HOW TEACHING WORD FAMILIES AFFECTS BEGINNERS' READING AND SPELLING

by

Esther Y. Scheiner

A dissertation submitted to the Graduate Faculty in Educational Psychology
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York

2005

This manuscript has been read and accepted for the Graduate Faculty in Educational Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date	Dr. Linnea Ehri, Chair of Examining Committee
Date	Dr. Alan Gross, Executive Officer

Dr. Linnea Ehri

Dr. David Rindskopf

Dr. Carol Tittle

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

HOW TEACHING WORD FAMILIES AFFECTS BEGINNERS' READING AND SPELLING

by

Esther Y. Scheiner

Advisor: Dr. Linnea Ehri

This study examined the effects of teaching words in families on beginning readers' reading and spelling. Four groups of beginning readers were randomly assigned to one of three word learning conditions or an untrained control group. The students in the word learning conditions practiced reading 16 words, comprised of four words from each rime-based word family (-ale, -uck, -ill, -ark) with initial letters b, p, m, d. Words were grouped into sets of four words for learning. The words were grouped by shared rime, or by mixed rime with the same initial letter, or by mixed rime with different initial letters. Several tasks measured participants' pretreatment and posttreatment reading and spelling skills.

Results showed that participants who were trained with mixed family word sets produced superior gains in word learning on posttests compared to students who were trained with shared rime words in pure families. Gains were slightly superior for students trained to read sets of words that contained mixed rimes but shared initial letters. Although the students trained with pure word families performed significantly better during the training, they did not secure the training words in memory in sufficient letter

detail to maintain superior learning on the posttest. Students entering the study with more advanced decoding ability were able to learn more words during training. This outcome supports Ehri and Robbin's (1992) finding that in order for readers to read new words by analogy to sight words, they must have some decoding skill.

Results suggest that teaching words in families as part of early reading instruction is not as effective as presenting students with lists of words that do not share rimes. Word reading practice needs to be structured so that beginning readers are pushed to form connections between individual graphemes and phonemes. This helps them attend to and remember the words.

Acknowledgements

I wish to thank my advisor, Linnea Ehri, whose insight and suggestions improved this dissertation enormously. I am appreciative of the wisdom and assistance from my committee members, Carol Tittle and David Rindskopf. Any remaining errors, are of course, my own.

I also wish to acknowledge my gratitude to the outside readers and others in the department for their helpfulness: Alpana Bhattacharya, Alan Gross, and Rhonda Palant; and I am indebted to Helen Johnson, who has done more things for me than I could possibly attempt to list here, and always with the utmost graciousness. Finally, I thank my parents and husband for their support and encouragement.

TABLE OF CONTENTS

Introduction	1
Literature Review	3
Method	43
Results	56
Discussion	73
Tables	97
Figures	119
Appendixes	121
References	135

List of Tables Page

1. Training Words from Pilot Study 97

2. Number of Training Words Read Correctly on Pilot Study Posttests 98

3. Training Words 99

4. Mean Performance of Low and High Decoders on Pretest Measures and Test Statistics 100

5. Mean Performance of the Word Family Treatment Groups and the Control Group on the Pretest Measures 102

6. Mean Number of Training Words Read Correctly During the 11 Trials for the Treatment Groups 106

7. ANOVA test Statistics for Reading Words During Training as a Function of Decoding Level and Treatment Session 107

8. Mean Performance of the Word Family Treatment Groups and the Control Group on the Posttest Measures 108

9. Mean Performance of Low and High Decoders on the Posttest Measures 110

10. Pearson Correlations Between Pretest Measures and Word Learning for All Participants 112

11. Pearson Correlations Between Pretest Measures and Word Learning for Participants in Treatment Groups 114

Table Page

12. Summary of Regression Analysis for Pretest Variables Predicting

Word Learning Gain (Pretest to Posttest) in Words Recognized Immediately 116

13. Children's Reading and Spelling of Words with Shared Rimes on the

Pretest and Posttest. 117

List of Figures Page

1. Mean Performance Per Session as a Function of Decoder Level and Treatment Group 119

2. Mean Number of Words Read Correctly by High and Low Decoders During Training Sessions as a Function of Treatment Group 120

Chapter 1

INTRODUCTION

The purpose of this study was to examine the effects of training beginning readers to read lists of words grouped into word families. “Word families” are groups of words that have identical rimes, with each member having a different onset. An onset is the initial consonant sound(s) that precede(s) the vowel in a word, and the rime is comprised of the vowel plus remaining consonants, for example, j-ump or str-ike (Trieman, 1985). Presenting word families is a common technique that teachers use in classrooms. The question of interest was whether students would be able to read and spell more words on posttests if they were trained with mixed lists of words that did not share rimes, particularly if the lists contained words that all begin with the same initial letter.

According to Ehri (1991, 1994, 1995) beginning readers do not yet have strong phoneme-grapheme mapping skills at the phoneme level, and therefore they are unable to process consolidated units such as rimes in words. However, according to Goswami (1986, Goswami & Bryant, 1986) beginning readers are able to read through analogy and are able to segment words into onsets and rimes more easily than they can segment words into individual phonemes.

Readers may process words in several different ways: through immediate recognition from memory, by sounding out and blending back together, by decoding common spelling patterns, by analogy, or through the use of context cues (Ehri, 1991, 1994). Typically, common words are recognized immediately (“on sight”), while unfamiliar words require one or several of the strategies listed above.

Although skilled readers can identify words quickly and correctly, beginners cannot. Beginning readers do not yet have full knowledge of the letter-sound system, and typically have weak segmenting and blending skills. They often rely on forming partial connections to read words. For example, they will use the initial letter and context knowledge when confronted with an unfamiliar word.

The present study explored ways to strengthen children's word learning by examining how various ways of grouping words in lists influences children's attention to graphemes in particular words and their memory for the words. It was anticipated that word recognition by readers who learned to read lists of words that did not share rimes would be superior to word recognition by readers who learned to read words in pure word families, and that an additional advantage would accrue for participants trained with lists of words that did not share rimes and shared the same initial letter. Differences in word learning between beginning readers with lower and higher decoding skills were also examined. Findings were expected to make a significant contribution to methods of early literacy instruction by providing evidence for instruction involving phoneme-grapheme segmentation and blending when word families are used as part of a beginning reading curriculum.

Chapter 2

LITERATURE REVIEW

The development of word reading ability

Acquiring literacy skills is a long, complex process. One important component involves acquiring the ability to read words. For fluent decoding, readers need to have knowledge of phoneme-grapheme connections, a large mental vocabulary of "sight" words stored in memory, knowledge about spelling regularities, as well as a great deal of explicit and implicit knowledge about grammar, syntax, and the pragmatics and semantics of language.

According to Ehri (1991, 1994, 1995; Ehri & McCormick 1998), typically developing readers progress through various phases of word learning as they gain proficiency in reading words. As alphabetic knowledge used to form connections for word identification develops, children gradually advance through the phases. Each phase is characterized by reliance on one type of connection as the predominant strategy for word identification.

The first phase, called the *pre-alphabetic phase*, is characterized by word recognition that is not dependent on the alphabet. Children may recognize a word because there are environmental cues surrounding it; for example, reading "stop" on the stop sign, but not recognizing the word in isolation. Children who can "read" the word would still read "stop" even if all the letters were altered as long as they were printed in white on a red, hexagonal-shaped sign (Masonheimer, Drum, & Ehri, 1984). Children may be able to "read" a word (they are actually identifying it) because they are able to remember certain visual features of the word, unrelated to the grapheme-phoneme information contained within it. For example, a child might know that the word "look" has two circles in it. However, a child who "reads" the word "look" on the basis of the visual cues will have trouble if presented with other words containing two o's in the middle, such as "tool." At this early phase, the child is oblivious to the grapheme-phoneme information contained in the word, and is unskilled in decoding.

In the second phase, which Ehri terms the *partial alphabetic phase* (*or partial cue phase*), children have gained knowledge about letters and their names or sounds. Children at this phase are able to process some of the grapheme-phoneme connections in a word. Although they are not yet fluent readers, they have enough knowledge to be able to make some good guesses when they have environmental clues. For example, if a child sees a picture of animal labeled "dog" and knows that d makes a /d/ sound, it is possible for the child to figure out the entire word. This strategy is the use of the initial-letter to predict the whole word. Of course, when confronted with several words that share the same onset, the child's use of this strategy will lose its effectiveness. For example, a child may read "house" for "home," or, when there is less contextual information to rely on, might even "read" the word as "horse," or "hop" instead. Readers in this phase can engage in phonetic cue reading. By remembering the initial and final segments in words they are able to read some words by sight.

During the third phase, termed the *full alphabetic phase*, children learn to associate sounds with almost all the letters that they see in words. Children gain knowledge of the most common grapheme-phoneme correspondences, and begin to decode, or "sound out" unfamiliar words. Although reading begins as a slow, and even sometimes laborious, process for children, they quickly gain speed with practice in decoding. Additionally, during this phase, they are able to fully analyze the letter-sound relationships in words to store them in memory as "sight words." A word becomes a "sight word" when the child has had enough exposure to the word so that strong connections have been made in their brains which enables the visual sequence of letters to be recognized automatically from memory. The child is able to recognize the entire word as one unit without having to consciously break the word into sounds. For example "pot" is segmented into the sounds /p/ /a/ /t/ and then blended together. Instead, these words are recognized just by sight, without the need for analyzing each letter's phonemic value and sounding it out. Because the child has had enough exposure to the word, it is recognized instantly as a whole word. Some words, such as "aisle" can only be read by sight, and are fixed in memory after enough exposures. Children gain the ability to read words through analogy at this stage, but this strategy becomes more common in the fourth reading phase.

The connections made in the third phase become consolidated into larger units as the same letter sequences are practiced. As more and more larger units become known, children shift from the third phase into the fourth, *consolidated alphabetic phase*. In the fourth phase, children begin to recognize chunks of letters as units, rather than having to process each individual grapheme-phoneme connection. For example, children recognize "ing" as a single unit in a word, as a result of reading this letter-sequence in several words. They are able to process common spelling patterns, such as "-est" and "-and" quickly. Sight vocabularies grow larger in this phase. Children also learn spelling regularities that assist them in decoding words. For example, they learn that a doubled consonant softens the preceding vowel (compare "later" and "latter.")

In order for children's reading to become fluent, they must develop a large sight word vocabulary. Without the ability to process many words automatically, reading is a slow, laborious process which involves much conscious effort and the need to focus on decoding, rather than on meaning, when reading. However, in order to reach the point at which words are recognized as whole words, children must go through the decoding process. That is, they must be able to make grapheme-phoneme connections quickly, and learn how to break down words to analyze the graphemes' phonemic representations, which then need to be blended back together to form the word. Without these experiences, the word will not become part of the child's sight vocabulary. This is very likely why the "whole word" reading method emphasizing use of visual feature of words is less effective than phonics instruction in helping beginners learn to read. (See Vellutino, 1991 for an extensive review of such studies).

Word Families

"Word families" are groups of words that have identical rimes, with each member having a different onset. An onset is the initial consonant sound(s) that precede(s) the vowel in a word, and the rime is comprised of the vowel plus remaining consonants (Trieman, 1985). For example, the word family "-ate" contains the following words: late, hate, mate, rate, fate, plate, crate, etc. These words all share the same sounds in the rime unit, and the rimes have the same spelling pattern. However, the words “weight” and “trait” are not included in the word family described above. Though “weight” and “trait” share a phonological rhyme to “-ate”, they do not share the rime spelling pattern, which is what determines whether or not a word is included in a word family. The fact that oral rhyming rime spelling families in words may or may not identify words in English very likely creates confusion when children attempt to transfer oral language knowledge to spell new words.

Words in word families share the same spelling patterns, and they rhyme with one another. In theory, a child who knows one word in a family could make an analogy to read a new, unfamiliar word in the same family. Using word families as part of early literacy instruction raises several interesting questions:

1. Are children naturally aware of onset-rime units in words?
2. If not, can children be trained to use onset-rime units to analyze words?
3. Is there a relationship between rhyming ability and word reading?
4. Do children read words by analogy?
5. If children do not spontaneously make analogies to read new words, can they be taught to do so?
6. Do children benefit from having words presented to them in families?
7. What processes or strategies do children use when confronted with lists of word families?
8. How does presenting words in lists of word families affect children’s ability to retain words in memory and spell them?

Presenting word families is a common technique that teachers use in classrooms. Word families have been used as an activity to heighten awareness of onset and rhyme and to stimulate phoneme awareness among young children (Edelen Smith, 1997).

However, teachers may lack an understanding about how to incorporate word families effectively into instruction. For example, during a recent visit to a first grade classroom, teacher-made word family posters were observed prominently displayed. The "ight" family contained words such as "night," and "right," as well as words beginning in consonant blends, such as "bright." However, nonmember "height" was also included.

Word families also possibly benefits dyslexic readers when they are taught to use rime-unit analogies to read new words. Research has shown that children with reading disabilities make achievement gains when trained to use rime-based orthographic analogies (Greaney, Tunmer, & Chapman, 1997).

However, if students lack sufficient knowledge of grapheme-phoneme relations, then providing them with word families may not be beneficial. Gaskins, Ehri, Cress, O'Hara, and Donnelly (1996-97) describe a reading program in which they taught struggling readers to learn as sight words a set of key words that included high-frequency spelling patterns, such as *–et, -in, -and*, and *–up.* However, they found that although children could read the key words containing these patterns, some failed to read new words containing the same known patterns by analogy. For example, the children learned the word *yellow*, but when presented with the word *fellow*, they did not recognize the similarity between the unknown word and the known word. The researchers suggest that this processing failure was due to a lack of memory for all of the letters in the words. Children may have simply remembered only the salient characteristics of the word, such as beginning and ending letters.

The findings of a pilot study by this author, (discussed in detail below), in which children were given practice reading several sets of whole words from word families under different conditions, demonstrated that students retained more words in memory when the words from the same families were mixed across the sets rather than taught together as pure word families sets. One explanation for the field study results is that students trained under the mixed-order list condition were forced to attend to the grapheme-phoneme connections in the words, while the children in the pure "word family" condition had the option of relying on a initial-letter strategy combined with a rhyming strategy to read the words during training.

