the cake, I just have to take one bite* the word order could be altered to make the final words rhyme: *In order to taste the cake, I just have one bite to take*. This change in word order likely increases syntactical complexity.

Also, in order to correctly fill in a missing word, children must have sensitivity to syntactic categories and be able to recognize whether the missing word is a noun, adjective, preposition, or verb (DaFontoura & Siegel, 1995; Lipka & Siegel, 2007). With these potential confounding issues of syntax, along with the semantic issues discussed earlier, there are many factors that must be considered when developing rhyme cloze items.

Factors Identified in the Finish-a-Rhyme-Story Pilot Study and Field Test

A Finish-a-Rhyme-Story pilot study which included approximately 700 preschool and kindergarten children was performed over a 6 year period prior to the current study. Potential confounding factors such as story content; semantic associations; vocabulary; and syntactic complexity were observed in the Finish-a-Rhyme-Story pilot study. These factors were then taken into account as new and revised items were created in an attempt to control these potential confounding issues; the items were then used in a Finish-a-Rhyme-Story field test and in the current study.

The pilot data indicated a developmental difference between preschool and kindergarten children's performance. That is, the majority of the preschool children

tended to produce responses that finished sentence completion items semantically but did not fit the phonological rhyme structure of the item. In contrast, though we do not know that they were directly paying attention to both the meaning and the sound structure, the majority of kindergarten children were able to produce a response that completed many of the items both semantically and phonologically. Several of the items that did not generate this trend were found to have confounding issues; the strength of these factors could have overpowered a correct (i.e., semantic and phonological) response. While we did not expect all preschoolers to respond semantically and all kindergarteners to respond semantically and phonologically, we did gather important information from the pilot study. Specifically, the pilot study showed that several of the items demonstrated confounding factors of semantic relatedness, vocabulary, and syntax.

Semantic associations. Pilot data demonstrated that children had difficulty completing several items due to the strong semantic association of the words in the item. This was observed in the pilot item, *I love my little bear, I show him that I care, when I brush his furry hair, I let him use my [chair]*. For this item, only 4% of the preschool children produced responses that fit the item both semantically and phonologically; 44% of the kindergarten children produced responses that fit both semantically and phonologically. Though kindergarten children performed better, many children, both in preschool and kindergarten, responded with *bed* which was considered a semantic but not phonologic response. It is thought that *bed* was a frequent answer because of its semantic association with *chair* and the possibility that it is more common for young children to have their own bed than their own chair.

Vocabulary. Data gathered during the pilot study showed that items with possibly unfamiliar vocabulary resulted in low semantic and phonological responses in a few Finish-a-Rhyme-Story items. For example, on the item, *When I get dressed, I wear the very best. I go to my chest, and pull out my [vest]*, both preschool and kindergarten children struggled to complete the item with a response that fit both semantically and phonologically; less than 10% of both preschool and kindergarten children produced responses to this item that fit both semantically and phonologically. It is assumed that children struggled with this item because they were unfamiliar with the words *chest* and *vest*; though children may have been familiar with a *toy chest*, they likely did not know of a *chest* for clothing or of a *vest* because these words do not regularly occur in young children's lives. This item demonstrates the confounding issue of vocabulary and emphasizes the need to use age-appropriate vocabulary when developing Finish-a-Rhyme-Story items for young children.

Syntax. Unnatural syntax was also shown to cause some problems during the pilot study. In the item, *What do I see? An angry bee. He's coming from that tree. Hope he doesn't sting [me]*, children struggled to complete the item with a response that was semantic and phonological. On this item, approximately 22% of preschool children produced a response that fit the item both semantically and phonologically. Kindergarten children performed better, with 65% of the responses fitting both semantically and phonologically, but this number was still lower than kindergarten performance on other items. It is likely that this struggle occurred because the intended response word was a pronoun as opposed to a more concrete noun or verb. This item may also be viewed as one that uses unnatural language because poetic license was

exercised to ensure that each line ended with a rhyme. Due to the word class issue and the unnatural language used in this item, it was not used in the current study.

Based on the data from the pilot study, items were revised or eliminated. The revised items, along with some newly created ones that controlled for the confounding issues as much as possible, were used in a small Finish-a-Rhyme-Story field test that included 75 preschool and kindergarten children. The field test was conducted to collect preliminary data on the new and revised items in order to create the stimuli for the current study. After the field test, items that produced inadequate results were eliminated (see Appendix A for a list of items used in the current study).

The purpose of the current study is to develop Finish-a-Rhyme-Story (i.e., rhyme cloze) items that control for confounding factors and to determine which of these items successfully discriminate between high and low level rhyming performance based on this task. Though the Finish-a-Rhyme-Story items are not developmentally advanced enough to be used for a kindergarten assessment, kindergarteners were used in the study in order to compare preschool versus kindergarten responses to the items. In addition, comparisons between monolingual English speakers (ENG) and English Language Learners (ELL) will be made. With this information, it will be possible to begin developing an assessment that may help teachers identify young children who are struggling with implicit rhyme awareness.

Method

Participants

The individuals participating in this study were preschool and kindergarten children with consent from a legal guardian (see Appendix B for consent form). The study was conducted over a 3 month period in the middle of the school-year (i.e., November, 2008 through January, 2009) in 22 preschool classrooms ($M = 9$ students with consent per classroom, $SD = 4$) and 27 kindergarten classrooms ($M = 14$ students with consent per classroom, $SD = 7$). This included 207 preschool children between the ages of 3:2 (i.e., 3 years 2 months) and 5:7 ($M = 4:7$, $SD = 6$ months) and 382 kindergarten children between the ages of 4:11 and 7:7 ($M = 5:9$, $SD = 4$ months). The majority of the children in the study attended schools in which at least 10% of the population was considered to be English Language Learners (ELL) and at least 30% of the population was receiving free or reduced school lunch indicating low socioeconomic status (SES). In order to obtain a large preschool sample, 4 home-based preschools were also included; 102 (10 classrooms) of the preschool children in the study were from these schools.

Based on teacher report, children who are considered ELL and children with known speech and language delays were included in the study; 13.87% of the preschool children and 20.94% of the kindergarten children in the study were labeled ELL by the teachers. The teachers also reported that 2.09% of the preschool children and 5.76% of the kindergarten children had language delays; 5.24% of the preschool children and 9.95% of the kindergarten children were reported as having speech delays.

Stimuli

The Finish-a-Rhyme-Story assessment was composed of 24 items. Each item consisted of a four-line rhyme that tells a simple story. The words at the end of each line rhymed; the final word in the last line was filled in by the child. When given the opportunity—signaled by a pause and change in intonation—the child had the opportunity to produce a word that both rhymed and completed the four-line story. The stories were illustrated graphically by two simple black-and-white line drawings (see Appendix B for an example of Finish-a-Rhyme-Cloze illustrations). The first drawing represented the first three lines of the story. The second drawing represented the line that included the missing word.

The item drawings were arranged in assessment booklets that were used during administration to keep the child's interest and to provide a visual representation of the story. Though each child in the study was assessed with all 24 items, the items were divided into four sets (A through D) of six different items so that administration order could be randomized. The items in each set were arranged in an attempt to make the difficulty level of the sets equivalent.

Items were arranged based on a ranking obtained from the percentage of children who gave a semantic and phonological (S&P) response during the pilot study (see Table 1 for the items rankings based on the responses from the pilot study). Based on these rankings, items with a higher rank (indicating a higher percentage of S&P response) were assumed to be easier than items with a lower rank. The rankings of the items were then used to balance the difficulty levels of the items within each set. Therefore, each set included items that range from easy to more difficult with the first and last items in each set being items that were predicted to be easier than those in the

middle (see Table 2 for the balance of item rankings and the order of items used in the current study). The items were arranged in this way in hopes that the assumed easier beginning and ending questions would help the child build confidence in the beginning of the task and would allow the child to finish the task with a sense of success.

Procedures

Administrators and Training

The sets of Finish-a-Rhyme-Story items were administered by trained undergraduate and graduate level students as well as two professors in the School of Education. The administrators participated in a 2-hour training meeting conducted by the author (i.e., a graduate student who was involved in the field test) and two professors in the School of Education who also were involved in the field test. During the training, administrators were trained as both ‘readers’ and ‘recorders’ (the specific jobs of each position will be discussed in the Test Administration section). The administrators reviewed the assessment protocol, learned the administration procedures, and practiced administering the assessment. Various scenarios were presented to the administrators and group discussions aided in clarifying the procedure for handling such situations as multiple responses, absent participants, children who do not respond to an item, and talkative children who continuously go off topic. Training also incorporated a comprehensive discussion and step-by-step guidelines on data collection procedures. The administration training meeting was used to increase the consistency of administrator performance by ensuring that all were trained in the same way. This consistency was monitored as the author periodically attended data collection sessions to examine the administrators’ performance, answer questions, and ensure that protocol

was being followed; in rare cases where protocol needed to be clarified, the author met with the administrators to review procedures.