The relationship between rhyming ability and reading

Studying word families raises the question of the role that rhyme may play in learning to read. Teaching word families to children is done with the expectation that they will rely on their rhyming knowledge and strategies to decode words rather than sounding out and blending the grapheme-phoneme connections in words. If rhyming is determined to be an important ability that contributes to reading, then perhaps exposing children to lists of word families will help them read words.

Read (1971) studied the rhyming ability of four-year-old kindergarten children. Children were presented with a rhyming game in which puppets were said to only like objects that rhymed with their names, (i.e. children were asked, "Would Ed like bed or bead?"), and the results demonstrated that the overwhelming majority of children had little difficulty with this type of task. However, other research has suggested that children have difficulty making rhyming judgements (Bowey & Francis, 1991; Knafle, 1974). This discrepancy may be due to the specific demands of the rhyming task.

Bradley and Bryant (1983, 1985), in a longitudinal study involving 368 children, tested the relationship between rhyme and reading. The results demonstrated that children who were more sensitive to rhyme at four or five, when they were prereaders, had more success as both readers and spellers when they were eight or nine years old, after controlling for I.Q. scores. The researchers used a rhyme oddity task in which the child had to select the word with a different rhyme from the rest of the words on the list. The children's early rhyming scores were related to later success in reading and spelling, but not to their mathematics achievement. Bryant, MacLean & Bradley (1990) found a correlation between rhyming skills of four-year-olds and their reading and spelling abilities two years later. Of course, these correlations do not prove causation.

Ellis and Large (1987), conducted a longitudinal study spanning three years with prereaders who were five years old at the start. At the end of the study they compared three groups of children, those with high I.Q.s who were poor readers, those with low I.Q.s who were poor readers, and those with high I.Q.s who were good readers. They found that the best predictor of reading level among the various tests that the children took during the three years of the study were measures of sensitivity and ability to rhyme.

Dyslexic children perform worse than age-level-matched peers on tests of rhyme oddity and rhyme production (Snowling, Goulandris, & Defty, 1998). Bryant and Bradley (1985) comment that

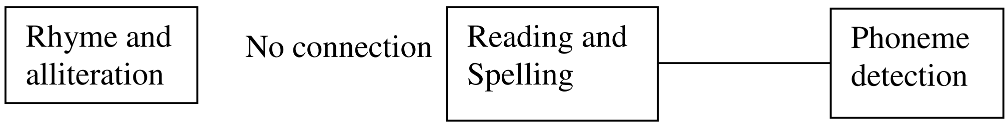
> "The discovery that backward readers are very poor at detecting rhyme…prompts two questions. One is whether this particular skill is an important one for every child … The second question is about the nature of the skill. Are backward readers

> poor only with rhyme or do they fail at any test of awareness of sounds in words?" (p.52.)

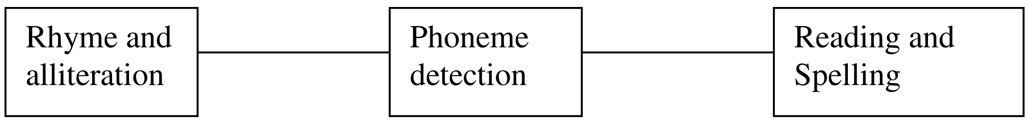
There is a link between children's rhyming ability and their skills in analogy-making when they read new words. Goswami (1990b) studied the relationship between measures of phonological awareness, including a measure of sensitivity to rhyme and alliteration and a measure of phoneme deletion skill, and children's use of analogies when reading. The results demonstrated that rhyme scores were more closely associated with children's analogies in reading than they were to the phoneme deletion task. Bradley and Bryant (1978), in a reading level match design, compared 60 ten-year-old readers who read on a seven-year-old level with normal seven-year-old readers. They asked children to select the word that ended differently from the others in a group of words. For example, children were given the list: "weed," "need," "deed," "peel," and were supposed to choose "peel," which has a different ending (or rime) than the others. The poor ten-year-old readers had more difficulty with this task (as well as the other tasks in the study) than did the seven-year-old normal readers. Hanley, Reynolds, and Thornton (1997), studied nine dyslexic children, aged nine to eleven, and compared their reading performance to reading-age matched comparison group of younger children. They found that the number of analogous words that the dyslexic children read was significantly correlated with their performance on a rhyme-detection test.

Goswami and Bryant (1990) describe three possible models connecting rhyming abilities and phoneme detection with reading and spelling based on the literature.

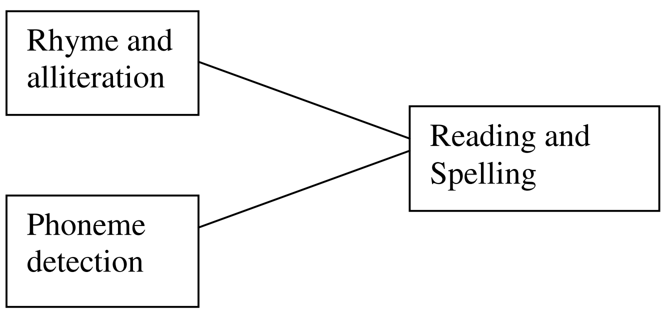
Model 1. Reading leads to phoneme detection.



Model 2. Rhyme leads to phoneme detection and thus to reading



Model 3. Rhyme and phoneme detection have separate paths to reading



Goswami & Bryant (1990) (p.110)

Goswami and Bryant (1990) argue that rhyme makes a direct and independent contribution to reading, because it allows children to connect sequences of letters with onsets and rimes, as displayed in model 3, above.

However, Wood (1999) found that rhyming ability contributed less to orthographic analogy use than did phonemic awareness. In a study examining the relationships between rhyming ability, phonemic awareness, reading ability and orthographic analogy use, 68 children (mean age = 5.8) were given a test of reading ability, and an orthographic analogies task, as well as other measures. Results of a regression analysis demonstrated that phonemic awareness and reading experience were best able to account for analogy use. Additionally, it was found that orthographic analogy use did not directly contribute to early reading ability, which contradicts the third model proposed by Goswami and Bryant. Perhaps Wood's findings support a fourth hypothetical model, in which phoneme detection leads to both rhyme and alliteration and reading and spelling.

Savage (2001) argues that the "current level of uncertainty" in research is not reaching teachers, who are using rhymes in teaching practices based on claims such as Goswami's. Savage points out the difficulties in proving a causal relationship between rhyme and reading, and critiques Goswami's (1999b) review of research studies involving research on children's knowledge and awareness of rimes.

Because a rhyming strategy may be used by children when they are taught word families, the present study included a measure of rhyming ability. Of interest was whether rhyming ability was correlated positively with learning words in families.

Onset-rime units in words

It is important to examine the orthography that readers encounter to clarify how beginning readers might process onset-rime units in particular words. The frequency and consistency of rimes has been studied through analyzing large numbers of words.

Glushko (1981) points out that reading words by analogy may not be a simple process in English, considering the sound-spelling rules. Suppose, for example, a reader is confronted with the unknown word "toves." The reader has a choice of choosing one of three pronunciations, all of which would be acceptable according to English spelling-sound rules. "Toves" may be pronounced as analogous to "roves," "loves," or "moves." (Glushko, 1981). Glushko (1979b) examined the speed with which skilled readers could identify words which were regular and consistent compared to words that were not. He presented 164 words in random order to 16 adult readers. Results demonstrated that regular and consistent words, like "wade" which do not have more than one analogous alternative, were pronounced faster when compared to words such as "wave," for which there are alternative analogs (i.e. "cave" or "have"). Exception words like "have," (which has the less usual pronunciation for the –ave unit) took the longest.

Glushko (1979a) found that presenting "pint" in a list of words slowed reader's pronunciation of "-int" words that they pronounced to rhyme with "hint," because it activated another alternative. Glushko (1981) indicated that he tried to "refrain from using the concept of analogy" in his "description of the workings of the mechanism." He reasoned that "analogy implies a conscious aspect to the process of activation that is not reflected in what readers think they are doing and could not occur in the half a second or so required to pronounce words (p.71-72)." Glushko's comments are interesting because he raises the issue of making a differentiation between a conscious use of strategy versus the spontaneous activation of a process because of properties inherent in a certain situation (i.e. shared rimes of new words which can be recognized as analogous to a known word). His comment about the speed of word-processing is particularly insightful in considering whether one wishes to believe that readers use analogies to read words. Glushko (1981) argues that readers must use grapheme-phoneme knowledge when confronting a new word because analogy would not always be a useful strategy, given the research evidence discussed above. In many cases, reading by analogy, or by analogy alone, is not possible.

Goswami (1998) compared seven, eight, and nine-year-old children's speed and accuracy in reading nonsense words like "dake" vs. "daik" and "bomic" (analogous to "comic") vs. "bommick." Words such as "dake" which have analogous neighbors were read more accurately and faster by children of all ages, and results were replicated with a group of university students.

Leslie and Calhoon (1995) studied rime "neighborhoods" and their effect on children's reading. A rime neighborhood is defined as sequences of orthographic triples that are found in multiple words, i.e. "-ood." The size of the neighborhood is determined by the triple's frequency in words. A second issue the researchers examined is the frequency with which particular words are found in English texts. Lastly, the consistency of the rime's pronunciations was considered. For example, "-ood" is an "inconsistent" neighborhood, because this triple may be pronounced two ways (as in "hood" versus "food)." The researchers demonstrated that rimes that come from large neighborhoods were read correctly more often than rimes from small neighborhoods, in both lists and stories. There were significant interactions between consistency and frequency. High frequency words were read equally well regardless of whether or not the neighborhoods were consistent. However, low frequency words from inconsistent neighborhoods were read less accurately than were low frequency words from consistent neighborhoods.

Stanback (1992) analyzed 17,602 frequently used words in English. The syllables in each word were examined. Syllables are units within a word that contain only one vowel. The syllable may be a lone vowel sound, a vowel preceded by a consonant sound or consonant cluster, a vowel followed by a consonant sound or cluster, or a vowel preceded and also followed by consonants. Stanback found that the same 616 orthographic rimes occurred in rime families of almost all the 43,041 syllables. Of these, 436 were both regular and consistent in pronunciation, while 86 had less than a 90 percent level of consistency.

Children's ability to segment words into onset-rime units

Wise, Olson, and Treiman (1990) conducted an experiment with 20 first graders, early in the children's school year. The children knew the letters and letter-sounds, but none could read more than a few preprimer words. Children listened to words on a computer. Children heard half the four letter, three phoneme words divided into onsets and rimes (i.e. "dish" segmented as "d-ish") or segmented between the second and third phonemes (i.e. "dish" as "di-sh.") On post-testing, results demonstrated that more short-term word learning occurred in the onset-rime segmentation condition. However, effects were no longer significant in test sessions conducted half an hour after the conclusion of training.

Goswami (1991) examined children's ability to read consonant clusters in new words after they had been trained to read words that contained that consonant cluster. 20 beginning readers (approximately equivalent to American first and second graders) were trained to read single-syllable "clue" words. Words were segmented for the students in a way that did not break down consonant clusters (i.e. "trim" was divided into three units "tr/i/m;" "desk" was divided into "d/e/sk.") Clue words were left in the child's sight to aid memory. Children were then given analogy words that shared consonant blends either at the beginning or the end of words (i.e. an analogy word for "trim" was "trap;" an analogy word for "wink" was "honk..") Results demonstrated that it was easier for the children to transfer knowledge of consonant cluster when the cluster was in the onset segment of the word rather than when it was a part of a rime unit. In a second experiment, children were given clue words such as "wink" with analogy words such as "pink," that shared the rime of the clue word. Alternatively, children were given clue words such as "trim" which shared heads (onset plus vowel) with analogy words such as "trip." Children were able to spell new words that shared the rimes of clue words better than words that shared the heads of the clue words.

Treiman (1985) conducted an experiment in which 8 year olds were asked to substitute two phonemes in words with two other phonemes. Children had a much easier time substituting consonant cluster onsets (i.e. gwe to sle) than they did substituting two phonemes in the second and third positions in the words (i.e. gwe to gli). In another experiment, Treiman (1985) asked children to decode nonsense words, some of which had consonant clusters and some of which did not. Although the words contained the same letters, first graders had an easier time decoding words such as *keer*, that did not have consonant clusters in the onset, than words like *kree*, which had consonant clusters in the onset. Treiman (1992) argues that this is evidence that children analyze words into onset-rime units for word reading before they can decode words by analyzing all the grapheme-phoneme correspondences.

Treiman (1992) makes the case for teaching decoding at the onset-rime level before training in individual letter segmentation and blending. The theoretical justification rests on studies demonstrating that young children have an easier time segmenting onset-rime units in speech than they do segmenting words into individual phonemes.

Making connections between words that share rimes

According to Goswami (1992), in order to use analogies, children need to have an awareness that words are composed of onset and rime units. Goswami (1999a) states that "onset-rime awareness is present in most four-to-five year olds (p. 177)." According to Goswami (1992), if children have knowledge of the division between onset-rime in spoken language, they can then make analogies to written language and can use spelling patterns that reflect onset-rime divisions to decode new words by analogy.

Goswami (1986) explored children's ability to make use of analogies by presenting three groups of children, ages five, six and seven, with a list of "clue" words. They were told what each word said, and then they were shown the analogous words and read them. Goswami compared children's ability to read the new words that shared letter strings with "clue" words. For example, children were taught the word "beak" and then asked to read a word that shared a rime (i.e. “peak”), a word that shared the head (i.e. “bean”), or a word that shared three letters with clue word (i.e. “bask”). There were significant differences in all three groups’ ability to read the analogy words that shared a rime and common spelling sequence (“beak”-“peak”) compared to the children’s ability to read the control words (“beak”-“bask”). The children did better when they were able to take advantage of the analogy to read the new word. This is evidence that even beginning readers do make analogies based on rime. Even pre-readers had an easier time reading analogous words that shared rimes with the clue words. However, this study and others like it have been criticized because the reading task does not take into account the effects of phonological priming (Savage, 1997, 1999).

Goswami (1990a) conducted a study in which she gave the participants clue words followed by words that rhymed, some of which shared the spelling pattern of the clue words and some of which had different spellings. For example, the clue word “head” followed by the words “bread” and “said.” Participants did much better on the words that had the same spelling pattern, indicating that participants in both of these studies were not simply being primed to say a rhyming word. Goswami (1990b) suggests that a phonological priming effect contributes to, but does not explain, the pattern of significant transfer between words that rhymed but did not share rimes (i.e. bread and said) demonstrated in this study.

Studies involving “clue” words presented immediately before reading transfer words do not explain what children do when confronted with a new word in the absence of a clue word, which is typically the way unfamiliar words are encountered in real life. Goswami (1988b) investigated transfer effects when children were asked to read new words embedded in text. Clue words were either taught in the title of the story, or taught in the title and repeated on the first page, or not taught at all, but children were not given information about how to segment the word. Results demonstrated that there were no significant differences between transfer to words that shared rimes (i.e. beak to peak) versus words that shared heads (i.e. beak to bean). This suggests that the "clue" words in training studies do serve as a prompt of some kind, likely a phonological one, as Savage (1999) argues.

Muter, Snowling, and Taylor (1994) asked children to read a set of words, and then taught them a set of clue words, which were analogous to half the words on the first list. The children were then asked to read the first list a second time. Words on the original list to which analogous clue words had been taught were read significantly better than words to which no analogous clue words were taught. The researchers also investigated the effects of keeping the clue words in the child's sight while the original list was read the second time. The children read the analogous words significantly better when the clue words were present. Similarly, Savage and Stuart (1998) presented six-year-old children with "clue" words, and then showed them the test words. The children were told that the "clue" words were "clues to the mystery of how to read" new words. Some children were given the pronunciations, but not the spellings, of the "clue" words, (for example, the child heard the "clue" word "beak," which was then followed by the test word "peak.") Under the alternative condition, children both heard and saw the "clue" words. Children read more of the test words correctly when they were able to see the words, compared to the children who only heard them (55% to 40%). This lends evidence to the position that something beyond phonological priming is occurring.

Goswami (1999a) points out that, “given the salience of the rime in the phonology of the English language and its effects in adult reading, it would be surprising if no rhyme priming effects were found with children. This salience is one of the factors that can be exploited by teachers who use analogy methods in the classroom” (p.193).

Factors accounting for use of analogies: age, reading skills, and instructional method

Bowey and Underwood (1996) found increased use of orthographic rime correspondence in nonword reading tasks between second and fourth grade, and their experiments demonstrated that this skill is strongly associated with word-level reading skills.

Nation and Hulme (1997) examined the relationship between different types of phonological awareness and reading and spelling ability. Their study involved 75 children, from Primary Years 1, 3, and 4 (approximately equivalent to the corresponding American elementary school grades). Children were given four phonological awareness tasks, two measuring sound categorization and two measuring segmentation ability. Participants were also tested on reading, spelling, and digit span. Sound categorization and phonemic segmentation performance increased with age; however, the ability to divide nonwords into onset-rime units did not. The children in this study had difficulty with onset-rime segmentation tasks across all three age groups, which the researchers argue conflicts with Goswami and Bryant’s (1990) claim that onset-rime awareness develops early and without systematic training. Hierarchical multiple regression was used to determine the extent to which the different phonological awareness tasks predicted variance in reading and spelling ability. Onset-rime segmentation failed to account for any statistically significant variance in spelling or reading ability. Phonemic segmentation was shown to be the best predictor of reading and spelling ability. The researchers note that the participants had been exposed to instruction that could be classified as a mixture of look-say and phonics-based methods, and that results might have been different for a group of children that had been exposed to literacy instruction that emphasized onset-rime segmentation skills.

Goswami (1999a) points out that Nation and Hulme reported that the youngest children in the study found onset-rime segmentation easier than phonemic segmentation tasks (averaging 55% correct versus 24% correct.) Additionally, Goswami believes that the findings of the study were consistent with the teaching methods the children were exposed to, and that is why results demonstrated an improvement in phonemic segmentation skills across the age groups but not in onset-rime segmentation skills. Goswami (1999a) concludes her discussion of the Nation and Hulme study by saying that the study “demonstrates the importance of the teaching method for the development of metalinguistic control over phonological structures” (p. 187).

Researchers disagree about whether or not children can use onset and rime units to decode words before they have mastered mapping individual phonemes to graphemes. Goswami (1986, 1990a, 1999a) argues that onset-rime units are accessible to young children. However, according to Treiman (1993), children’s ability to use their knowledge of onsets and rimes is dependent on their ability to make connections between letters and sounds in isolation. Ehri (1998a) analyzed the subskills necessary in order for children to be able to read words by analogy. She argues that children do not simply recognize a recurring letter pattern across two or more words. Children must be able to segment and blend the letters that match the onset and rime (spoken) segments of words, recognize the sounds of letters in familiar words, and have enough word knowledge to recognize how words are similar and different.

Ehri and Robbins (1992) taught children to read one-syllable words with nonstandard spellings and examined the effect on children's ability to transfer word learning to analogous words and to words that shared the experimenter-created orthographic unit. For example, children were taught the word "cave" spelled "KAAV" and later asked to read "SAAV" and "RAAN." The researchers found that the participants who lacked decoding skills could not read words by analogy. Ehri explains that the nondecoders only stored partial representations of the training words in memory, and so they failed to recognize that they shared rime spellings with unfamiliar words.

Gaskins, Ehri, Cress, O'Hara, and Donnelly (1996-97) describe a reading program in which they taught struggling readers to learn as sight words a set of key words that included high-frequency spelling patterns, such as *–et, -in, -and*, and *–up.* However, they found that although children could read the key words containing these patterns, some failed to read new words containing the same known patterns by analogy. For example, the children learned the word *yellow*, but when presented with the word *fellow*, they did not recognize the similarity between the unknown word and the known word. The researchers suggest that this processing failure was due to a lack of memory for all of the letters in the words. Children may have simply remembered only the salient characteristics of the word, such as beginning and ending letters.

The study by Leslie and Calhoon (1995) involving 40 first and second graders demonstrated that children reading at or below second grade level were less affected by rime-neighborhood size than children who were reading at or above a third grade level, suggesting that the younger readers relied more on other strategies such as individual grapheme-phoneme mapping, and possibly less effective cues such as initial letters rather than applying their knowledge of rimes.

Marsh, Desberg and Cooper (1977) examined how younger and older readers process nonwords that are analogous to real words. They presented three groups of participants, ages 10, 16, and college undergraduates, with a list of ten nonsense words. The experimenters offered the participants two ways to pronounce these nonsense words. One followed the grapheme-phoneme rules and the other choice was a pronunciation analogous to a real word. For example, participants could choose to pronounce "puscle" as though it rhymed with "muscle" or they could pronounce it as if it was spelled "puskuhl." The researchers found that younger readers made fewer analogies and used grapheme-phoneme rules more frequently than the older groups of readers did.

A second study by Marsh and his colleagues (Marsh, Friedman, Welch, & Desberg, 1980a) used a similar task with seven-year-olds, ten-year-olds, and college students. Participants were presented with nonsense words that were embedded in text passages in a way that indicated that the nonsense words stood for nouns. The results of this study demonstrated that the seven-year-old group used analogies an average of 14% of the time. In contrast, the ten-year-olds used analogies 34% of the time and the college students 38% of the time.

However, Goswami and Bryant (1990) argue that there are two reasons why these results should be questioned. They point out that in order for a reader to take advantage of an analogous word, a reader must know the spelling pattern of the original word that the new word resembles. So if a seven-year-old does not know the word “muscle,” it is not possible for him to make an analogy to this word. The older readers, in contrast, knew more of the words to which the nonsense words were analogous. A second point is that readers are more likely to make use of grapheme-phoneme rules when exposed to nonsense words than to real words, so these results may not hold true for typical strategies in real-life situations when readers are confronted with unknown real words.

When Marsh and his colleagues (Marsh, Friedman, Welch & Desberg, 1980b) examined children’s use of analogies to spell nonsense words, they found similar results to their studies involving reading. For example, when the three groups of participants, seven-year-olds, ten-year-olds, and college students, were asked to spell a nonsense word such as “jation,” the following patterns emerged. The researchers found that none of the seven-year-olds spelled this word as analogous to “nation,” but rather, they used phoneme-grapheme rules, coming up with a spelling such as “jayshun,” while 33% of the ten-year-olds made analogies and 50% of the college students did. Again, the same criticism may be raised: The words that the participants were asked to spell were nonsense words, and the younger participants were unlikely to know the spellings of the words that these nonsense words were analogous to, which therefore excludes the possibility of the younger readers using the analogous real words as a basis to create spellings.

A fourth study by Marsh and his colleagues (Marsh, Friedman, Desberg, & Saterdahl, 1981) demonstrated that children can make analogies when they understand how to do so. Two groups of children, ages seven and nine, were given a list of nonsense words. The participants were also provided with the real words to which the nonsense words were analogous, and were told that each nonsense word was created by changing one letter of a real word. In this experiment, the seven-year-olds made analogies 78% of the time and the nine-year-olds 92% of the time. Although these results demonstrate that children can make analogies when they are instructed to do so, it is still unclear whether children are likely to do this spontaneously when reading.

Goswami and East (2000) examined the effects of training children to use onset-rime segmentation. Researchers gave an onset-rime segmentation task to 32 five-year-olds. Half of the sample was unable to perform the segmentation task at all prior to instruction (receiving scores of zero). The children were then trained to read words based on rhyme and analogy methods, and results demonstrated that all children were able to divide words into onset-rime segments following this intervention. The participants' initial onset-rime segmentation ability and improvement in onset-rime segmentation ability were not correlated with reading ability. Goswami and East interpreted these findings to mean that onset-rime segmentation performance is not correlated with reading ability, because all children can learn to segment words into onset and rime units, regardless of their reading level.

Reading instruction that children are exposed to also affects whether or not children will adopt analogy-making as a decoding strategy. Deavers, Solity, and Kerfoot (2000) examined the role of early reading instruction on the nonword reading strategies employed by beginning readers. 45 beginning readers from classrooms emphasizing small units instruction (grapheme-phoneme correspondences), onset-rime units in conjunction with rhyme awareness, or a combination were asked to read nonwords. Children being taught with a curriculum emphasizing onset-rime units demonstrated a preference for rime-based strategies and made significantly more use of them compared to the other groups.

Comparison of onset-rime segmentation and phonemic segmentation

Levy, Bourassa and Horn (1999) compared onset/rime segmentation, phonemic segmentation, and whole word training on poor second graders' reading. Results demonstrated that participants showed poorer retention of training words under the onset/rime treatment condition when compared with phoneme level training. Generalization to new words was best following phonemic segmentation training, but rime/onset was better than whole word training.

Post, Carreker, and Holland (2001) compared phoneme-based training and rime-spelling instruction. They trained 63 first graders to spell five final letter patterns in monosyllabic words over the course of ten days. The participants were trained under one of two conditions: using phoneme-based instruction that emphasized the direct relationship between final speech sounds and their spelling patterns, or linguistically implicit instruction that focused solely on the spelling of the rime. The group trained with phoneme instruction outperformed the rime instruction group on posttest reading and spelling measures, suggesting that phoneme-based spelling instruction is superior to rime-based instruction.

Bruck and Treiman (1992) trained 39 first grade children to read pairs of words with shared phonemes. The pairs either shared rimes (pig and big), vowels (pig and rib) or heads (pig and pin). The researchers measured how quickly children could master the words and also retention of the words. The group taught to read words with shared rimes learned to read words fastest, but were significantly worse than the other two groups at retaining those words. Interestingly, the students were drawn from schools which the researchers describe as ones that “strongly adhered to the whole-language philosophy” and their observations revealed “no systematic teaching of phonics” (Bruck & Treiman, 1992, p.378.)

Walton, Walton, and Felton (2001) compared the effects of teaching pre-readers to use rime analogy or letter-recoding. Results of their study, involving 127 children in kindergarten and first grade, showed that both groups benefited on word reading ability measures when compared to a non-contact control group. This pattern remained consistent even for participants with weak letter sound and phonological skills. Similarly, O’Shaugnessy and Swanson (2000) found positive outcomes for reading fluency among reading-disabled second graders following six-week treatment interventions that involved either onset-rime instruction *or* phonemic instruction.

Some reading programs, such as the National Literacy Strategy in the UK, advocate teaching both grapheme-phoneme correspondences and onset-rime strategies at the same time, in recognition of the fact that both are ways that successful readers use to process words (Bangerter, 2002). Goswami (2001) states that she agrees that “phonemic awareness is related to reading," and concedes that it is "important to teach children print-sound correspondences at both onset-rime and phoneme levels." These statements represent a stronger emphasis on the importance of phonemic awareness than Goswami's writing contained previously. In her earlier writing Goswami (1999) suggested that "teaching reading by analogy will also mean teaching a child letter-sound correspondences, but only when these letter-sound correspondences are onsets." (Goswami, 1999, p.218).

On the basis of these findings, it seems that the ability to take advantage of the fact that lists of words share common rimes is most probably only useful in the presence of solid grapheme-phoneme connection-making ability. Teaching beginners to make rime-based analogies should certainly not replace instruction designed to develop knowledge of individual letter-sound relationships. Additionally, awareness of onset-rimes may need to be enhanced through training in order for children to use onset-rimes as a strategy for dealing with words in word families. This should be taken into consideration when designing instruction, especially for education programs intended for novice and disabled readers.

Teaching words families by relating single unknown words to known words

Peterson and Haines (1992) examined the effectiveness of teaching analogy-making using word families. They trained kindergarten children to use rime analogies by using a word family approach. Children were presented with a single word (i.e. "ball") and then taught to divide the word into onset and rime ("b-all"). Then the participants were shown a second word (i.e. "hall"), and encouraged to decode it by dividing it into onset and rime units ("h-all"). The children were asked to compare the rime patterns. This instruction continued with other words in the word family. Over the course of a month, children were taught ten word families. Results revealed that the treatment group significantly outperformed the control group on a measure of reading new words by analogy, as well as word segmentation ability and letter-sound knowledge. This indicated that onset-rime training was generalized to transfer reading tasks and participants were able to use onset-rime segmentation as a decoding strategy on subsequent reading tasks.

White and Cunningham (1990) developed a program in which they taught low SES children to read by introducing what they termed an "analogy strategy" early on, and then systematically teaching children word families, such as "-at," "-ine," and "-ob." Words were posted on a word wall, and when new words were presented in these families, the children had to categorize them by family. Results from testing a year later demonstrated that the group in the analogy program had significantly higher scores on measures of decoding and comprehension when compared with a similar control group that was exposed to regular classroom instruction. This study lends support for the use of word families in early reading instruction.