Test Administration

All 24 items were administered to each child in the study; this was done in two sessions in which 12 items (two sets of 6 items each) were administered at one time; the four sets were administered in a random order. The assessment booklet of pictures was situated so that the child could see the pictures while listening to the stories.

The assessment was administered in a quiet setting in the child's preschool or kindergarten classroom during school hours. During each of the two administration sessions, each child was pulled out of class for approximately ten minutes to participate in the task. Administration was carried out in a two-on-one situation where one administrator (i.e., the reader) read the items to the child while a second administrator (i.e., the recorder) recorded the child's responses. During administration, the reader sat in a chair across from the child. The recorder also sat at the table, but to the side.

As each child followed the reader to the table, the reader engaged in an introductory conversation with the child and explained that they would be reading some stories together. This conversation was meant to make the child feel at ease so as to establish a comfortable situation. When they arrived at the table, the reader introduced the recorder as a friend who would be helping with the stories. This procedure was implemented because during the earlier field test, it was found that this two-on-one protocol worked well as it created a more naturalistic interaction where the reader was able to focus all attention on the child without the distraction of breaking the flow of the interaction to record the child's responses. This allowed the reader to concentrate solely on using an engaging manner to present the assessment.

Table 1

Pilot Study Item Response Ranking

Pilot Rank	Item	Pilot Rank	Item	Pilot Rank	Item
1	nest	9	fun	17	log
2	head	10	hair	18	sock
3	hot	11	ring	19	pet
4	eat	12	cake	20	fall
5	fit	13	shoe	21	stuck
6	mad	14	face	22	gone
7	sleep	15	bed	23	flat
8	sad	16	crawl	24	cat

Note. A higher rank denotes a larger percentage of preschool and kindergarten children who gave a semantic and phonological (S&P) response during the pilot study.

Table 2

Balancing and Ordering of Items Within Sets

Balancing of Items			
Set A	Set B	Set C	Set D
1 (nest)	2 (head)	3 (hot)	4 (eat)
8 (sad)	7 (sleep)	6 (mad)	5 (fit)
9 (fun)	10 (hair)	11 (ring)	12 (cake)
16 (crawl)	14 (face) ^a	15 (bed) ^a	13 (shoe)
17 (log)	18 (sock)	19 (pet)	20 (fall)
24 (cat)	23 (flat)	22 (gone)	21 (stuck)

Order of Items			
Set A	Set B	Set C	Set D
8 (sad)	7 (sleep)	6 (mad)	5 (fit)
17 (log)	18 (sock)	19 (pet)	20 (fall)
1 (nest)	2 (head)	3 (hot)	4 (eat)
24 (cat)	23 (flat)	22 (gone)	21 (stuck)
16 (crawl)	14 (face)	15 (bed)	13 (shoe)
9 (fun)	10 (hair)	11 (ring)	12 (cake)

Note. Items were balanced between sets to give each set a range of item difficulty; items with a higher rank were assumed to be easier because a larger percentage of preschool and kindergarten children gave a semantic and phonological (S&P) response to these items during the pilot study. Items were ordered within each set based on the rank; the rank order was 2nd most difficult followed by the 5th, 1st, 6th, 4th, and 3rd, respectively.

^aThe face and bed items were switched in order to keep Set B from having 2 items with the rime –ed.

Demonstration item. Once attempts were made to make the child as comfortable as possible with the reader, the reader explained that the stories they would be reading had some words that sounded the same and that the child would get to help finish each of the stories (see Appendix C for the complete assessment administration protocol). The reader then introduced the child to a demonstration item so the child would be familiar with the way in which the stories worked. During the demonstration item, the reader read the story aloud while showing the child the first drawing that accompanied the item. Prior to producing the final word in the short rhyme poem, the reader paused for a second and then completed the story with a word that both rhymed and fit the story semantically. After filling in the final word, the reader showed the child the second picture which represented the final word. The reader then repeated the entire story to ensure that the child understood that the story was complete.

Sample item. After presenting the demonstration, the child participated in a sample item. The sample item was administered in the same manner as the actual assessment items; the reader read the story and paused to allow the child to produce the final word. Once the child produced a word, the reader showed the child the second picture. If the child responded with an answer that fit the story both semantically and phonologically, the reader verbally reinforced the response by commenting on the fact that the child's response completed the story and sounded like the other words in the story. The reader then proceeded to the assessment items. If the child responded to the sample item, the reader did not give negative feedback, but simply acknowledged the child's response and then re-read the story and filled in the response that fit semantically and phonologically. The reader then explained to the child that this word

both finished the story and sounded the same as some of the other words in the story; this explanation was only used during the demonstration and sample items (see Appendix C for an explanation in the assessment administration protocol).

Assessment items. Once the demonstration and sample items were administered, the reader proceeded to administer the assessment items in the same way as the sample item. However, during the assessment, if the child gave an answer that did not fit both semantically and phonologically, the reader did not repeat the item. For all responses, the child's answer was recorded word-for-word and then the reader continued on to the next item. Though each response was kindly acknowledged during assessment, the reader did not give specific feedback about the correctness of the child's response. However, the second picture of the item did provide the child with some feedback because this picture, which represented the intended semantic and phonological response, was shown to each child regardless of their response. The administrator would also say the intended response so that the child was able to hear the rhyme. Therefore, children who did not produce the intended response would see a picture of the intended response and hear the word that fit the story semantically and phonologically; this was done in a playful manner and was not elaborated on or explained in a way that provided additional instruction.

Scoring

Children's responses to each item were scored by the author. Inter-rater reliability between the author and a member of the thesis committee was found to be 93.98% based on a random sample of approximately 10% ($n = 63$ students) of the participants. The scoring of the items was based on the type of response given. For each Finish-a-Rhyme-Story item there was a specific intended response that fit the story both

semantically and phonologically. Children's responses were analyzed in terms of the semantic and phonological fit with the items. Responses were scored as one of five different types: incorrect, semantic but not phonological (S-notP), phonological but not semantic (P-notS), semantic *and* phonological (S&P), or repeat.

Incorrect responses were responses that did not fit the task either semantically or phonologically. For example, on item A1, *This is Jake, he likes to bake. What should he make? How about a ___* an incorrect response was *hat*. This response was seemingly random and has no semantic or phonological relation to the content. Non-responses and responses of *I don't know* were also scored as incorrect.

Semantic but not phonological (S-notP) responses were words that fit the context of the item but did not fit the rhyme. A response of *pie* was considered an S-notP response to item A1 (listed above). Obviously this response completed the meaning of the story, but did not complete the story with a rhyming word.

Phonological but not semantic (P-notS) responses were words that completed the rhyme but did not fit semantically. For example, *snake* was considered a P-notS response to item A1 because although this response rhymed with the other words in the story, it did not correctly complete the meaning of the story.

Semantic and phonological (S&P) responses to Finish-a-Rhyme-Story items were the words that completed the item by rhyming with the other words and fit the context of the story. In some cases, it was possible for children to come up with a word other than the intended word that also met the criteria of rhyming and completing the meaning of the story. These responses were also counted as S&P. For example, for item

D6, *My apples went bad, I'm not glad. I'll go tell my dad, that I am very ____*, the intended response was *sad* but *mad* was also counted as S&P.

Repeat responses were instances where rather than generating a word to complete the story, children simply repeated one of the rhyming words within the story. Though this type of response was rare in the pilot study, it was observed.

Data Analysis

Data analysis tests such as Item Response Theory (IRT) analysis have been created to measure underlying or latent abilities (Baker, 1985; Henard, 2000; McKinley & Mills, 1989). IRT is beneficial because it has the ability to examine items individually or as a whole (Baker, 1985; Bond & Fox, 2007; McKinley & Mills, 1989). For this data set, an IRT Rasch analysis was combined with some classical test theory (CTT) methodologies to measure the items as a whole including difficulty ranking and differential item function (DIF) analysis. Upon completion of the validation of the items, several analyses of group differences were conducted using analysis of variance (ANOVA).