It should be noted that the details of the methodology involved in the studies with successful outcomes can not be overlooked. For instance, in the above example, students were *not* given the entire word family to process, but rather were forced to analyze specific words and *then* relate them to words with the same rimes that had been learned previously.

Reading by analogy among dyslexic readers

The term “dyslexia” has been used as a way of describing various reading difficulties. According to the American Psychological Association, DSM-IV, dyslexia is a failure to learn to read despite the fact that the individual has received adequate reading instruction, and this problem is not caused by sensory, intellectual, emotional, or socioeconomic handicaps. Related problems, such as the inability to spell, and to segment phonemes are viewed as evidence of language processing difficulties underlying dyslexia (Scarborough, 1990). Dyslexic readers’ ability to use analogies as a decoding strategy is of interest when considering young readers, because although dyslexics are older, their reading level is similar to that of younger students.

Lovett, Warren-Chaplin, Ransby and Borden (1990) systematically trained dyslexic children to read a set of words, such as “cart” and “peak.” Afterwards, they presented the participants with a set of words that shared the rime spellings but not the onsets of the training words, such as “part” and “weak.” They found that the dyslexic children performed no better on this second task than did a matched control group that had received no training in reading at all. The researchers concluded that the dyslexics failed to spontaneously make use of the rime-based analogies to identify the unfamiliar words because they were unable to divide the syllable of the word into the onset and rime parts.

Hanley et al (1997), presented nine dyslexic children, aged nine to eleven, with “clue words” followed by analogous test words, similar to those used by Goswami

(1988b). The results of their experiment indicated that the dyslexic readers read significantly fewer of the analogous words than did a reading-age matched comparison group of younger children. In fact, even the dyslexic child with the highest score on this task read fewer words correctly than did the lowest-scorer in the comparison group. In a second experiment, the researchers presented clue and test words to 23 dyslexic children, ages eight to eleven, after the children were brought to criterion in reading the clue words. Again, the researchers found that the dyslexics read fewer words than did the reading age-matched comparison group. Based on their results, the researchers argue that a failure to make analogies may be a main cause of the reading impairment displayed by children who have developmental dyslexia.

If analogies are so important for building early reading skills, as is argued by some researchers, then it is critically important to understand what happens when children are exposed to analogous lists of words (families) without explicit instruction, especially since this is a common classroom instructional technique.

Training Students to Read Lists of Words

The selection of words in training studies is a complex issue and can influence the ability of children to learn the words. The present study incorporated the patterns “ck,” and “ll,” and also two that were harder “le” for /l/ and “rk.” Of interest is whether the former or the latter will prove easier for children to learn.

Post, Carreker, and Holland (2001) found that their participants had more difficulty learning rime patterns that included two phoneme blends (i.e. “sk,” “nk”) and final silent e preceded by a long vowel, as in "cake") than they did learning words which contained sounds represented by digraphs (i.e. "sh" and "ck.") This suggests that training effects in studies are influenced by the specific words involved, their level of complexity, the children's familiarity with spelling rules, and the sound sequences.

Word lists are more confusing for beginners when the words are similar to one another. Spring, Gilbert, and Sassenrath (1979) gave kindergartners either a list of similar words (*pots, post, spot, stop*) or dissimilar words (*play, fire, bugs, honk*) and found that the group learning similar words needed more trials to master the words in whole-word training that did not involve teaching segmentation and blending. In comparing children's word learning when providing children with similar and dissimilar lists, Gilbert, Spring and Sassenrath (1977) found that when they included a group that was trained with dissimilar lists, but trained beyond the point of reading criterion (criterion plus six overlearning trials), results did not differ on transfer tasks compared to results for the students who were trained to the point of criterion. It takes children longer to master words in sets when the words are similar to one another. It is therefore important to provide enough trials in a training study to allow for participants to secure words in memory. However, in the final results, training involving similar words is likely to result in greater memory for those words. Children in the Spring et al (1979) study who were trained with similar words made fewer errors on subsequent reading tasks.

Pilot Study

A pilot study was conducted to examine the effect of presenting word families on children's word reading and spelling. To answer this question, three training conditions were created. In the pure "word families" condition, students were taught to read four lists of words, each comprised of four words from one family. The second group was taught to read the same words in mixed rather than pure lists. A third, untrained group participated in pre and posttesting only. Table 1 displays the words and learning conditions used in the pilot study.

In the pilot study, 32 kindergarten children were the participants. First, they were given pretest measures to assess letter-names, letter-sound knowledge, ability to read preprimer words and words taught during training, and ability to invent spellings for the training words. On the basis of scores on these measures, two groups of students were formed, called lower and higher achieving readers. There were sixteen children in each group. From each group, 12 randomly selected participants were then assigned to one of two treatment groups at random: a word family group, or a mixed-order group. A control group was created from the remaining eight students. This control group received no training and only completed the posttest measures when the other students had completed training. Half of the students in each training and control condition were lower achieving readers and half were higher achieving readers.

Following these pretest tasks, participants in the experimental groups were taught to read four lists of words from four word families. (See Table 1). Students were shown each word, told its pronunciation, and given the word in a sentence to explain its meaning. Participants then practiced reading the words for eleven additional trials. The first group was taught four words from one word family on each of the four treatment days,(i.e. night, light, right, tight) while the second group was taught four words, one from each word family on each treatment day (i.e. night, hump, bark, seed). A control group was not exposed to the words between pre and posttesting. Posttesting tasks involved reading training words, reading new words and nonwords analogous to families that were taught, inventing spellings of selected training words, and a spelling recognition task involving selected training words. Participants' scores on these tasks were analyzed through the calculation of a mean score for each group on each task, and then scores of the three groups were compared. Of interest was whether training impacted posttest performance and whether effects were similar for the lower and higher achieving readers in the groups.

The pilot study examined the effects of presenting word families on children's word reading and spelling. Based on findings of Spring et al (1979), it would be expected that training students to learn dissimilar lists of words would result in slower mastery of the words during training, but greater recall of words on subsequent reading tasks. Kindergarten children would be expected to fit Ehri's partial alphabetic phase, characterizing students who identify words by processing partial letters, often only the initial letter in each word. Participants learning words in pure families at a minimum would only need to identify the initial letters and then blend those sounds to the oral rhyme of the rime. However, participants in the mixed lists groups would need to examine other letters besides the initial letter to identify words during training.

One possible explanation for more word learning following training with mixed lists is that presenting words in mixed lists forces children to attend to more of the phoneme-grapheme correspondences available in a word to remember how to read them, something that may not happen when children are taught lists of pure word families. Because instruction involving word families is a common part of early literacy education, it was of interest to explore if grouping words in this way served a useful purpose for emerging readers, especially since Goswami and others have argued that word learning is facilitated through onset-rime awareness.

Results of the pilot study displayed in Table 2 indicated that lower-achieving readers who underwent word training with pure "word families" remembered how to read an average of 1.5 words (out of a possible 16), on the posttest which was not significantly different from the control group who received no training and who learned no words. However, the average number of words learned by low-achieving children in the mixed-order condition was 5.66, which was significantly greater than the other two groups. Thus, training children to read words from word families by mixing the words taught on each day helped them distinguish among the words and remember how to read them better than training children to read words all from that same family on each day.

Results of the pilot study indicated that for higher-achieving beginning readers, teaching words in pure lists of word families did not impact word learning any more or less than presenting words in mixed lists. Both types of training produced superior learning when compared to the third group that received no training. (See Table 2). One possible explanation for this finding is that the more advanced readers were already able to process all the phoneme-grapheme connections contained in individual words to remember how to read them, so that the makeup of the sets of words taught together had no significant effects upon word learning.

There were no significant differences on the spelling posttests between the three groups, with the exception of one task requiring students to spell four nonwords analogous to training words. On this task, the two training groups' performances were not significantly different from one another but were significantly better than the control group's performance, and this finding held true even when results were analyzed separately for low achieving and high achieving readers. The spelling recognition task, based on a task designed by Masterson (1983), turned out to be a poor choice for testing the young participants in the study. The task overwhelmed them by requiring that they examine eight choices simultaneously.

Additionally, it should be noted that there were ceiling effects on scores, since the training only involved 16 words. This affected some participants more than others, since some participants entered the pilot study already knowing how to read a few of the training words. In the present study, children who could read more than two training words by sight, or more than one training word from any family, were excluded.

Pesent Word Families Study

The present investigation of word families was unique in that it compared beginners' ability to remember how to read and spell the same words under different learning conditions in a manner that could provide evidence for the justification of teaching methodology. This project also allowed for the exploration of the various strategies of word reading and spelling that beginners use. Of particular interest was an examination of beginners' reliance on initial letters for word reading, since previous studies have suggested that the letter in the first position is the most likely to be learned (Marchbanks & Levin, 1965; Rayner, 1976.) In the author's pilot study, both experimental groups had the opportunity to rely on an initial letter strategy during training, since the four words in each set in both conditions all began with four different letters. In this study, a third experimental group was added to control for this. In this condition, the students were taught words in sets where the initial letters of all four words were the same, but the rimes were different. The words that were taught are presented in Table 3. Otherwise procedures in the study closely resembled the pilot study procedures.

Rationale and Hypotheses

Although beginning readers can use decoding skills to process words, they often identify words by relying on context or only utilizing some of the alphabetic information available, such as initial and final letters. Words need to be fully analyzed a number of times by beginning readers before they can be well-secured in memory as sight words, without the need for sounding out and blending (Ehri, 1994, 1995). In the present study, words were practiced in isolation, removing the possibility of reliance on pictures or textual information as a source for identifying words. The study was also designed to explore participants' reliance on an initial letter strategy.

The present study explored ways to strengthen word learning of beginning readers by examining how the construction of word lists affects beginners' reading and spelling. Three types of training were compared. All trained participants read the same 16 words drawn from four word families, but the words were organized into three different types of lists. In the first treatment condition, participants practiced reading lists of words grouped by family, meaning that the words on each list shared rimes. In the other two treatment conditions, participants practiced reading mixed lists of words that did not share rimes. These latter treatments involved lists of words either with different initial letters or the same initial letter. Also, a no-treatment group control was included.

In recent years a number of studies have been designed to test whether children have an easier time learning to read words that share rimes (Treiman & Bruck, 1992; Goswami, 1986). However, these studies compared students' ability to learn rimes compared with other words. The present study was designed to teach all participants the same list of words to examine the effects of grouping words by family.

It was hypothesized that teaching students to read lists of words organized into pure families would result in superior reading performance during the training sessions when compared with students taught mixed lists. This is because participants learning pure families limit their attention to the initial letters of words and then apply the oral rhyme to say the rest of the words. By joining the initial letter sound with the rhyme, words could be identified correctly. It was expected that it would take students in the mixed lists conditions longer to master the words during training, but ultimately they were expected to demonstrate superior performance on posttest reading and spelling when compared with students trained with pure word families. This outcome was expected because the mixed list conditions required students to attend to other letters in the words besides the initial letters, and more extensive processing of letters should lead to greater memory for the words. Bruck and Treiman (1992) found that students trained to read words that shared rimes mastered words quickly during training. However, they were unable to secure these words in memory.

Findings from previous studies (i.e. Gilbert et al, 1977) suggest that it takes students longer to master lists of words that are similar to one another. However, students trained with similar words which required careful attention to the letters resulted in greater memory for the words.

Both mixed list treatments were expected to prove superior to training involving word families lists, with an added advantage for the participants exposed to mixed lists that shared initial letters. Mixed lists with shared initial letters removed the possibility that participants were relying on initial letters to identify words.

Beginning readers with better decoding skills can process words by examining all the letters in words and assigning sounds to them. In contrast, beginners with poor decoding skills often do not have solid knowledge of all the grapheme-phoneme relationships, particularly short vowels. It was expected that there would be a decoder-level by treatment interaction revealing that high decoders read words equally well regardless of how word lists were organized, whereas low decoders showed superior word learning when words were presented in mixed lists. This was the pattern observed in the author's pilot study.

It was expected that high decoders would have an easier time learning to read and spell words compared to low decoders. This study included participants who had full knowledge of the letter names and sounds as well as participants who lacked full knowledge of letter names and sounds. Numerous studies have demonstrated that students with good decoding skills outperform students with poor decoding skills on reading and spelling tasks. It was of interest to examine which other pretest factors best explain variance in word learning. It was expected that rhyming ability would explain some of the variance because research has shown that beginning readers with good decoding skills usually demonstrate competency on rhyming tasks (Ellis & Large, 1987; Bradley and Bryant 1983, 1985).

For beginning readers, some orthographic patterns present a great challenge. It was expected that students in this study would have a harder time processing and remembering words with the rimes –ark and –ale compared to words with the rimes –ill and –uck. Post et al (2001) found that beginners had difficulty learning words with rime patterns that included two phoneme blends (i.e. "sk," "nk") and the long vowel final e pattern (as in "cake") than they did learning words which contained sounds represented by digraphs (i.e. "sh" and "ck.") Treiman, Zukowski, and Richmond-Welty (1995) demonstrated that children have a harder time spelling words that contain two sounds in CVCC words, and often use only one consonant to represent the final two consonants when spelling words with this pattern.

The present study tested the following six hypotheses.

1. Presenting students with lists of words organized in pure families will result in superior word learning during the training sessions when compared to presenting students with mixed lists.
2. On posttests, students trained with mixed lists will demonstrate superior performance on reading and spelling tasks when compared to students who were trained with pure word families.

3. Participants trained with mixed lists that share initial letters will have an added advantage over the group trained with mixed lists having different initial letters. Slightly superior memory for words and spellings on posttests will be demonstrated.
4. Results will reveal a treatment by decoder level interaction. High decoders will learn to read and spell words equally well, regardless of how the word lists are structured, but poor decoders will demonstrate superior performance when trained with mixed lists.
5. Two pretest factors will best explain the variance in word learning: decoding skill, and rhyming ability.
6. It will be more difficult for participants to learn to read and spell words containing the rimes –ark and –ale when compared to words containing the rimes –ill and –uck.

Chapter 3

METHOD

Participants.

The sample included one pre-kindergarten student, 62 kindergarten students, 8 first graders, and 2 second graders. There were 34 males and 39 females. The mean age of participants was 69.7 months.