Results

Data analysis provided several ways of examining the information obtained during this study. The scores were used to analyze the items as a whole and determine the way in which certain groups of children handled the items (i.e., preschool versus kindergarten children and ENG versus ELL children). Results are presented below.

Item Performance

In order to gain an understanding of how the items perform, IRT Rasch analysis was used to provide a difficulty level for each item and an ability level on this task for each child. That is, Rasch analysis was used to calculate a statistical difficulty level associated with each item; due to the theoretical nature of Rasch analysis, this level is a fixed difficulty level and was not dependent upon the sample from which it was generated. The student ability levels for this task relate to the fixed item difficulty levels; the ability levels estimate student rhyming ability based on the Finish-a-Rhyme-Story items. To make comparisons between the difficulty level of the items and the student ability levels on this task clear, the Rasch model was programmed to impose a mean of 50 and a standard deviation of 10 on the item difficulty mean. The following results report findings on item performance.

Item Difficulty

A difficulty measurement for each item was calculated based an imposed mean ($M = 50$ and $SD = 10$); items with higher means were more difficult than items with lower means. The items ranged in difficulty with item D4 as the most difficult ($M = 72.45$) and item C1 as the easiest ($M = 30.32$) (see Table 3 for difficulty ranking of the items and Appendix A for the text of each item). In addition to difficulty

measurements, infit and outfit, discrimination statistics, and internal consistency were calculated.

Infit and Outfit. The Rasch model calculated infit and outfit statistics. Infit and outfit examined how well the items fit within the theoretical Rasch model and how well the items matched the ability levels of the given sample. Ideally, infit and outfit should be less than 1.5. Though one item hit the upper limit (A4 outfit = 1.54), all items were within the parameters of infit and outfit; mean infit and outfit values were strong (infit $M = 1.00$, outfit $M = 1.01$) indicating that the items fit the theoretical model and matched the ability levels of the sample. In addition, discrimination statistics were calculated using CTT statistics.

Discrimination. The CTT statistics of point-biserial correlation and discrimination estimation based on a 2-parameter model were used to measure discrimination (see Table 4 for the statistics of each item). Point-biserial correlation measured the extent to which the sample's performance on a given item is consistent with their performance on the items as a whole; essentially, the point-biserial correlation determined whether each item functioned properly compared to how the sample performed on the other items. The calculated point-biserial correlations were greater than .3, as desired, for all items in the study. However, two items, A4 and D4, had point-biserials that were very close to the desired value (i.e., A4 = .33, D4 = .33).

Internal consistency. The internal consistency of the items was measured using Cronbach's alpha. This correlation was strong for the combined sample of preschool and kindergarten children ($\alpha = .90$). The items also showed strong reliability (reliability = .85) based on a Rasch model reliability test that excluded the 17 children who

Table 3

Item Difficulty Ranking based on Rasch Model Analysis

Difficulty Level	Item			
Challenging	D4 (stuck)			
	A4 (cat)	B4 (flat)		
	B2 (sock)	B5 (face)		
	A5 (crawl)	C2 (pet)		
	C4 (gone)			
	D2 (fall)			
	A2 (log)			
	Average	C5 (bed)		
		D6 (cake)		
		A3 (nest)	B3 (head)	
B1 (sleep)		C3 (hot)	C6 (ring)	
A1 (sad)				
D1 (fit)				
A6 (fun)		D5 (shoe)		
B6 (hair)				
D3 (eat)				
Easy		C1 (mad)		

Note. Difficulty is based on $M = 50$ and $SD = 10$; higher mean is more difficult. See Appendix A for complete item texts.

Table 4

Item Performance Based on Rasch Model Analysis

Item	Difficulty	Infit (<i>MS</i>)	Outfit (<i>MS</i>)	Correlation	Discrimination
A1 (sad)	39.73	.88	.80	.59	1.14
A2 (log)	54.57	.87	.79	.58	1.26
A3 (nest)	47.42	.89	.76	.59	1.19
A4 (cat)	69.19	1.10	1.54	.33	.77
A5 (crawl)	60.89	1.03	1.25	.45	.93
A6 (fun)	34.97	.99	1.01	.51	1.00
B1 (sleep)	44.67	1.11	1.08	.47	.86
B2 (sock)	63.34	1.05	1.07	.43	.91
B3 (head)	47.31	.99	.96	.53	1.02
B4 (flat)	68.51	1.01	.97	.40	.98
B5 (face)	63.13	.93	.83	.50	1.16
B6 (hair)	33.29	.79	.61	.61	1.20
C1 (mad)	30.32	1.05	.94	.48	.97
C2 (pet)	62.11	.92	.81	.52	1.18
C3 (hot)	44.31	1.06	.97	.51	.95
C4 (gone)	59.98	1.05	.96	.46	.92
C5 (bed)	49.81	1.15	1.25	.43	.73
C6 (ring)	44.67	.91	.86	.58	1.12
D1 (fit)	38.04	.82	.72	.61	1.19
D2 (fall)	56.18	1.07	1.11	.45	.85
D3 (eat)	31.85	.91	.78	.54	1.08
D4 (stuck)	72.45	1.07	1.42	.33	.87
D5 (shoe)	34.31	1.16	1.47	.42	.82
D6 (cake)	48.95	1.06	1.19	.49	.88

Note. Difficulty is calculated based on $M=50$ and $SD=10$. Discrimination is estimated based on a 2-parameter model.

obtained either minimum or maximum scores. Overall, the Rasch model and CTT tests produced results that fell within standard statistical ranges.

Differential Item Function

Item function was also examined using differential item functioning (DIF) which provided information about the way each item performed with preschool and kindergarten children. This was accomplished by comparing the scores of preschool and kindergarten children who were matched for ability level on the task. Comparing children matched for ability level made it possible to detect differences in performance on certain items that may indicate that something in the item content unfairly biases members of one group (i.e., biases either preschool or kindergarten children); therefore, items with low probabilities need to be examined carefully for possible bias. The majority of the items had Welch and Mantel-Haenszel probabilities well above .05. However, there were five items with Welch and Mantel-Haenszel probabilities lower than .05: A6, B4, C5, D2, and D4 (see Table 5 for all Welch and Mantel-Haenszel probabilities) indicating that something about these items causes preschool and kindergarten children who were matched for ability level to perform differently.

Student Ability

The student ability levels produced by the Rasch analysis provided a more in-depth way of examining the items because they presented data on the way in which this sample responded to the items. The student ability scores estimate student rhyming ability based on the items. As mentioned above, the analysis made comparisons based on an imposed scale ($M = 50$ and $SD = 10$). The Rasch analysis showed that the combined sample (i.e., preschool and kindergarten children) achieved a mean ability score ($M = 54.44$, $SD = 17.81$) slightly higher than the imposed item mean ($M = 50$, SD

= 10). The discrepancy between the student sample mean and the item mean was small (i.e., within one standard deviation) indicating that the items were well matched to the ability level of the students.

The Rasch model also used the ability score for each child to examine the consistency of the items. The individuals in the sample received a relatively low mean standard of error ($MSE = 5.98$) for the ability scores the task. However, as was expected, children with very high or very low ability scores (i.e., those that were far from the mean) tended to have a larger standard error because there are not as many items in their ability range so calculating their ability is less accurate. It is interesting to note that three children, all ENG kindergarteners, achieved the maximum score on the items (i.e., produced an S&P response for all 24 items). Fourteen children, 13 of which were preschoolers, received the minimum score (i.e., did not produce an S&P response for any of the 24 items).

Preschool versus Kindergarten Performance

By nature, the development of rhyme awareness in preschool and kindergarten should differ. To determine if the 24 Finish-a-Rhyme-Story items tapped into the developmental difference between the two groups, performance of preschool and kindergarten children was compared using a one-way analysis of variance (ANOVA) (see Table 6 for a comparison of preschool and kindergarten performance). It was found that kindergarten children were more likely to produce S&P responses ($M = 15.76$, $SD = 5.08$) than preschool children ($M = 10.33$, $SD = 5.96$); this difference was statistically significant, $F(1, 588) = 135.22$, $p < .001$, $d = .983$. Kindergarten children also produced P-notS responses ($M = 0.80$, $SD = 1.12$) significantly more often than preschool children ($M = 0.57$, $SD = 1.04$), $F(1, 588) = 6.09$, $p = .014$, $d = .213$. That is,

Table 5

Differential Item Functioning of Preschool and Kindergarten Performance Measured by Welch (p_W) and Mantel-Haenszel (p_{MH}) Probabilities

Item	p_W	p_{MH}	Item	p_W	p_{MH}
A1	.567	.708	C1	.224	.189
A2	.604	.593	C2	.203	.434
A3	.439	.965	C3	.051	.100
A4	.531	.081	C4	.713	.396
A5	1.000	.886	C5	.013*	.180
A6	.001***	.001***	C6	.656	.282
B1	.656	.877	D1	.248	.943
B2	1.000	.714	D2	.007**	.027*
B3	.391	.335	D3	.725	.590
B4	.028*	.014*	D4	.037*	.156
B5	.265	.724	D5	.117	.689
B6	.315	.757	D6	.9211	.897

* $p < .05$. ** $p < .01$. *** $p \leq .001$.

kindergarten children produced more S&P and P-notS responses than preschool children. Conversely, preschool children were significantly more likely to produce S-notP responses ($M = 4.98, SD = 3.52$) than kindergarten children ($M = 3.33, SD = 3.24$), $F(1, 588) = 32.42, p < .001, d = .488$. Preschool children were also significantly more likely ($M = 6.62, SD = 6.26$) to produce incorrect responses than kindergarten children ($M = 2.76, SD = 3.46$), $F(1, 588) = 93.32, p < .001, d = .794$. There was no significant difference between preschool ($M = 1.50, SD = 1.46$) and kindergarten ($M = 1.35, SD = 1.33$) children's production of repeat responses, $F(1, 588) = 1.56, p = .21, d = .108$.

Monolingual English Speakers versus English Language Learners Performance

Because of the large number of ELL children included in the study, two separate ANOVAs were conducted to reveal differences between the performance of ENG and ELL children. The first ANOVA considered only language classification and not grade level. That is, the total sample was separated into two groups: ENG children and ELL children (see Table 6 for a comparison of ENG and ELL performance). It was found that ENG children ($M = 15.56, SD = 4.98$) produced S&P responses significantly more than ELL children ($M = 7.98, SD = 5.46$), $F(1, 588) = 228.45, p < .001, d = 1.542$. In fact, no ELL child produced S&P responses for all 24 items; the highest number of S&P responses produced by an ELL child was 20. In addition, the minimum number of S&P responses produced by ENG kindergarten children was four, while the minimum number of S&P responses produced by ELL kindergarten children was zero. ENG children also produced P-notS responses ($M = 0.80, SD = 1.14$) significantly more often than ELL children ($M = 0.41, SD = 0.87$), $F(1, 588) = 13.30, p < .001, d = .388$. However, ELL children produced S-notP responses ($M = 5.39, SD = 3.30$) significantly more than ENG children ($M = 3.48, SD = 3.35$), $F(1, 588) = 33.76, p < .001, d = .574$.

That is, ENG children produced more P-notS responses and ELL children produced more S-notP responses. ELL children were also significantly more likely to produce repeat and incorrect responses than ENG children. It is important to note that every ELL preschool child gave at least one incorrect response.

A second ANOVA looked at language classification as an independent variable within the kindergarten and preschool groups. This allowed for examination of performance differences between ENG and ELL preschool children and between ENG and ELL kindergarten children (see Table 7 for a comparison of ENG and ELL in preschool and kindergarten). Based on the data presented above, it was expected that ENG preschool children would perform better (i.e., give more S&P responses) than ELL preschool children and that ENG kindergarteners would perform better than ELL kindergarteners. The ANOVA revealed that kindergarten children did follow the predicted trend: ENG kindergarteners produced S&P responses ($M = 17.31$, $SD = 3.76$) significantly more often than ELL kindergarteners ($M = 9.89$, $SD = 5.17$), $F(1, 380) = 208.61$, $p < .001$, $d = 1.662$. ENG kindergarteners also produced significantly more P-notS responses than ELL kindergarteners. However, ELL kindergarten children gave incorrect, S-notP, and repeat responses significantly more often than their ENG counterparts. Interestingly, preschool children did not follow these same trends.

ENG preschool children produced significantly more S&P responses ($M = 12.13$, $SD = 5.29$) than ELL preschool children ($M = 5.11$, $SD = 4.59$), $F(1, 205) = 74.07$, $p < .001$, $d = 1.421$. And ELL preschoolers produced significantly more incorrect response ($M = 12.08$, $SD = 6.91$) than ENG preschoolers ($M = 4.75$, $SD = 4.76$), $F(1, 205) = 73.08$, $p < .001$, $d = 1.256$. However, there was no significant

difference in their production of P-notS, S-notP, or repeat responses. This shows that the difference between ENG and ELL preschoolers was not as great as the difference between ENG and ELL kindergarteners.

Table 6

*Analysis of Variance (ANOVA) Comparing Preschool and Kindergarten**Performance and Monolingual English Speakers (ENG) and English Language**Learners (ELL) Performance*

Response	Preschool		Kindergarten		<i>df</i>	<i>F</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
S&P	10.33	5.96	15.76	5.08	1	135.22***	<.001	.983
P-notS	0.57	1.04	0.80	1.12	1	6.09**	.014	.213
S-notP	4.98	3.52	3.33	3.24	1	32.42***	<.001	.488
Repeat	1.50	1.46	1.35	1.33	1	1.56	.212	.108
Incorrect	6.62	6.26	2.76	3.46	1	93.32***	<.001	.794

Response	ENG		ELL		<i>df</i>	<i>F</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
S&P	15.56	4.98	7.98	5.46	1	228.45***	<.001	1.542
P-notS	0.80	1.14	0.41	.87	1	13.30***	<.001	.388
S-notP	3.48	3.35	5.39	3.30	1	33.76***	<.001	.574
Repeat	1.34	1.36	1.62	1.45	1	4.02*	.045	.199
Incorrect	2.81	3.51	8.59	6.50	1	180.81***	<.001	1.155

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Analysis of Variance (ANOVA) Comparing Monolingual English Speakers (ENG) and English Language Learners (ELL) in Preschool and Kindergarten

Response	ENG vs. ELL Comparison					
	<i>M</i>	<i>SD</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>d</i>
Preschool						
S&P						
ENG	12.13	5.29	1	74.07**	<.001	1.421
ELL	5.11	4.59				
P-notS						
ENG	.59	1.05	1	.37	.55	.097
ELL	.49	1.01				
S-notP						
ENG	5.07	3.57	1	.44	.51	.106
ELL	4.70	3.42				
Repeat						
ENG	1.46	1.47	1	.48	.49	.110
ELL	1.62	1.44				
Incorrect						
ENG	4.75	4.76	1	73.08**	<.001	1.256
ELL	12.08	6.91				
Kindergarten						
S&P						
ENG	17.31	3.76	1	208.61**	<.001	1.662
ELL	9.89	5.17				
P-notS						
ENG	.91	1.18	1	15.81**	<.001	.097
ELL	.36	.77				
S-notP						
ENG	2.67	2.93	1	72.37**	<.001	1.044
ELL	5.85	3.16				
Repeat						
ENG	1.28	1.29	1	3.85*	.050	.241
ELL	1.61	1.45				
Incorrect						
ENG	1.82	2.06	1	145.07**	<.001	1.254
ELL	6.29	5.07				

* $p < .05$. ** $p < .001$.

Discussion

The results of this study suggest that as a whole, the Finish-a-Rhyme-Story items functioned as expected; the items revealed children with both high and low level rhyming ability based on this task. The following section discusses item performance and discriminative power.

Item Performance

When looking at the items as a whole, it was found that the combined preschool and kindergarten sample achieved a mean that was a bit higher than the item mean. The small difference in means suggests that the items were well matched, if not slightly easy for the group; this result is not surprising considering that a larger number of kindergarten children ($n = 382$) compared to preschool children ($n = 207$) was included in the sample. However, the fact that the student ability mean was so close to the item difficulty mean is an indication of the strength of the items. This indicates that the items were well matched to the sample; the means were close therefore the items were not too easy or too difficult. These findings suggest that as a whole, the items are appropriate for preschool and kindergarten children.

In judging the consistency of the items, it was found that there was a low mean standard of error for the individual scores on the 24 items. The relatively low standard of error indicates that the items fit the sample well and have good consistency of measure. If the standard of error had been high, it would suggest that the items did not appropriately fit the ability level of the sample; however, this was not the case. Notably, children with scores that were very different from the mean (i.e., very high or very low) had larger standards of error; this typically occurs at the farthest ends of a scale where there are not enough items in these extreme ability levels to provide more accurate

scores. This finding should not be of great concern because it is often the case when assessing ability levels (Bond & Fox, 2007). However, if these items are to be used as an assessment in the future, it would be important to remember that the scores of children in the highest and lowest ability levels would not be as accurate as the scores of children with average ability.