The participants were drawn from two urban public schools, one in Queens, N.Y. and the other in Brooklyn, N.Y. Both schools were in middle-class neighborhoods. However, children from lower socioeconomic areas were bussed in to attend the schools. The Brooklyn school was a large school (several hundred students) and the Queens school was a small, new school, (with a total of six classrooms of about 20 students each) to which students needed to apply and were chosen by lottery, with siblings of already enrolled students given placements. (Insideschools, 2004)

Both schools had many children who had second languages spoken in their homes, which is fairly typical in New York City. This was assessed by asking children if someone in their home spoke another language. There were 55 students who reported a second language in the home, which is about 75% of the sample.

Additionally, both schools had a student body comprised of children from ethnically diverse backgrounds**.** The reported ethnic group breakdown for the school in Queens was 11% White, 47% Black, 17% Hispanic, and 27% Asian; the Brooklyn school had a student body that was 18% White, 3% Black, 24% Hispanic, and 57% Asian. Both schools reportedly average a 97% attendance rate. The larger Brooklyn school had about 90% of children eligible for free lunch, while the smaller Queens school only had 35% eligible for free lunch. Additionally, the Brooklyn school was ranked with four out of five stars on reading and math scores. No information about test scores was available for the Queens school. (Insideschools, 2004).

Informal observations were used to assess the types of instruction and activities used to teach literacy. A variety of literacy activities were evident, some of which may be described as phonics-based. Posters displaying "word families" were present in almost every classroom. The kindergarten classrooms seemed to devote only a small amount of instructional time to children's reading and writing.

The research was conducted in January at the Queens school and in June at the Brooklyn school. Letters describing the research project, along with consent forms, were sent home to children. In the first school, which was a small, new school with one pre-K class, two kindergarten classes, two first grade classes, and one second grade class, letters were distributed to all Pre-K through second grade students. Several first graders and two second graders participated, and were at the midway point of the school year. In the second school, which had grades Pre-K to 5, there were seven kindergarten classes. Since research was being conducted in June, it was most likely that few first or second graders would meet study criteria, and therefore letters were given only to kindergartners. All students whose parents signed consent forms were given pretests. To qualify for the study, students had to be able to identify at least thirteen letters of the alphabet but be unable to read more than two of the training words from different families. There were 73 students included in the study. There were 17 students who were pretested but did not match criteria, seven because their knowledge of the letters was too limited (they could not identify more than ten letters) and ten because they could already read several training words.

In the sample of 73 students, there were three outliers in terms of their grade level, one per-K student and two second graders. Their pretest scores were compared to kindergarten and first grade students comprising the bulk of the sample. The pre-K student's scores on all measures were comparable to the average of the first grade group. The only difference between the pre-K student and the first graders was that they recognized, on average, one training word (usually "duck,"), while he did not recognize any. Additionally, at 56 months, he was significantly younger than the rest of the sample (i.e. average for kindergartners, 68 months old; average for first graders, 78 months).

The two second graders' scores on pretest variables sometimes resembled the first grade group and sometimes were closer to the average of the kindergarten group. In general, their scores were more similar to the first grade group, on measures of letter names, Roswell Chall Auditory Blending, rhyming tasks, and invented spelling. However, the average for the second graders' letter-sound knowledge was slightly less than the kindergarten mean (20, versus 20.37 for kindergarten and 22 for first grade). The second graders knew only two of the short vowels, compared to 2.6 for the first graders and 2.8 for the kindergartners. One second grader (the girl) was unable to read any nonwords, but the other second grader (a boy) read three nonwords, recognized two of the training words, read more preprimer words than first graders (20, compared to an average of 15 for first graders), and read more Woodcock Reading Mastery words (27 compared to an average of 23 for the first graders). The two second graders were significantly older than the rest of the sample (89 and 90 months, compared to the first graders, who were 78 months old on average).

It should be noted that age did not distinguish between high and low readers. High decoders were on average, 71 months of age, while low decoders were 69 months. In fact, the majority of students in each group were between five-and-a-half and six-and-a-half, a typical age for beginner readers. The critical variable in this study was not age but the literacy level of participants, and for this reason, numerous pretest measures were administered.

Treatment groups were created using random assignment stratified by achievement level and gender. Based on pretest scores on the Woodcock Reading Mastery Word Attack test, students were classified as low or high achieving beginning readers. Students who scored zero on the test were considered low, and those with a score of one or higher were considered high. There were two exceptional students with very high scores on the Woodcock Reading Mastery word identification subtest but scores of zero on the word attack subtest. For the sake of consistency, these two students were assigned to the lower decoder group.

Members of high and low groups were separated by sex, and randomly assigned to one of the four conditions. The stratification procedure resulted in unequal numbers of males and females in the high and low groups. There were 19 males and 20 females (39 students, or 52.1% of the sample) in the high group, and 15 males and 19 females in the low group (34 students, or 47.9% of the sample).

The students' grade levels were not a factor in assigning students to treatment groups. The pre-K student and one second grader were classified as high decoders and one second grader was a low decoder, on the basis of their Woodcock Reading Mastery Word Attack scores. As a result of random assignment, the two second graders were in the control group, and the pre-K student was in the Pure Word Families group. Because of the possibility that these students were exceptional (possibly gifted or reading disabled), after regular analyses were done to determine findings, the results were re-analyzed with these three students excluded.

Materials.

Single syllable words with four letters were selected to serve as training words. The words were from four word families (rimes) and the set of words in each family had the same four initial letters. (See Table 3). These words were presented on cards in various orders during the training sessions.

Pretests. Several tasks were administrated individually in the order listed below. Unless otherwise noted, no training trials or feedback was provided. (For a complete record of experimenter-designed tasks and scripts, see appendixes A-E).

1. Letter-name knowledge. Students were shown and told to name 25 upper case letters plus lower case *i* presented in random order on a sheet of paper. The letters were presented in this way (all uppercase and lowercase i) because beginning readers typically have an easier time with capital letters, with which they are more familiar, but in order to avoid confusion, capital I, which may be read as a lowercase l, was replaced by lowercase

i, which is distinct. This format was followed on all tasks being presented to children throughout this study. This task had a maximum score of 26.

2. Letter-sound knowledge. Participants were asked to read across the same randomly-ordered alphabet sheet and tell the letter-sounds. The maximum score was 26. For letters that had two sounds (c, g) if one or both sounds were given, this was scored as correct. Vowels were scored as correct if the short or long sound was given. It was impossible to determine whether the child was naming the letter or elongating the letter-name to represent the long vowel sound. Accepted answers for X included (ks, z, and /k/), and Q was scored as correct if the child responded with /kw/ or /k/. If a child responded to a consonant by identifying the letter name, he or she was told, "The name of that letter is ____, can you tell me what *sound* the letter makes?" If the incorrect letter name was given, no feedback was provided.

3. Vowel Task. This task assessed students' knowledge of short vowel sounds. Students were asked to look at three nonword choices and circle the one that made the sound the experimenter pronounced. (For example, they were presented with uv/ev/av and asked, "Which one says *ev*?") The students were introduced to the tasks with a sample item, to be sure they understood the task. For this item, they were shown the letters M, K, and R, and asked to circle the one that made the sound /k/. (Only two students did not get this item correct; those who did not had K pointed out to them and were told, "That one says /k/, so that's the one you are supposed to circle." No other practice items were given).

4. Reading Words. Participants were asked to read the 16 training words and 24 preprimer words (such as "little" and "me") presented on cards. The words were presented in a mixed order to minimize attention to the training words and their recurring

rimes. Correct responses and errors were recorded. Also, any evidence of strategies was noted, for example, when a child "sounded out" a word out loud. If children said that they did not know a word, they were told, "That's okay, let's try another one." Correct responses were scored in two ways: first, scoring words as correct only if they were recognized immediately, and second if they were either recognized immediately or sounded out and blended back together correctly. Immediate recognition of the word was interpreted to mean that the word was read from memory by sight. Any child who read three or more training words, or two from the same family, was excluded from the study, in order to leave room to observe training effects on word learning.

4. Woodcock Reading Mastery Tests –Revised (Woodcock, 1987). Subtest of Word Reading, Form H. Participants were asked to read a list of words of increasing difficulty presented on cards.

5. Woodcock Reading Mastery Test, Word Attack, Form H. Participants were asked to read a list of nonwords presented on cards.

6. Roswell-Chall Auditory Blending Task. (Roswell & Chall, 1992). Participants listened to a list of words, each spoken so that the phonemes were separated by a half second (i.e. "too" presented as /t/ - pause - /oo/). They were asked to state the word that they heard. Each correctly identified word was assigned one point, for a maximum score of 30.

7. Rhyming task. There are no standardized tests to determine rhyming ability. Therefore, participants were given researcher-created rhyming activities.

The first task involved an oddity task, requiring student to listen to a set of four words and choose the one that had a different rhyme from the other three. The sequence of words was repeated three times to minimize ordering effects on choice. For example, the child heard, "Fell, go, no, blow, fell, go, no, blow, fell, go, no, blow," and had to select "fell." This task was introduced with an explanation of rhyming words ("words that sound the same") and two practice items. If the child got the practice items right, he or she was told, for example, "That's right. Cap, lap, and tap all sound the same, they rhyme, but *book* sounds different." If the child chose an incorrect response on the practice items, the correct answer was given and then the explanation was provided. One point was assigned for each correct response, with a maximum score of 12. No feedback was provided during test trials.

The second task asked participants to invent rhymes for given words. Practice items were provided. If the child did not give a correct response, he or she was given the correct answer and an explanation. For example, the child was asked to provide a rhyme for "me." If they did not give a correct response, they were told, "If I say 'me,' you could say 'key.' Or if I say 'me' you might say 'bee.' Or you might say a silly word, like 'ree." Care was taken to explain that there was not just one correct answer and that nonwords were acceptable responses. Any rhyming word was accepted, whether it was a real or nonsense word, and rimes were accepted as correct answers (i.e. "eep" as a rhyme for "deep.") Each correct response was given one point, for a maximum score of 10. The combined rhyming tasks were totaled for a maximum score of 22.

8. Spelling Production. Participants were asked to write spellings for 12 words. This list did not include training words to avoid priming students, but was a more generalized test of spelling, adapted from Santa and Hoein (1999). Students were instructed to listen for the sounds in the words and make their best guess, writing some if not all of the letters. They were introduced to the task by having the spelling word "nose" modeled for them,

with an explanation about listening for the sounds in the words and writing letters for those sounds. This task was scored according to the method of Santa and Hoein, with some slight changes in the number of points awarded for certain spellings (see Appendix F for a complete explanation). The general rubric called for awarding one point for providing an initial letter or letters for the first sound in the word that was deemed to be correct or phonetically appropriate. Two points were given for providing, in addition, a letter or letters that represented a second sound in the word, either correctly or phonetically. Three points were given when a third and/or fourth sound was represented in this way, and four points were awarded for representing all the major sounds in the word in way that demonstrated knowledge of the orthographic system, even though the words were not spelled correctly (i.e. adding a final letter – “dresse” for dress, or representing a vowel with a phonetically correct pair of letters that was not the correct spelling – “roed” for road).

The maximum possible score was 60, and this was not achieved by any of the children. These spellings were scored by this author and by an independent rater, following the rubric and samples given by Santa and Hoein (1999), with inter-rater agreement occurring 99% percent of the time.

<u>Training Procedures.</u>

Children were taught to read 16 words, comprised of four words from each word family (-ale, -uck, -ill, -ark) with initial letters b, p, m, d. The words were presented to each child individually on cards. During the first study trial, the children were told the word once (“That word says ‘park.’”), were given a sentence to clarify meaning (i.e. “I like to go on the swings in the park.”), were told the word again and asked to repeat it (“Now you look at it and read it to me.”). They repeated this process with the four words in that day’s set. Any comments such as, “These words rhyme” or “they all start with b” were recorded. (See Appendix G for script.) During subsequent test trials, they practiced all four words, presented in random order, for 11 trials. This number of trials was chosen because scores from the pilot study indicated that this number of trials was more than sufficient for every participant to master the words. The experimenter did not point out any letter-sound correspondences, rhymes, or strategies for remembering/decoding words, but rather, treated the words as wholes, and each word as separate, unconnected to the others. Any time the child failed to read a word correctly, it was stated for him/her (i.e. “That word says, duck.”). Negative feedback, (i.e., saying “no,” or “wrong”) was avoided. Words were counter-balanced across the days within the groups. (See Table 3.)

Each student read the four words in their daily set in different orders across trials. Participants were given 11 trials for practice regardless of how many words they read correctly. Results of the pilot study indicated that children in both experimental groups were able to master words well before this number of trials, and this ensured that both groups saw the words an equal number of times.

Posttests. The posttests were given 3 or 4 days following the last training session. (For a complete record of experimenter-designed tasks and scripts, see Appendixes C, H, and I.)

1. Reading. Students were asked to read all 16 training words and the set of preprimer words, which were ordered and scored the same way as in the pretest.
2. Spelling Production. Students were asked to write eight of the training words and four new nonwords. Participants were told, "Remember the words we did together?" as an introduction to the task, but were not told to try to remember the training word spellings. This language was chosen deliberately to hint to the trained groups that they were supposed to try to remember spellings based on the training words yet still make sense to the untrained control group, so that the language would be identical for both trained and untrained participants. Participants were instructed to write some letters if they did not know all of them, and to "write the ones you think are right." The four nonwords were introduced with an explanation of "silly" words ("a word you can say, but it doesn't mean anything") and an example showing how the nonword "yad" could be written ("See, "yad" doesn't mean anything, but I can say it and I can write it down.") More words were not tested because writing spellings took some children an extraordinarily long time. More than 12 words would have consumed more than 20 minutes, and would have exceeded children's capacity to remain focused.

   This task was scored in two ways: first, by counting the number of words spelled correctly, for a maximum score of 12, and second by counting the total number of correct letters in each word, for a total score of 48. The second scoring method meant that a child who wrote the four correct letters but added an extra letter (i.e. a response of "dille" for "dill") or a child who confused the order of letters (i.e. spelling of "dukc" for "duck") received as many points as a child who spelled the word correctly. None of the participants reached ceiling.

3. Spelling recognition task 1.

A multiple choice spelling recognition task was administered. The targets were the 16 training words, and eight new transfer words. Half of the new words had two consonants in the onset. All new words contained the same rimes as the training words.