Though children with very high and very low ability levels did not have many items in their ability range, the analysis showed various levels of difficulty among the items. In fact, the items spread fairly equally around the average difficulty level providing a wide range of difficulty. Furthermore, statistical tests produced fit and correlation results within standard statistical ranges indicating that the items are functioning appropriately for young children. Relating these findings to specific items provides a more concrete understanding of how the Finish-a-Rhyme-Story items are functioning with preschool and kindergarten children.

The easiest item was C1 (*This is Chad, he's being bad. Don't tell Dad, cause he'll get ____*). In this item, story content and vocabulary may have contributed to the low difficulty level. For example, the story is about a boy being bad and dad getting mad which is fairly common and concrete for young children. Also, because there are not many other words in young children's vocabulary that could complete this story, it is possible that *mad* was the first word children thought of regardless of whether or not they were aware of the rhyme. This means that children who may not yet have strong rhyme skills could have produced an S&P response to this item without necessarily being aware of the rhyme.

Story content and vocabulary also may have affected the difficulty level of item D4 (*This little duck, has run out of luck. He walked in the muck, and now he is ____*). However in D4, these elements may have played a role in making the item the most difficult item of the group. The story content of D4 could be explained as a cause and effect. Though this is also true of item C1, the cause and effect in C1 is very familiar to young children: if you are bad, parents get mad. The cause and effect depicted in D4 is not as common and is dependent on children also knowing the vocabulary. That is, in order to understand the cause and effect of D4, children must know the vocabulary word *muck*; they need to understand that muck is like mud and is not only dirty (which was a common response to this item), but it can also be difficult to walk through. Thus *muck* may cause a duck to get stuck. Therefore, the vocabulary word *muck* and the cause and effect story line may be the reason D4 is more challenging than other items. Though D4 is more challenging, it was not impossible. Children were able to produce S&P responses to D4 making this item one that may help discriminate between high and low ability levels based on this task.

One purpose of this study was to determine which of the 24 items were able to discriminate between children with high and low level rhyming performance on the Finish-a-Rhyme-Story items. In determining this, it was important to establish which items are consistent and functioned appropriately. The strong Cronbach's alpha for the 24 Finish-a-Rhyme Story items points to good internal consistency. The point-biserial correlations within standard statistical ranges indicate that the items functioned properly compared to how the sample performed. However, it is important to note that items A4 and D4 had low point-biserial correlations. This suggests that items A4 and

D4 may not measure ability as well as other items. As noted above, D4 is the most difficult item and may have a low point-biserial because of the smaller number of children who were able to produce S&P responses; this finding should not eliminate D4 from future use. Item A4 was also a very difficult item; it received many repeat responses. The high level of difficulty and tendency for repeat responses for A4 likely stems from two things: story content and word usage.

The story in A4 depicts a rat that should be afraid of a cat (see Appendix A for the complete item). However, young children are likely more familiar with a similar association between cat and mouse and may not generalize this association to cat and rat, thus causing confusion about the content of the story. Also, *rat* was used twice in the story (i.e., in the first line and third line). In other stories rhyme words were not typically re-used; the fact that *rat* was used twice may have caused a difference in the way children responded. A similar situation occurred in D5.

It was not realized until scoring, but in D5 the intended response was actually a repeat of a word in the story thus causing S&P responses to also be repeat responses (see Appendix A for the complete item). Interestingly, the statistical analysis did not indicate any problems with D5, which was one of the easier items. Perhaps the low difficulty level of D5 is due to the fact that S&P responses were prompted because children heard the intended response earlier in the story. The unusual situation of having rhyme words that are duplicated within the story is something that may need to be considered in future research.

To further investigate the items, differential item functioning (DIF) was used to compare the scores of preschool and kindergarten children who were matched for

ability level. The Welch and Mantel-Haenszel probabilities were high for the majority of the items indicating that preschool and kindergarten children with similar ability levels responded to the items as expected. Five items – A6, B4, C5, D2, and D4 – produced low Welch and Mantel-Haenszel probabilities suggesting that on these items preschool and kindergarten children with similar ability levels did not perform in the same way (i.e., they did not perform as expected). This may indicate some type of bias in these items that caused the two groups to respond differently. It is likely that a developmental difference causes the bias; perhaps these items have story content or vocabulary that are more familiar to kindergarten than preschool children, even when they demonstrated similar ability on other items. With this age group, it is appropriate to continue to use items with these differences in order to create a scale that is broad enough to measure the ability of high functioning preschoolers and low functioning kindergartners. Therefore, these items should continue to be used in future research.

Preschool versus Kindergarten Performance

It was no surprise that preschool and kindergarten children performed differently; these findings were expected and the results were strong. The fact that kindergarten children produced S&P and P-notS responses more often than preschool children suggests that kindergarten children were more aware of the rhyme component of the items; even when they did not give S&P responses, kindergarten children were more likely to produce a rhyming word (i.e., a P-notS response) than a word that fit semantically. In contrast, preschool children appeared to be more aware of the semantic component of the stories as they produced more S-notP responses than kindergarten children. Though these findings support the findings from the pilot study, it is important

to address the fact that the preschool and kindergarten children in this study did not come from completely similar backgrounds.

As mentioned in the Methods section, the kindergarten children used in the study were from public schools with low SES and high ELL populations. While many of the preschool children were also from similar schools, nearly half of the preschool children were from home-based preschools. Unfortunately, information about the makeup of the home-based preschools was not as readily available as in the public schools, but it is realistic to say that the children from the home-based schools were less likely to be ELL and more likely to come from higher SES backgrounds. Because the Rasch model was used to analyze the items, the difference in the preschool and kindergarten groups should not have affected the results for item difficulty levels. However, the difference may have influenced the relationship between preschool and kindergarten performance. In actuality, the demographic differences strengthen the results because if a more similar sample of preschool and kindergarten children had been used the difference between their ability levels would probably have been greater since it is likely that the preschool children from home-based preschools performed better than their peers from public preschools.

Monolingual English Speakers versus English Language Learners Performance

The high percentage of ELL children in this study created opportunities to not only examine the way the items worked with groups of different ages but also with groups from different language backgrounds. As in the preschool versus kindergarten comparison, it was found that one group out-performed the other in production of S&P and P-notS responses: ENG children produced more S&P and P-notS responses. These results suggest that ENG children are more aware than ELL children of the rhyme

component. This result is not surprising considering the primary language of the majority of the ELL children in the study was Spanish and rhyme is not a stressed and salient phonological unit in Spanish (Gorman & Gillam, 2003). Conversely, ELL children produced more S-notP responses and therefore showed more awareness of the semantic component. This same pattern was revealed when looking only at ENG and ELL kindergarten children. However, the comparison of ENG and ELL preschool children did not yield the same result.

While ENG preschool children produced more S&P responses and ELL preschool children produced more incorrect responses, there was no significance between the P-notS, S-notP, and repeat responses. This unexpected result seems to suggest that ELL preschool children are not lagging behind their ENG counterparts in development of rhyme awareness as much as ELL kindergarteners are lagging behind theirs.

Possible Factors Confounding Performance

The results of this study are very promising for the discriminative power of the Finish-a-Rhyme-Story items. However, there are several possible confounding factors that may have influenced the results. These factors may be caused by characteristics of the items, characteristics of the young sample, or characteristics of the administration.

Item Characteristics

As suggested in the Review of Literature, rhyme cloze items have several possible confounding factors. Though these factors were taken into consideration when creating the Finish-a-Rhyme-Story items, there were still a few pieces of evidence that suggest these factors may have had an influence on the responses. Two of these factors, story content and vocabulary, were discussed above. A third possible confounding

factor may stem from the fact that some items had a greater range of semantic responses which may or may not have influenced the difficulty level of the item. The difference in the number of possible semantic responses is apparent when comparing items such as A3 and B3 (see Appendix A for the complete item), both of which earned the same difficulty level just below average.