Each multiple choice item contained the correct spelling for a word and two foils. One foil was created by changing one of the middle letters of the word (i.e. bill to "bell" or "bull," bark to "bank" or "back") and the second foil was created by constructing a phonetic spelling (i.e. "bil" for bill, "duk" for duck). For example, one item consisted of the words "duck," "dunk," and "duk." Test items were presented one at a time in random order.

Procedure. Each test item was presented to the participants orally while they looked at the three choices, and students circled their choice with a pencil. For example, participants were told, "These words all say 'dill.' Circle the one that you think is the *right* way to spell 'dill.'"

Scoring. Each correct answer was awarded one point, for a maximum score of 24.

3. Spelling Recognition Task 2. A second spelling recognition task tested participants' ability to select correct words. Four words were shown to the children. Each was presented twice, once in its pure family (i.e. selecting duck out of puck, duck, muck, buck) and once in a list of words sharing the same initial letter (i.e. duck from duck, dark, dale, dill). Each correct response was awarded one point, with a maximum of four for the first type of choice and four for the second, with a combined score of eight. This task also

allowed children a 25% chance of guessing the correct answer, which would have given them a score of 2 by chance alone.

Procedures.

Children were trained and tested individually. Pretests and posttests were completed in one session lasting approximately 30 minutes. Training was conducted across four successive days, each session lasting approximately 18 minutes. The author trained and tested all children. The time elapsing between pretesting and training and between the completion of training and posttesting was three or four days.

Chapter 4

RESULTS

Pretest Performance.

Participants were all beginning readers as indicated by the fact that they identified most of the letters and many letter-sounds but were not yet able to recognize the training words immediately upon seeing them. To distinguish beginning readers with stronger and weaker decoding skills, participants were divided into two groups, low and high decoders, on the basis of their Woodcock Reading Mastery (1987) Word Attack Scores: those who scored 1 or higher were considered high decoders, those who scored 0 were considered low. Performance of high and low decoders on pretest measures is presented in Table 4. Separate ANOVAs were conducted using decoding level as the independent variable and the pretest measures as the dependent variables. These analyses confirmed that the high and low decoders differed significantly on every literacy measure (all ps < .05) except age and the rhyme oddity task. F values are reported in Table 4.

From mean values in Table 4, it is apparent that despite the finding of a statistical difference, both high and low decoders knew the names of most letters. They differed more in their knowledge of letter-sounds but both groups still knew the majority. However, neither group had mastered short vowels. Differences between the groups were substantial on three phonological awareness measures, though performance fell short of mastery. The one task failing to distinguish the two groups was the rhyme oddity multiple choice task, with both groups performing at a chance level. On the word reading tasks, high decoders could read many more words than low decoders. Their grade equivalent scores on the Woodcock Reading Mastery word identification subtest corresponded to mid-first grade for the high decoders and early-first grade for the low decoders. Slightly more than half the sample scored at or below 10 Woodcock Reading Mastery words correct, but a large number of students (another 40% of participants) read between 12 and 28 words correctly, and five students even read between 31 and 39 words correctly. Scores on the Woodcock Reading Mastery nonword task revealed that while the low decoders did not read any words, 34 of the high decoders read between one and 7 words, with only four participants reading between nine and thirteen words correctly. It should be noted that students were observed to "sound out" some words on both of these tasks, rather than to recognize all the words "by sight." Both groups recognized few if any training words immediately by sight, but the high decoders were able to decode a few of these words. The two groups also differed substantially in their ability to spell words.

Students in the higher decoding group may qualify as being in Ehri's full alphabetic phase, since most had strong letter-sound knowledge and some decoding skill, while the lower decoders were clearly in the partial alphabetic phase**.** On average, high decoders identified more than twice as many preprimer words as low decoders did. However, this short vowel knowledge was not fully developed.

Word learning groups were assigned on the basis of stratified random assignment, which ensured equivalent numbers of high and low readers in each of the four groups. As much as possible, once the reading level was taken into consideration, a balance between males and females in each group was created.

The dependent measures were taken from pretests, training tasks, and posttests. Effects of the independent variables were assessed through two-way ANOVAs, which considered both reading level (high versus low) and treatment group (4 groups).

Performance on pretest measures was examined to verify that the treatment groups did not differ in their letter knowledge, phonological awareness, and reading skills. Mean performance of the four treatment groups on pretest tasks are presented in Table 5.

Two-way ANOVAs with treatment group and decoding level on the independent variables were applied to test whether or not the four groups differed on any skills measured by pretests. There were no significant differences between the experimental groups on any of the measures (all ps > .05). (See Table 5.) This means that the four groups did not differ in their skills when they entered the study, and hence differences in word learning can be attributed to training.

Performance During Training.

In the three word learning conditions, students practiced reading words during each of the four training sessions, each set of four words on a different day. The first group was given sets of words in a family (i.e. mill, dill, bill, pill), the second group was given sets of words from different families with different initial letters in each set (i.e. duck, park, male, bill), and the third group had sets of words from different families but words in each set shared initial letters (i.e. male, mill, muck, mark). (See Table 3 for complete word lists.)

One question of interest was whether or not students trained under different conditions performed similarly in learning to read words during the training sessions. Specifically, were students able to learn to read training word sets with equal ease across the treatment conditions, and did their performance improve to the same extent across training days in the three learning conditions?

To examine children's performance learning to read words during the four days of training, the number of words they read correctly across 11 trials in each session was counted. All children were given 11 trials to read the words. Words read immediately as well as words sounded out and blended were counted as correct. Also, self-corrections were counted as correct. Mean scores are presented in Table 6.

Inspection of children's mean scores across the various treatments revealed ceiling effects (i.e. mean scores above 85%) among high decoders in all treatments and among low decoders in the Pure Families treatment group. That is, children in these groups read the four words taught in each session almost perfectly on most of the 11 trials. However, performance fell short of perfect among low decoders learning the mixed lists. Figure 1 displays mean performance of the groups during training.

Scores of the participants were subjected to a repeated measures ANOVA with treatment group (Pure Families versus Same Initial Letter versus Different Initial Letter in the Mixed Lists), decoding level (high versus low) and Session (1-4) as the independent variables.

Results of the ANOVA are reported in Table 7. Significant main effects were detected for decoding level and for treatment, as well as a significant interaction between these two variables. In order to pinpoint differences, ANOVAs were conducted separately on high and low decoder groups, and post-hoc Tukey pairwise comparisons were conducted to compare mean performance of the three treatment groups.

There was a main effect of treatment group. This shows that it was easier to learn to read sets of words consisting of single families of rhyming words than to learn to read sets containing words in mixed lists. There was also a main effect for decoding level. Higher decoders, taken as a group, read significantly more words during training than did low decoders. The interaction between these two variables was significant.

Figure 1 reveals the source of the interaction. The main effect favoring the Pure group over the two mixed groups held only among the low decoders, but not among the high decoders, who all performed similarly. Moreover, Figure 1 reveals that the low decoders in the pure group performed as well as all the high decoder groups. Post hoc Tukey tests revealed that the two mixed low decoder groups did not differ from each other and both performed significantly worse than the low-decoder Pure Families group. One reason why reading words was so easy for low decoders trained with pure word family sets is that the children only needed to pay attention to the initial letters of the words in the sets because the words all rhymed (i.e. park, mark, dark, bark). In contrast, none of the words within a set rhymed in the mixed lists. Among the higher decoders, the group trained to read words in families demonstrated slightly higher performance than the other two groups, in support of the pattern found for low decoders.

The main effect of Session fell short of significance, and there were no significant interactions between Session and the other two variables. Mean performance during training sessions is plotted in Figure 2, and F statistics are reported in Table 7. One reason that Session was not significant and no interactions were found between Session and the other variables was because the performance of the low decoders was highly variable across sessions. A few children were not able to read a complete set of four words correctly during a single trial by the time they reached the 11th trial. Some of the low decoders in the mixed lists groups were not even able to read every word correctly at least once across the 11 trials. Although mixed lists participants might be expected to improve as training progressed since they were learning analogous words across sessions, there was a great deal of variability among the low decoders, with uneven patterns of performance across sessions. The extent of this variability is depicted in Figure 2. This indicates that the analogous relations among words were not benefiting word learning in low decoders.

Posttest Performance.

The posttest measures assessed the effectiveness of the various word learning conditions on participants' ability to recognize training words and to spell them. In these analyses, performance of a no-treatment control group was included. Posttests were given 3-4 days following the conclusion of the training. This follow-up session was conducted close to the training sessions, because the participants were young and their memories not fully developed.

It was expected that training that involved Mixed Lists with Different Initial Letters would result in superior word memory on posttest reading and spelling tasks when compared to training involving Pure Word Families, and that training with Mixed

Lists whose words began with the Same Initial Letters would show the strongest performance of all.

Reading. A two-way ANOVA with treatment group and decoding level was conducted. Two different measures of target word memory were calculated: (1) the number of words read immediately upon seeing them, as an indicator of sight word learning, and (2) the number of words read immediately plus the number of words overtly sounded out and blended (i.e. a response of "/m/ /ai/ /l/ …male" for "male.") Because some participants had identified a word or two correctly on pretests, gain scores for each participant were calculated. There were no ceiling effects for any participants.

For the measure of words recognized immediately, results revealed that there were significant main effects for treatment group F (3,65) = 9.67, p = < .05, and decoding level, F (1, 65) = 28.30, p = < .05, but no interaction F (3, 65) = 2.08, p > .05. Mean performance is reported in Tables 8 and 9. From Table 9, it is apparent that high decoders recognized more words immediately than low decoders.

Post hoc Tukey tests revealed that participants trained with the two types of mixed lists remembered how to read significantly more words immediately than the untrained control group. In contrast, students in the pure word families group did not read more words immediately than the untrained control group. The two mixed list groups were not significantly different from one another. However, the students trained in the mixed lists condition with words that shared an initial letter did read significantly more words than students in the pure word families group.

These findings support the hypothesis that teaching words in pure families is not as effective for remembering how to read words by sight compared to teaching words in mixed families. Also, they suggest that a special advantage may accrue when words are taught in mixed families whose members all begin with the same letter. The preferred explanation is that mixing families and imposing the same initial letter on sets of words forces students to process letter-sound connections more completely within words to store them in memory.

For the measure of words recognized immediately or sounded out and blended, results revealed that there were significant main effects for treatment group $\underline{F}$ (3, 65) = 7.34, $\underline{p}$ = < .05 and decoding level, $\underline{F}$ (1, 65) = 18.17, $\underline{p}$ = < .05, but no interaction $\underline{F}$ (3, 65) = 1.225, $\underline{p}$ > .05. Mean performance is reported in Tables 8 and 9. High decoders made significantly greater gains from pretest to posttest in reading training words than lower decoders. Post hoc Tukey tests revealed that all three trained groups significantly outperformed the untrained control group, but there were no significant differences between the trained groups.

These findings show that exposure to the words during the learning trials enabled all of the groups to read the words better, either by decoding or by sight, compared to the control group that was not taught the words. However, adding results of decoding skill to the dependent measure eliminated differences among the treatment groups that were evident when immediate recognition of words was scored (see above). This suggests that differential effects of the treatment held mainly for sight word reading.

The sample included three grade-based outliers who were drawn from second grade and pre-K, whereas the majority of the participants were end-year kindergartners and mid-year first graders. The analyses described above were re-calculated with these students excluded because of the possibility that they were exceptional. Results revealed the same pattern as described above for differences among the treatment groups.

Spelling. Separate ANOVAs were conducted to examine posttest scores on ten spelling measures. (See Table 9). The independent variables were treatment condition (Pure Families versus Mixed Lists with Different Initial Letters versus Mixed Lists with Same Initial Letters versus No Treatment Control) and students' decoding level (low versus high). The dependent measures were drawn from the two spelling recognition tasks, and also from the spelling production task in which students were asked to write some of the training words and some transfer nonwords (i.e. nonwords sharing the rimes of the word families that were taught). To score the spelling recognition tasks, the child was awarded one point when he or she circled the target word from a group of three or four words, for a total of 24 on the first task and eight on the second task. The former recognition task had children select training words and transfer words from word choices s spelled three ways: correctly, phonetically, or non-phonetically. (For example, selecting bill from *bill, bil, and bull*). The second spelling recognition task required children to select four of the training words, first from a word family list comprised of training words, and then from words sharing the same initial letter as the target words, also comprised of training words. (For example, selecting bill from *bill, pill, dill, mill* and then selecting bill again from *bill, bark, buck, bale*.)

Spelling production was scored in two ways. The number of words spelled correctly was calculated, for a maximum score of 16, and also the total numbers of correct letters across all the words was calculated, for a maximum score of 64.

There were no main effects or interactions involving treatment groups (p.s. > .05). All scores were generally low, both for spellings created and correct spellings recognized. Table 8 reports means and F values.

Results of the ANOVAs for high and low decoders on the spelling measures are presented in Table 9. Findings revealed significant effects on all the measures, indicating that the high decoders outperformed low decoders. However, inspection of means revealed that even high decoders did not remember how to spell most of the words they learned to read. They remembered 64% of the letters. On average they recalled only 2 out of 16 words correctly. Of course, the spelling tests were given several days after the word learning sessions, so they may have forgotten some of the words. Also, being that the majority of students were kindergartners, and others were first graders at the midpoint of first grade, with limited knowledge of the alphabetic system, especially short vowels, long vowels, and controlled vowels, their ability to retain complete spellings of words in memory is weak.

Students performed extremely poorly on the first spelling recognition task. It was obvious that some were simply circling the word in the first position in the row of three words. It is likely that these children were too inexperienced with multiple choice tasks and were unable to examine all the choices. Also, some points may have been achieved due simply to guessing, since there was a 33% chance of guessing each item correctly. In fact, most children scored at or close to chance level, (i.e. a score of 8.) Only six children got 20 or more words correct out of a maximum of 24, while 44 children scored between 5 and 11 correct.

On the second spelling recognition task, 64 out of the 73 participants achieved a score of 5 or more correct (8 maximum). This was an easy task, not sensitive to training effects. There were too few task items to detect statistically significant group differences.

Relationship between Pretest Measures and Posttest Performance.