The story in A3 required children to tell where a robin goes to rest; other than the intended *nest*, there were not many semantic response options for this item (e.g., *house* or *bed*). In contrast, B3 required children to complete a story about what Ed hurt when he fell out of bed; this had many semantic possibilities other than the intended *head* and thus received a wide range of S-notP responses: *elbow, knee, foot, face, etc.* Because these items achieved the same difficulty level, it is hard to draw conclusions about whether or not the number of possible semantic responses influenced them. However, it would seem that items with multiple semantic possibilities provide more cognitive distracters that children must sort through in their minds in order to produce an S&P response. Thus B3 may require more cognitive sorting and may actually be more challenging than A3. In future research it could be interesting to investigate the difference between items with various ranges of possible semantic responses. Item characteristics such as story content, vocabulary, and various semantic possibilities were not the only possible confounding factors. Characteristics of the young sample also may have influenced the results.

Sample Characteristics

Young children are unpredictable and interactions with them often entail unexpected occurrences; this study was no exception. Administrators reported many stories of unexpected responses. Several of these anecdotes illustrate possible factors

that could have influenced performance. For example, it is not uncommon for young children to continue a behavior that has received a humorous reaction. This was true of one preschool boy who continued to give the response *poop* after the administrator giggled at this response for item A2 which occurred early in the administration. Thus, all subsequent answers of *poop* or *pooped* were not a response to the subsequent items, but rather a response to the administrator's reaction. Though this child may have had rhyming ability, his distracted responses in this situation made it impossible to accurately measure his ability level. This anecdote is a reminder that young children's maturity level and unpredictable nature must be considered when assessing their abilities.

Another characteristic of young children is occasional shyness around unfamiliar adults. Throughout the study, it was not uncommon to find children who were slow-to-warm or even unresponsive to the administrators. Though the administrators expected and prepared for this by conveying kindness and encouragement, some children were reluctant to respond. However, in some cases, administrators were able to entice responses from these children by giving them prompts. The prompts were either verbal (e.g., asking children "What do you think?") or visual (i.e., showing the child the second picture). Because not all children received prompts, any response following a prompt was not eligible for scoring; administrators were trained to record responses prior to a prompt. That is, if a child gave no response before a prompt but was able to produce a response after seeing the second picture the administrator recorded the response as *no response*. Though it was not required, some administrators not only recorded responses prior to a prompt but also those after a

prompt. This information could not realistically be used to draw conclusions because not all administrators recorded responses after prompts; however, it is interesting to look at the responses given after prompts because for some items nearly 24% of children who provided non-S&P responses were able to produce S&P responses after a prompt was given. Perhaps this information may be useful in future research that examines use of these items in an instructional setting; future research could compare pre- and post-prompt responses or could examine if children with low level abilities are able to produce S&P responses to difficult items when given a prompt.

In some cases, use of a visual prompt was initiated by administrators to persuade shy children to respond. In other cases, children initiated a visual prompt by demonstrating the desire to see the second picture before giving a response. Interestingly, these occurrences happened with preschool and not kindergarten children. Preschool children demonstrated their desire to see the second picture in several ways: some tried to reach across the table and turn the page themselves while others said “open it” instead of providing an actual response. The preschool children’s desire to see the second picture before providing a response may show an immaturity in their ability to understand the task as the administrators demonstrated that a response was to be given before the second picture was revealed; kindergarten children seemed to understand response protocol more clearly. As discussed above, only responses prior to prompts were eligible for scoring. Because of the way in which these responses had to be scored, children who may have been able to provide an S&P response were scored as *incorrect* because they were distracted by the need to turn the page and see the second picture; this obviously has an influence on the results. One unfortunate outcome

of this situation is that the picture booklets which were meant to be a child-friendly aid were actually more of a distraction for some children; this showcases another challenge in assessing young children and provides circumstances that future researchers may consider.

While some children struggled to provide a response, other children were very successful. An interesting event that occurred often with both preschool and kindergarten children was the phenomenon of actually seeing them think before responding. This was observed by many administrators: after the item was read, children would pause and think (usually demonstrated by looking up toward the ceiling) before giving a response. The thinking time only lasted seconds, but it was fascinating to watch and imagine what was going through their minds. Sometimes they would give one response and immediately change it to another, usually S&P, response. In these cases, if the second response was provided before the second picture was exposed the response was scored as S&P. If the Finish-a-Rhyme-Story items are used in the future, administrators should be prepared for responses of this type.

Administration Characteristics

Though administration was as standardized as possible, there are a few situations that may have affected the results. First, the items were administered by multiple people. While there does not appear to be a systematic error because the administrative pairs were randomized as much as possible, there is a small chance that results may have been influenced by the administrators. For instance, for reasons that can not be explained, children may have been more comfortable with some administrators than with others. However, the variations in administrative pairs should have eliminated this.

Unfortunately, another administrative factor, the setting of the administration, likely had more of an impact on the results. In most classrooms, the administrators met with each child in a quiet corner of the class where there was little interruption. However, in some classrooms this setting was not as quiet or private and some children in the class had the opportunity to linger within earshot of the administration. Though administrators did their best to keep this from occurring, in some situations there was nothing they could do; therefore some children may have overheard the items before it was their turn. Thus the responses from these children may have been influenced by how their classmates responded. In the future this is something that should be prevented.

Future Research

In the previous sections, suggestions for future research have been given as minor changes to the current study. However, there are also some larger issues that future research could investigate. These issues address the *what*, *where*, and *how* of Finish-a-Rhyme Story items.

What role does the semantic component play in the Finish-a-Rhyme Story items? The current results suggest that different groups of children respond differently to these items. Future research may investigate whether or not this difference is influenced by the semantic component of the items. That is, does the story framework of the Finish-a-Rhyme Story items impact performance and if so, how? This may be addressed by comparing children's rhyme ability levels based on a rhyme cloze task compared to their ability levels based on a generative rhyme task with no semantic structure. Research that addresses the impact of the semantic component of these items

may help researchers and educators better understand where and how these items may be used in the future.

Where can Finish-a-Rhyme-Story items be used in preschool curriculum?

Future research that addresses the use of these items may help teachers find a way to utilize the items in the classroom. This utilization may take the form of an instructional tool or as an assessment. In any case, future research that investigates ways to use these items in the classroom has the potential to be beneficial to both teachers and students.

How do children's responses to Finish-a-Rhyme-Story items relate to later reading success? Comparing the early reading skills of children who gave mostly S-notP responses and children who gave mostly P-notS responses may reveal some interesting findings. This type of research could help researchers discover relationships between response types and later reading skills which may enable the items to be used in a type of predictive assessment.

By building upon the information gained in this study, future research has the potential to develop tools that may be used in the assessment and instruction of early rhyme skills. Tools of this kind could benefit both students and teachers.

References

- Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology, 96*(1), 43-55.
- Baker, F. B. (1985). *The basics of item response theory*. Portsmouth, NH: Heinemann.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor, MI: The University of Michigan Press.
- Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*(3), 419-421.
- Brady, S., Fowler, A., Stone, B., & Winbury, N. (1994). Training phonological awareness: A study with inner-city kindergarten children. *Annals of Dyslexia, 44*(1), 26-59.
- Brady, S., & Shankweiler, D. P. (1991). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bryant, P. (1998). Sensitivity to onset and rhyme does predict young children's reading: A comment on Muter, Hulme, Snowling, and Taylor (1997). *Journal of Experimental Child Psychology, 71*(1), 29-37.
- Bryant, P., MacLean, M., Bradley, L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology, 26*(3), 429-438.
- Byrne, B., & Fielding-Barnsley, R. (1993). Recognition of phoneme invariance by beginning readers: Confounding effects of global similarity. *Reading and Writing: An Interdisciplinary Journal, 5*(3), 315-324.
- Cardoso-Martins, C. (1994). Rhyme perception: Global or analytical? *Journal of Experimental Child Psychology, 57*(1), 26-41.
- Carroll, J. M., & Snowling, M. J. (2001). The effects of global similarity between stimuli on children's judgment of rime and alliteration. *Applied Psycholinguistics, 22*(3), 327-324.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45*(6), 1142-1157.

- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, 17(1), 205-215.
- Charles-Luce, J., & Luce, P. A. (1995). An examination of similarity neighborhoods in young children's receptive vocabularies. *Journal of Child Language*, 22(3), 727-735.
- Christie, J., Enz, B. J., & Vukelich, C. (2007). *Teaching language and literacy: Preschool through elementary grades* (Third Edition ed.). Boston: Pearson Education, Inc.
- Culatta, B., Hall, K., Kovarsky, D., & Theadore, G. (2007). Contextualized approach to language and literacy instruction. *Communication Disorders Quarterly*, 28(4), 216-235.
- Culatta, B., Setzer, L., Wilson, C., & Aslett, R. (2004). Project SEEL: Part III. Children's engagement and progress attainments. *Communication Disorders Quarterly*, 25(3), 127-144.
- Cunningham, J. W., Cunningham, P. M., Hoffman, J. V., & Yopp, H. K. (1998). *Phonemic awareness and the teaching of reading: A position statement from the Board of Directors of the International Reading Association*. Newark: DE: International Reading Association.
- DaFontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Reading and Writing: An Interdisciplinary Journal*, 7(1), 139-153.
- De Cara, B., & Goswami, U. (2003). Phonological neighbourhood density: Effects in a rhyme awareness task in five-year-old children. *Journal of Child Language*, 30(3), 695-710.
- Dorsey, M. E. (1972). *Reading games and activities*. Belmont, CA: Fearon Publishers.
- Dowker, A. (1989). Rhyme and alliteration in poems elicited from young children. *Journal of Child Language*, 16(1), 181-202.
- Ehri, L. C., & Snowling, M. J. (2004). Developmental variation in word recognition. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 433-437). New York: The Guilford Press.
- Fernandez-Fein, S., & Baker, L. (1997). Rhyme and alliteration sensitivity and relevant experiences among preschoolers from diverse backgrounds *Journal of Literacy Research*, 29(3), 433-459.
- Gillon, G. T. (2004). *Phonological awareness: From research to practice*. New York: The Guilford Press.

- Gorman, B. K., & Gillam, R. B. (2003). Tutorial for clinicians on phonological awareness development in Spanish. *Communication Disorders Quarterly*, 25, 13-22.
- Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology*, 42(1), 73-83.
- Goswami, U. (1988). Orthographic analogies and reading development. *The Quarterly Journal of Experimental Psychology*, 40A(2), 239-268.
- 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: The Guilford Press.
- Goswami, U., & Bryant, B. (1990). *Phonological skills and learning to read*. Hove, East Sussex, UK: Lawrence Erlbaum Associates.
- Goswami, U., & Bryant, B. (1992). Rhyme, analogy, and children's reading. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 49-63). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of Child Language*, 15, 395-410.
- Henard, D. H. (2000). Item response theory. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 67-97). Washington, DC: American Psychological Association.
- Justice, L. M., Invernizzi, M. A., & Meier, J. D. (2002). Designing and implementing an early literacy screening protocol: Suggestions for the speech-language pathologist. *Language, Speech, and Hearing Services in Schools*, 33(2), 84-101.
- Kieff, J. E., & Casbergue, R. M. (2000). *Playful learning and teaching: Integrating play into preschool and primary programs*. Boston: Allyn and Bacon.
- Layton, L., Deeny, K., Tall, G., & Upton, G. (1996). Researching and promoting phonological awareness in the nursery class. *Journal of Research in Reading*, 19(1), 1-13.
- Lipka, O., & Siegel, L. S. (2007). The development of reading skills in children with English as a second language. *Scientific Studies of Reading*, 11(2), 105-131.
- MacLean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly*, 33(3), 255-281.
- Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities*, 17(10), 592-599.

- McKinley, R. L., & Mills, C. N. (1989). Item response theory: Advances in achievement and attitude measurement. In B. Thompson (Ed.), *Advances in social sciences methodology* (Vol. 1, pp. 71-135). Greenwich, CT: JAI Press.
- Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3-19.
- Moss, H. E., Hare, M. L., Day, P., & Tyler, L. K. (1994). A distributed memory model of the associative boost in semantic priming. *Connection Science: Journal of Neural Computing, Artificial Intelligence and Cognitive Research, 6*, 413-427.
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665-681.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304-330.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly, 20*(2), 233-253.
- Neuman, S. B., Copple, C., & Bredekamp, S. (2000). *Learning to read and write: Developmentally appropriate practices for young children*. Washington, DC: National Association for the Education of Young Children.
- Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5*, 1-15.
- Owens, R. E. J. (2005). *Language development: An introduction* (Sixth ed.). Boston, MA: Pearson Education, Inc.
- Paul, R. (2007). *Language disorders: From infancy to adolescence* (Third ed.). St. Louis, MO: Mosby, Inc.
- Peverly, S. T., & Kitzen, K. R. (1998). Curriculum-based assessment of reading skills: Considerations and caveats for school psychologists. *Psychology in the Schools, 35*(1), 29-47.
- Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76*(6), 1309-1324.
- Siegel, L. S. (1993). The development of reading. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 63-97). San Diego, CA: Academic Press Inc.

- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stadler, M. A., Watson, M., & Skahan, S. (2007). Rhyming and vocabulary: Effects of lexical restructuring. *Communication Disorders Quarterly*, 28(4), 197-205.
- Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly*, 19(3), 278-303.
- Storkel, H. L. (2002). Restructuring of similarity neighborhoods in the developing mental lexicon. *Journal of Child Language*, 29(2), 251-274.
- van Kleeck, A., Gilliam, R. B., & McFadden, T. U. (1998). A study of classroom-based phonological awareness training for preschoolers with speech and/or language disorders. *American Journal of Speech-Language Pathology*, 7(3), 65-76.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2), 192-212.
- Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, 13(3), 286-350.
- West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, 28(1), 34-50.
- Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552-559.
- Ziefert, H. (2000). *Peephole Rhymes*. Los Angeles: Price Stern Sloan.

Appendix A

Finish-a-Rhyme-Story Assessment Items

Demonstration Item

I'm climbing to see
the house in the tree.
If I don't go carefully
I might skin my [knee].

Practice Item

Here comes the spring!
What does it bring?
Flowers in a ring
and birdies that [sing].

Set A

- A1. My apples went bad.
I'm not glad.
I'll go tell my dad,
that I am very [sad/mad].
- A2. I took my dog,
and went for a jog.
We found a frog
sitting on a [log].
- A3. Robin red-breast,
as a bright red chest.
When he wants to rest,
he flies to his [nest].
- A4. Oh no! A rat.
He's very fat.
Say "bye-bye" rat,
Here comes a [cat].
- A5. Baby's not tall.
She's really small.
If she walks she'll fall.
But she can [crawl].
- A6. The summer's begun,
so we play in the sun.
We hop, skip, and run.
We have so much [fun].

Set B

- B1. I see some sheep
on a hillside that's steep.
They jump and they leap.
When they're tired, they'll [sleep].
- B2. When I take a walk
around the block,
I might get a rock
in my shoe or my [sock].
- B3. In his little bed,
Sleeps a boy name Ed.
Oh no! He fell out of bed,
and hurt his [head].
- B4. I see a cat.
He is very fat.
He sat on my hat,
and now it is [flat].
- B5. We're in our place
ready for the race.
If I set a fast pace,
I'll feel win on my [face].
- B6. I love my little bear.
I found him on the stair.
So I put him in a chair
and brushed his curly [hair].

Set C

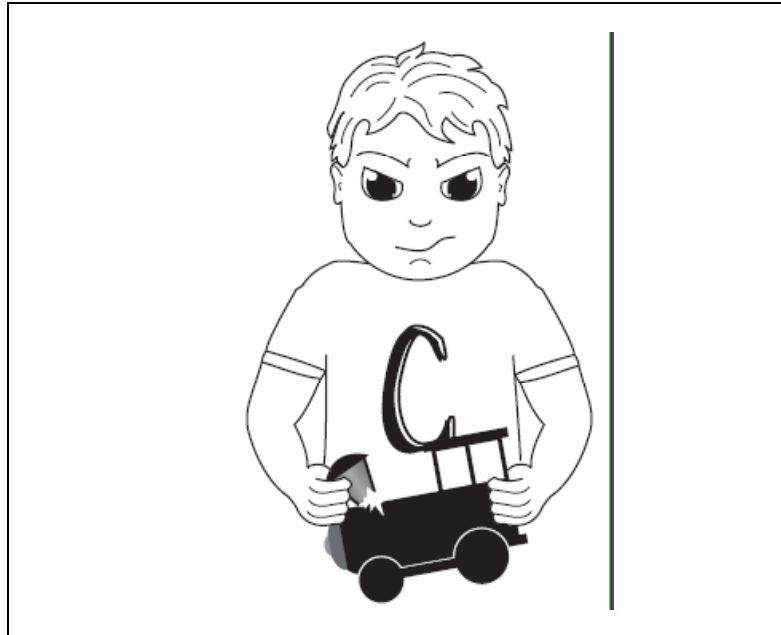
- C1. This is Chad.
He's being bad.
Don't tell Dad,
cause he'll get [mad].
- C2. This fish is Chet,
And when we met
he was caught in a net.
But now he's my [pet].
- C3. Look what I've got.
It's a brand new pot.
I'm cooking a lot.
Watch out, it's [hot].
- C4. I saw a fawn,
Standing on my lawn.
She came at dawn,
but now she's [gone].
- C5. My friend Ted,
Hurt his head.
His mother said,
"Go to [bed]."
- C6. Holidays bring
a wonderful thing
when people sing
and bells [ring].