Another question of interest was whether any of the pretest measures predicted word learning performance on the delayed posttest. A bivariate correlation was performed, using the difference between pretest and posttest measures of word learning to assess gains in words read correctly immediately, and words read immediately plus words sounded out and blended back together. In a two tailed test of significance, Pearson coefficients revealed that the following variables were correlated significantly with word learning gains: letter sound knowledge, Roswell-Chall Auditory Blending, Rhyme production, Woodcock Reading Mastery Word Identification, Woodcock Word Attack, and Pretest Spelling. All of these correlations were significant at the .01 level. Table 10 displays the correlations. This correlation was then repeated for the participants in the trained groups only, because word learning gains were primarily observed for participants in the groups who underwent training. Results were the same as reported above, with one additional variable, the Rhyme Oddity Task, correlating significantly with both measures of word learning at the .05 level. These results are reported in Table 11.

A linear regression analysis was conducted to determine which variables best predicted word learning performance among the participants in the three training groups. The dependent variable was gain in word reading between pretest and posttest sessions. This was calculated for words read immediately. Because the correlation was very high between the gain scores of words read immediately and words read immediately plus words sounded out and blended (see Table 11), results of the regression were expected to be similar on the two measures. The independent predictor variables were the pretest measures included in the correlational analysis shown in Tables 10 and 11 with one exception. Word attack scores were not included, because all low decoders had been defined as participants who scored zero on this task, and decoder level was considered a as a factor in this analysis.

In order to examine whether or not the pattern of low and high decoders' word learning was determined by the same or different pretest variables, a numeric variable was created for high and low decoder participants, with high decoders assigned a score of 1 and low decoders assigned a score of zero. Centered pretest variables were created by calculating the mean for each variable and then subtracting from each variable its mean. Cross-products were calculated by multiplying these centered variables with the numeric variable representing decoder level.

Regression analyses were performed by entering the centered pretest variables, cross-products, and numeric variable for decoder level. The procedure of backwards elimination was employed, in which the least significant cross-products were removed one at a time until only the significant variables remained, and then centered variables whose cross-products were not significant were removed one at a time, again by first removing the least significant variables.

Results of the regression analysis revealed that two variables explained significant unique variance in the final model. Findings are reported in Table 12. The Woodcock Word Identification test was a significant predictor of the gain in training words recognized immediately for both high and low decoders. Additionally, the Roswell Chall Auditory Blending cross-product was a significant predictor, indicating that this pretest measure predicted differently for high and low decoders. The slope of the regression equation for the low group was close to zero, meaning that it did not predict word learning for low decoders. In contrast, the Roswell Chall was a significant predictor of word learning for high decoders. The total variance explained by these two variables was 0.56, which is substantial.

Exploration of Errors on Reading and Spelling Tasks.

It was of interest to explore the errors participants made when reading training words during the treatment sessions and when reading and spelling words during the posttests. This is because errors often illustrate the thinking processes and known strategies, as well as the phoneme and rime knowledge used by children to carry out the reading tasks.

In the majority of cases, when children made errors during word reading tasks, the error response was a non-response (children either said nothing or stated that they did not recognize the word). In general, the low decoders made many more errors than did the high decoders. During the training portion of the study, the low decoders trained with word families made very few mistakes. Although the data collected on children's errors cannot be quantified statistically, some of the patterns and examples observed are suggestive of various cognitive processes and reading-related skills at work.

Pretests. Students were shown the 16 training words. In most cases, error responses were non-responses. Some errors were blind guesses ("the" for bark). Errors sometimes involved the child giving the first sound of the word (/m/ for mark) or using the first letter to guess a word ("paint" for puck, "mom" for male). Use of partial letter-sounds are examples of phonetic cue reading characterizing students at the partial alphabetic phase. In some cases, children produced words that shared some letters and sounds with the target word ("ballet" for bale, "pull" for pill, "pool" for pale) but it was unclear at times if this was because they were trying to sound the word out and came up with a partial alphabetic reading ("pal" for pale) or because they mistook the target word for a familiar "sight" word ("duck" for dark, "milk" for mill, "date" for dale). High decoders were more likely to give error responses as opposed to non-responses, and gave more responses that involved multiple phonemes present in the words.

Training. The training words lent themselves to some obvious mix-ups: in all groups, at both decoding levels, children confused –ark with –uck, and -ill with –ale. Errors mixing up the first phonemes (i.e. "bill" for pill) were most often found in the pure word families condition, not surprisingly since some children hardly looked at the words at times. Children in the low decoder group, but not the high decoder group, gave responses that could be characterized as random guesses ("can't" for duck, "ghost" for mill), and again, high decoders gave more responses that involved multiple phonemes present in the words ("make" for male, "book" for buck).

Posttest Reading. Incorrect responses followed the pattern observed during training, though there were many more errors to observe among the pure families group. It did not seem to be the case that participants in one treatment group made different types of errors than those in the other groups, though differences in responses between high and low decoders were noticeable.

Some errors were random guesses or guesses using the first letter only ("Joey" for pill, "bore" for buck) or attempts to sound out words (/m/ for mark, "mock" for muck). Other errors were substitutions of other training words. These could be classified into different categories: confusions within the families ("pill" for bill), confusions between training words that shared initial letters ("mill" for male), confusions between training words that did not share either rimes or initial letters ("duck" for pill). Low decoders made more mistakes that involved confusing words within the same families. There were no differences observed between the three treatment groups for these types of errors; however, the untrained control group participants had no errors of these kinds with the exception of one child who made two errors within a family (and perhaps that was due to simply confusing the letter p and b).

Posttest spelling. One question of interest was whether or not children demonstrated evidence of orthographic knowledge through their spelling. Three of the families involved in the study were not strictly phonetic. The rimes –ale, -uck, and –ill required remembering orthographic patterns of English.

The low decoders, but not high decoders, sometimes produced spellings that were random letter strings ("NOYX" for dale), or simply wrote the first letter ("m" for mark), or produced a combination of these two ("BPA" for buck, "MNF" for male). Many spellings produced by low decoders were semi-phonetic ("PK" for park, "BRC" for bark). One error involved adding a random letter to a correctly spelled word ("BiLLK" for bill). There were unusual instances where children did evidence orthographic knowledge, even when they were producing mostly semi-phonetic spellings ("DLL" for dill).

High decoders showed much more awareness of orthographic patterns than low decoders. This was true even for the untrained control group, so it can not be attributed to training, but rather to high decoders' knowledge of the conventional spelling system. Incorrect spellings included patterns such as a final e, double l, and ck ("barck" for bark, "maille" for male, "dall" for dale).

<u>Orthographic complexity of rimes</u>

The degree of difficulty of rime spellings in the training words was explored by examining the numbers of students who identified each item correctly on pre and posttests. It was of interest to see which orthographic patterns were more difficult for beginning readers. These scores are tallied in Table 13. 23 participants identified "duck" by sight on pretests. Five other training words, including "park," were also identified on pretests by a few children. The number of times each word was recognized on posttests was counted. Finally, the proportion of times a word was recognized immediately on posttesting after subtracting the number of times it was identified correctly on pretests was calculated. Results reveal that it was much easier for participants to learn words ending in –ill and –uck than it was to learn words ending in –ark and –ale. As shown in Table 13, students learned about twice as many words (39.23%) from the two easier patterns, -ill and uck, than they did from the–ale and –ark families (averaging 20.12%). Words with the –ale rime containing a silent letter were the most difficult for beginning readers.

Examination of children's spelling scores revealed that although most children could not spell the training words or transfer words correctly on the posttest, they did recall some of the final letters that they saw in words. Some children who wrote only semi-phonetic spellings were able to recall information about the word ending. For example, many children spelled "dill" as "dell" or even "dll." Due to a typo in the script, children were asked to spell "bark" instead of "buck," which meant that instead of two real words and one transfer nonword from each family, there was one extra word in the –ark family and one less in the –uck family. Nevertheless, it is obvious from Table 13 that children had a greater memory, or else more familiarity, with the doubled l than with other patterns. In fact, many children spelled –ale words with final doubled l. The final –le was the most difficult. Children also were able to spell the final letters in –ark with relative success. Many children put rc as the final two letters instead of rk. This may indicate that they were sounding out, rather than remembering the spellings that they saw, to spell the words.

Chapter 5

DISCUSSION

Summary of Findings.

Findings in the present study contribute to our understanding of the role that rimes and word families play in early literacy processes and instruction. The purpose of the study was to determine whether or not teaching children to read words in sets grouped by shared rimes, or "word families," benefits the reading and spelling skills of beginners. This type of training was compared to no instruction and to instruction that involved the same word families but grouped by combining words from different families. No strategies for decoding or word analyses were taught. Children simply practiced reading whole words for several trials.

The effect of training beginning readers to read word families under different training conditions was examined with posttests to assess the following dependent variables: number of words remembered, number of words spelled correctly, number of correct spellings recognized, and number of transfer words read and spelled correctly. It was hypothesized that students would learn words better with mixed lists than with word families, and additionally, that there would there be an advantage for students trained to read mixed lists of words that share the same initial letters. It was of interest to see if the relationship between training and dependent measures would hold true for both lower and higher achieving beginning readers. Finally, pretest variables were examined to see what skills predicted word learning.

Word reading on the pretest and posttest was scored in two ways: number of words recognized immediately, without overt sounding out, and number of immediate words plus words that were overtly sounded out and blended back together correctly. Gain scores from pretest to posttest were calculated.

For words recognized immediately, indicating that they were recognized by sight (memory), results revealed that students trained with mixed lists performed similarly to each other but significantly better than the untrained control group. In contrast, students trained with pure word family lists did not outperform students who received no training at all. Finally, the students trained with mixed lists that shared initial letters significantly outperformed the students trained with pure word families.

For the second measure combining immediately read words with sounded out words, results showed that all training groups performed similarly to each other and significantly better than the untrained control group. Thus, sight word learning was most affected by differences in how students practiced reading the words. Adding decoding skill to the measure had the effect of leveling differences in word reading.

Patterns of performance differed from performance during training. During training, beginning readers, even those who lacked decoding skill, learned to identify words organized by pure word families with relative ease. However, on the posttest several days later, they had difficulty remembering how to read the words they had learned. In contrast, students who were trained with mixed word lists had more difficulty during training, especially students who lacked decoding skill, yet they remembered how to read significantly more words on the posttest than did the students who read words in pure families.

There were no differences in performance among the treatment groups on any of the posttest spelling tasks, but high decoders outperformed low decoders on every measure. Additionally, there were no decoder level by treatment interactions in the posttest analyses. This indicates that the word family manipulations that influenced reading did not also influence spelling. One reason may be that these beginners did not spell words from memory but rather invented spellings by writing the sounds they heard.

Results were analyzed to see if any of the entering skills measured on pretests predicted word learning gains from pretest to posttest. No matter which way word learning was calculated, gains were highly correlated with the following pretest variables: letter sound knowledge, Woodcock Reading Mastery word reading, Woodcock Word Attack (nonword reading), spelling production, Roswell Chall Auditory Blending, and rhyme production.

A regression analysis was conducted to examine which pretest variables predicted word learning. Results indicated that the Woodcock Word Identification test accounted for significant unique variance in word learning. The analysis also revealed that the Roswell Chall Auditory Blending task was a significant predictor of word learning gains among high decoders but not among low decoders.

Rimes used for training were not equally accessible for beginners. In an effort to understand how orthography influences word learning, these differences were explored. Results revealed that it was much easier for participants to learn to read words ending in –ill and –uck than it was to learn words ending in –ark and –ale. An examination of spellings revealed that some participants demonstrated incomplete memory for orthographic patterns, by either confusing letter order (kc instead of ck) or by using orthographic patterns inappropriately (i.e. adding final E to –ill, or doubling l in –ale). Students were more likely to spell the final two letters in –ill words correctly than the other families, with the finaltwo letters in –ale words proving the most difficult. Students did demonstrate relatively high performance on –ark words, perhaps because all the letters corresponded to phonemes. This contrast in difficulty reveals the advantage for reading and spelling when the letters conform to the graphophonemic system. When digraphs such as ck or extra letters such as a final silent E are present, words are much harder for beginners to process and remember.

The differences in word learning between –ale words (the hardest) and –ill words (the easiest) were particularly dramatic. This finding supports the findings of Post et al (2001), who found that beginners had difficulty learning words with rime patterns that included two phoneme blends (i.e. "sk," "nk") and "ke" for /k/ when preceded by a long vowel (as in "cake") than they did learning words which contained sounds represented by digraphs (i.e. "sh" and "ck.") Since beginners generally read words with short vowels, the concept of final E marking a long vowel sound was probably unfamiliar to them.

Explanation of Findings.

Reading. Results of the training portion of the study revealed that students trained with pure word families identified the words nearly perfectly across the 11 trials. In contrast, students trained with mixed lists performed significantly worse, with the lowest scores for the group trained with mixed lists that shared initial letters. When results were analyzed by subgroups of high and low decoders, results revealed that high decoders in all treatment groups read most of the words in their sets during the 11 trials correctly regardless of treatment condition. Low decoders demonstrated a very different pattern: when trained with pure word families, they were able to identify most words correctly, but when trained with mixed lists, they performed significantly worse, with the poorest performance in the group trained with mixed lists of words having the same initial letters. (See Table 6 for means).

One possible explanation for these findings is offered by Ehri's (1991, 1994, 1995; Ehri & McCormick 1998) phase theory. According to Ehri, novice beginning readers who are in the partial alphabetic phase lack adequate knowledge of the letter-sound system and do not yet recognize consolidated units (rimes) in words. They identify words in some cases by processing only some of the alphabetic information available, such as the first letters of the words. Additionally, Goswami's (1992, 1999a) research has shown that young children can identify oral rhymes with relative ease. As such, it is likely that participants in the pure word families group used their partial alphabetic knowledge (identifying initial letters) in conjunction with their knowledge of oral rhymes to read the words during training and they paid little attention to the other letters in the words. This explanation is strengthened by the comments that children made during training. Many children stated that the pure word families lists were "easy" or said something to effect of, "These words all rhyme!" In contrast, only two children trained with mixed lists verbalized observations of the relationship between rimes of words during one training session and the next. Children in the pure families group could read the words when embedded in their families, but the initial letter strategy broke down on the word reading posttest when the four sets were mixed up and initial letters were shared by several words.