Set D

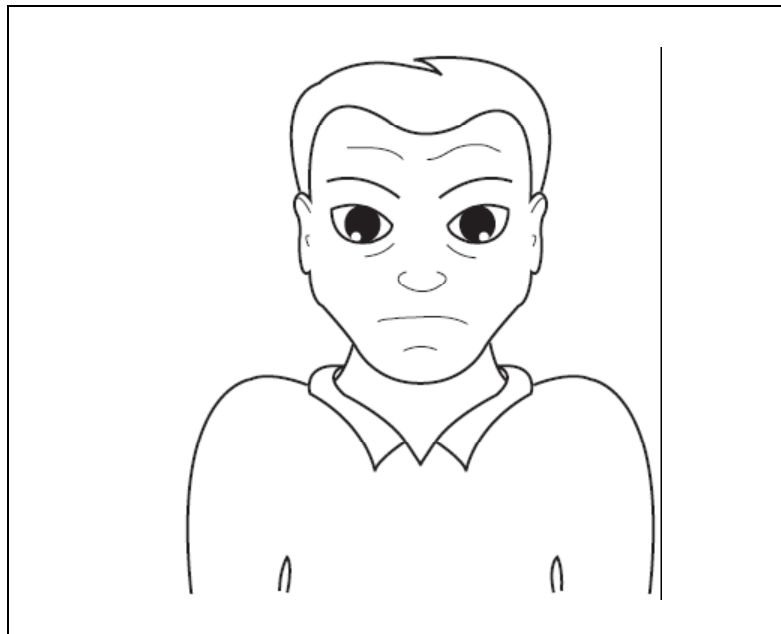
- D1. My friend Brit
likes to knit.
She made me a mitt,
but it's too big and doesn't [fit].
- D2. Paul threw a ball.
It hit the wall.
Then it hit my doll
and made her [fall].
- D3. Have a seat.
I have a treat.
It's something sweet,
that you like to [eat].
- D4. That poor little duck,
has run out of luck.
He walked in the muck,
and now he is [stuck].
- D5. I have one red shoe,
and one that's blue.
What should I do?
I need another blue [shoe].
- D6. My name is Jake.
I like to bake.
What should I make?
How about a [cake].

Appendix B

Example of a Finish-a-Rhyme-Story Item Illustration (Item C1)



This is the first illustration for item C1. The child is shown this picture while the item is read.



This is the second illustration for item C1. The child is shown this picture after producing a response.

Appendix C

Consent to be a Research Subject

Parent Consent for Student

Introduction

This project is being conducted by Dr. Barbara Culatta at Brigham Young University to:
a) identify a set of tasks that can be used to assess young children's ability to recognize words that rhyme as they complete a "story" (fill-in-the-blank) and b) to use the identified tasks to develop an enjoyable rhyme assessment instrument.

Procedures

Your child will participate in two assessment sessions lasting 15-20 min each. The assessment tasks are playful in nature and young children typically find them very engaging. Your child will be told short stories in rhyme form, with accompanying pictures, and asked to complete the story with a word. After filling in the final word in the story, your child will be shown a picture that shows how the story ends. Your child will also be asked to generate words that rhyme. For example, what word rhymes with cake?

Your child's teacher will also complete a short survey which will provide basic information about your child's language background and performance in school, and provide the researchers with the results of the beginning and end of the year assessments as administered by the district, school, and/or preschool program. Your child's teacher will also be given assessment results so that she/he can use the information for instructional purposes.

Confidentiality

Children's names will be removed from all protocols and replaced with initials prior to data analysis. Only the research team will have access to raw data and the names of institutions, schools districts, or individuals will not be disclosed in reports of the research.

Risks/Discomforts

There are minimal risks for participation in this study. Your child may become tired or bored during the assessments. However, if your child becomes tired or bored during the assessments, we will transition him/her back into regular classroom activities.

Benefits

While there are no direct benefits to your child his/her participation in this study, it is hoped that this study will improve our knowledge and understanding of how to more effectively assess young children's early literacy skills.

Participation

Participation in this research study is voluntary. You or your child have the right to withdraw at anytime or refuse to participate entirely without any consequence.

Questions about the Research

If you have questions regarding this study, you may contact Dr. Barbara Culatta barbara_culatta@byu.edu at 422-6262.

Questions about your Rights as Research Participants

If you have questions regarding your rights as a research participant, you may contact Christopher Dromey, PhD, IRB Chair, 422-6461, 133 TLRB, Brigham Young University, Provo, UT 84602, Christopher_Dromey@byu.edu.

I have read, understood, and received a copy of the above consent and desire of my own free will and volition to participate in this study.

Print Name: _____

Signature: _____

Date: _____

Appendix D

Finish-a-Rhyme-Story Administration Protocol

Introduction/Overview

The purpose for this protocol is to keep the conditions of item administration as uniform as possible and to ensure that the students understand what is expected of them in completing these tasks. We need to ensure that item administrators are giving essentially the same information to each student so each student has an equal chance of answering items correctly. It is not absolutely necessary to use the protocol word for word, but doing so is acceptable. The only reason for variation from the protocol would be to ensure that each child understands what is meant by the instructions. Test administrators should not deviate from the intended meaning of the instructions, and should not give any more or less information than what is provided in the protocol. Clarification of word meaning or the meaning of the instructions for a student is acceptable.

When you arrive, ensure that there is a place set up for administration of the assessment. Preferably there will be a place where the student and the administrator will be essentially alone and that will allow you to sit across from the student so the pictures can be easily seen.

Administration Script

Begin by greeting the student and making him/her feel at ease as much as possible. For example, you might say:

“Hi _____ (call child by name). My name is _____ and this is _____ (introduce partner). We’re going to do some fun things together—read some stories and look at some pictures. Will you help me?”

You may need to carry on a brief conversation to further put the student at ease. When you think the student is ready, move into administration of the items as follows:

“Today I’m going to tell you a story with words that sound alike—words like men, pen, ten, and hen. I’ll tell you most of the story but you’ll get to finish it.”

“The first little story I will read will show you how we’re going to do this. Look at the picture and listen as I read you the story.”

Place the picture on the table and read the *knee* demonstration item:

*I’m climbing to see,
the house in the tree.
If I don’t go carefully,
I might skin my _____.
(pause)*

“Knee”

Then flip to the picture of the target word and say:

“Knee sounds like tree and see and carefully, doesn’t it? And it finishes the story.”

NOTE: If the student happens to supply a word as you say the target word—either the correct word or a different word—thank the student and proceed as indicated above. Now, flip back to the first picture and say:

“So listen once more to the whole story:

*I’m climbing to see,
the house in the tree.
If I don’t go carefully,
I might skin my knee.*

(Flip to the target picture again as you say the word.)

“Now let’s try another one where you help me finish the story with one word. Look at the picture and listen as I read the first part.

Read the **Sing** practice item and wait for the student to fill in the last word.

*Here comes spring!
What does it bring?
Flowers in a ring
and birdies that ____.*

Positively acknowledge the child’s response (without indicating whether it is right or wrong). Say something like:

____ (repeat the word the child said). Thank you. Let’s see if that fits.

Then flip to the picture of the target word and, if the child said “sing” say:

“Oh, it is sing. Sing sounds like ring and it finishes our little story, just like the other one.”

NOTE: If the child does not reply or says a different word, flip to the picture and say:

“Oh, it’s sing. Sing sounds like ring and it finishes our little story, just like the other one. Okay, are you ready to help me finish some more little stories like this one?”

If the child seems unsure, repeat the demonstration item or do whatever you can to try to build the child’s confidence. Reassure the child that you will help, and that there is nothing to worry about. When the child is ready, proceed with the remaining items set.

NOTE: For the second administration, do only what is necessary to remind the students of the task (use the same example items), and then proceed to administration of the items.