In the condition where words in the lists did not share rimes or initial letters, students could utilize an initial letter strategy to decode words, but not an oral rhyming strategy to identify words during training. In contrast, the word lists that contained words from different families but shared initial letters did not allow participants to rely on either strategy, but forced children to make use of multiple phoneme-grapheme connections in the words in order to distinguish among them. During training, students learning mixed lists with and without having the same initial letters did not differ in word learning. However, on the posttest, participants trained with mixed lists that shared initial letters demonstrated superior performance compared to the participants trained with mixed lists that had different initial letters. This means that giving beginners lists of words to learn that share initial letters is useful because it forces them to examine other letters in words besides the initial ones.

The implication of these findings is that, as hypothesized, presenting words grouped by family to children does not present an advantage for word learning when compared to presenting words in mixed lists. Students, even low decoders, who were trained with words grouped in pure families identified them with relative ease during training, yet could not recall them during posttesting a few days later. This means that the students were able to identify the words when they were grouped by rimes, yet did not develop memory for the rimes or improve decoding skills. Low decoders in the pure families group and in the mixed lists with different initial letters were able to rely on initial letters as a word identification strategy during training. The superior performance on the posttests for the groups trained with mixed lists suggests thatrhyming was an important factor that allowed students trained with pure word families to ignore all the letters except the initial letters when reading the words during training. In contrast, the participants trained with mixed lists with different initial letters were not able to identify words only by processing the initial letter, but had to rely on other letters to remind them of the rimes in words. This was evident by the posttest results showing superior word learning for participants when words contained different initial letters and different rimes compared to the pure word families group, where there were different initial letters but the same rimes.

This supports Bruck and Trieman's (1992) finding. Their study demonstrated that although first graders were able to quickly master a list of words that shared rimes, they demonstrated the worst memory for words when compared to groups trained with words that shared only heads or only vowels.

Gilbert et al (1977) found that it takes children longer to master lists of words that are similar to one another because the children confused the words with one another. Similarly, children in the present study needed more trials to master words in the mixed lists conditions compared to the pure families condition. The 11 trials for word training did not allow all the children to master the mixed lists, though it was more than enough trials for children to master the pure families lists.

Theoretically, children in the groups trained with mixed lists had the opportunity to rely on previously learned rimes as the training sessions progressed. However, high decoders in these groups, the ones more likely to be able to retain full spelling of the words in memory and use them as a source for decoding the next set of words, had scores close to ceiling. This left no room to detect improvement from earlier to later sessions. Low decoders trained with mixed lists did not demonstrate improvement as sessions progressed, but they would not be expected to detect common rime spellings among the words since they remember only partial letter cues. This suggests another study to see whether high decoders, if given more words to learn, would benefit from shared rimes distributed across several days of training.

Participants in the group trained with mixed lists and the same initial letters recognized significantly more words immediately (without overt sounding out) on posttests than did the group trained with pure families. This supports the study hypothesis. The likely explanation is that students trained with mixed lists with shared initial letters were forced to process additional grapheme-phoneme connections in the words during training to read the words, and thus, their memory for the words was strengthened.

The finding that memory for words learned in mixed lists was superior to memory for words learned in families contradicts Goswami's theory about young children's awareness of onset-rime units in words and their ability to take advantage of analogies in order to read words. She is correct that words with shared rimes can be identified more easily when grouped by rime families, as happened during training sessions. However, words with a shared orthographic pattern do not become chunked in memory, since these students had not yet reached Ehri's consolidated alphabetic phase.

When examining word learning on the posttest by calculating words identified immediately by sight plus words that were sounded out and blended orally, there were no significant differences for word learning among the three trained groups. Whereas words identified by sight are read from memory, words that are read by sounding out and blending are read by application of decoding skill. The lack of an effect was not surprising because the groups were equivalent in decoding skills before training, and the training methods were not expected to improve general decoding skill. However, training did exert some impact on word learning when decoding was included, as evidenced by the fact that the untrained control group performed significantly worse than did the trained groups. The benefit of training may be explained in a number of ways: experience with the 16 words may have established a closed set in memory that took time to access. Or training may have raised students' comfort level with the experiment.

Juel and Minden-Cupp (2000), in a qualitative study involving four first grade classrooms, found that children in the low reading groups had difficulty seeing chunks in words. In the case of one low reading group, despite a great deal of onset-rime instruction, students could not recognize many rimes to which they had been exposed. In contrast, lower reading group students who were given a great deal of instruction involving the analysis of phonemes within rimes demonstrated an on-grade level ability to decode words, employing segmenting and blending as a frequent strategy. This finding relates to the reading program described by Gaskins et al. (1996-97) in which struggling readers were taught to learn as sight words a set of key words that included high-frequency spelling patterns, such as *–et, -in, -and*, and *–up.* However, although children mastered the sight words containing these patterns, some failed to read new words containing the same known patterns by analogy. For example, the children learned the word *yellow*, but when presented with the word *fellow*, they did not recognize the similarity between the unknown word and the known word. The researchers suspected that this processing failure was due to a lack of careful attention to the letters in the words. Children may have simply remembered only the salient characteristics of the word, such as beginning and ending letters.

One expectation motivating the present study involved a treatment by decoder level interaction for word learning as measured on the posttests. The author's pilot study found that high and low decoders were affected differently by training, whereas no such interaction was detected in the present study. In the pilot study, high decoders who learned words in pure families read the words as well on the posttest as did high decoders who learned mixed lists of words, whereas low decoders showed poorer performance with pure lists than with mixed lists. In contrast, in the present study, both high and low decoders showed the same pattern of word learning favoring mixed over pure lists on the posttest, a pattern identical to that shown by low decoders in the pilot study.

This may be explained by the fact that the pilot study did not involve strict criteria for excluding participants. Some of the high decoders in the pilot study were familiar with the rimes and were able to read many of the training words. In the present study, all the students were screened to ensure that they were not familiar with the training words.

These differences suggest that, in contrast to the high decoders in the pilot study, high decoders in the present study had not quite advanced to the full phase in their word reading. According to Ehri (1991,1994) full phase readers would be expected to retain sight words in memory regardless of whether words were learned in pure or mixed lists. This is because they possess the alphabetic knowledge and decoding skill to fully bond the spellings of words to their pronunciations in memory. As evident in Table 4, high decoders in the present study had limited word attack skill, their knowledge of short vowels was incomplete, and they did not produce phonetic spellings consistently. These deficiencies raise doubt that they possessed the word learning skills characterizing full phase readers. Additional research is needed to determine whether the pattern of word recall favoring mixed over pure lists disappears in students who clearly qualify as full phase readers. Alternatively, it may be that simply proving more word reading practice to the point of overlearning is sufficient to eliminate differences regardless of whether readers are in the partial or full phase (Spring et al, 1977).

Spelling. It was expected that children trained with Pure word families would demonstrate poorer memory for spellings of words and would have a more limited ability to recognize correct spellings than would children trained with mixed lists. However, no differences emerged on posttest spelling tasks. Participants in this study lacked full knowledge of how vowels are represented by graphemes, did not have complete letter-sound knowledge and were unfamiliar with orthographic patterns. Students' decoding level influenced performance, with high decoders significantly outperforming low decoders on both spelling production and recognition tasks. However, no differences as a function of training groups emerged.

Spelling generation tasks were too difficult for many participants who lacked full knowledge of the grapheme-phoneme system, particularly knowledge of short vowels. Many participants did not come up with any letters for some words, or used random letter strings, or were unable to get further than the initial letter of words. Some participants created semi-phonetic spellings, often omitting vowels, used the wrong vowels, or provided spellings that were phonetic, but incorrect. Poor decoders, especially, were usually unable to create any correct or even phonetic spellings. High decoders also made many errors in spelling by confusing the orthographic patterns from the different words they learned (i.e. “darck” for dark, “dille” for dill.) The presence of these types of errors contributed to very few spellings that were completely correct. This was why the number of correct letters was calculated as an alternative measure, but even those scores indicate that participants had great difficulty with the task. Participants were beginning readers, who typically are poor spellers. Although participants in the trained groups were exposed to the words, they could not produce spellings that were any better than students who were not exposed tot he words. This indicates that these beginners lacked long-term memory for word-specific information about the words they learned to read. From Table 8, it is apparent that students remembered few if any words, with most students scoring zero. In short, students in all the groups were inventing spellings rather than accessing remembered spellings.

Spelling recognition tasks that required students to select the correct spelling from a group of three words proved to be too unfamiliar or difficult to allow many of the children to be successful. Based on the results of this study, it is advised that future studies not include this type of multiple choice task with young children. Many of the children, particularly the low decoders, made blind guesses (i.e. circling the first choice for every item.) Unfortunately, using multiple choice tasks that are simpler (i.e. two choices) only raise the possibility that students will choose the correct answer based on chance.

Predictors of Word Learning.

As would be expected based on findings of many other studies (i.e. Muter, Hulme, Snowling & Taylor, 1998, Byrne, 1998), word learning in this study was highly correlated with students' scores on phonemic awareness tasks, spelling production, auditory blending, non-word reading, and rhyme production. (See Tables 10 and 11). There were stronger correlations between the students' pretest scores on the variables mentioned above and their word learning as measured by words recognized immediately than there were between the pretest variables and word learning as measured by words recognized immediately plus those sounded out. This can be interpreted to suggest that students with better reading-related skills are able to secure more words in memory.

Examining the number of words recognized immediately is an indication of words secured in memory. If words decoded are added to these scores, other underlying factors besides word learning then need to be considered. Examining how many words a child successfully decodes reflects not only word learning but also other factors, such as a child's motivation to struggle with unknown words and his or her ability to apply strategies.

Word learning was predicted by scores on the Woodcock Reading Mastery Word Identification test. This is consistent with the findings of O'Shaugnessy and Swanson (2000). Their study involved 45 disabled readers trained to analyze words at the phoneme or rime level. They found that the most significant predictor of growth in oral reading fluency was the participants' initial level of word identification skill. Logic would suggest that success on a word learning task should be predicted by previous successful word learning.

In addition, word learning for high decoders was predicted by the students' Roswell Chall Auditory Blending scores. This task measured phonemic awareness, something numerous studies have identified as a contributor in learning to read and in building sight memory for words. The usefulness for this measure as a better predictor for high decoders than for low decoders is explained by the fact that low decoders scored poorly on this task (mean 9.86 out of a maximum score of 30, with a standard deviation of 8.5). Thus, individual variation among the scores this low made little difference in low decoders' ability to learn words.

The findings of this regression analysis indicate that three pretest variables relating to language and literacy skills were useful in predicting word learning during this study. The Woodcock Reading Mastery Word Identification test, indicating participants' entering word reading ability, was a significant predictor. This is further evidence for the "Mathew effects" in growth in decoding fluency, as described by Stanovich (1986) and others. The more words the participants already knew when entering the study, the more words they were able to learn over the course of the study. Additionally, students who were high decoders (defined by having demonstrated evidence of decoding skill as measured by the Woodcock Reading Mastery Word Attack), showed greater gains in word learning than did low decoders, who were unable to sound out any nonwords.

Finally, for beginning readers who were high decoders, the Roswell Chall Auditory Blending task was a useful predictor of word learning. This finding supports numerous studies indicating that phonemic awareness, the ability to manipulate phonemes in spoken words, is an underlying component of the development of decoding skill.

The relationship between rhyme and reading.

Learning training words in this study was not predicted by rhyming in the regression analysis, although generating rhymes was moderately correlated with the number of words learned. This contradicts Goswami's (1995, p.135) argument that rhyme awareness is predictive of reading ability. Perhaps this finding can be used as further evidence in support of proposing a fourth hypothetical model connecting rhyme, phoneme detection, and reading and spelling in addition to the three described by Goswami and Bryant (1990). A fourth model could position phonemic awareness as an underlying skill that leads to both rhyme and reading and spelling skills. This fourth model would be supported by Wood's (1990) findings, as well as by numerous studies in the literature which examine predictors of success in reading and reading-related abilities. This explanation is further supported by the pattern of findings in the present study. High decoders, who displayed strengths in phonemic awareness tasks, had good rhyme-generation skills, whereas low decoders, who were weak in phonemic awareness, did not have good rhyme-generation skills.

The participants had difficulty with the rhyming tasks used in this study. The first task required students to choose a non-rhyming word out of a set of three rhyming and one non-rhyming word (i.e. bee, see, tee, car). Students heard the words three times in a row. Most students had great difficulty choosing the non-rhyming word. The overall mean for all students in the study was at about chance level. Perhaps their short-term memories could not hold the words long enough for them to consider and select the correct choices. Or perhaps they were attending to the meanings of the words. Also, 3 of the 12 items were extremely difficult because they included two-syllable words. (See Appendix D for item list.) There were significant differences between high and low decoders on the second rhyming task, which asked children to produce a rhyme to a given word. The rhyme production task proved difficult for some participants. Perhaps the open-ended format created uncertainty because participants were unable to determine if they were coming up with the right answers in the absence of feedback.

A research review by Macmillan (2002), who analyzed 61 studies involving the role of rhyming as a predictor of reading, raises some troubling questions about testing rhyme awareness. MacMillan points out that many oddity tasks involving rhymes actually are testing phonemic awareness, since children may be selecting the correct rhyme because they notice a different phoneme (i.e. choosing *sit* from *hat, cat, bat, sit* or selecting *had* from *bat, cat, hat, sad*.), creating a confounded rhyme measure (p.11).

Furthermore, she explains that

> …there may be at least four other contaminants operating within these tests. These include: the child's working memory (Oakhill & Kyle, 2000), speech perception (Snowling et al., 1994), and socio-economic background (Burt et al., 1999), as well as the examiner's quality, volume, or pace of speech production while articulating the three or four options per item (McDougall et al., 1994)
> p. 11

Macmillan's study found that rhyme oddity tasks in the studies she evaluated were highly correlated with performance on phonemic awareness tasks, with correlations as high as .67. In contrast, correlations between rhyme generation tests and phoneme tests were no higher than .35 (p.11). This explains why the rhyme generation task in this study may have been significantly and moderately correlated with word learning, while the rhyme oddity task was not. (See Tables 10 and 11 for correlations.)

The spelling production task scores were more highly correlated with word learning than the rhyme production task, and furthermore, the spelling production task was much more highly correlated with every other measure of reading knowledge, such as Woodcock Reading Mastery word reading and letter-sound scores (see Tables 10 and 11 for a comparison). Finally, although many critics of studies involving testing rhyme awareness have argued that recognition and production of rhymes is largely a function of age (i.e. MacMillan 2002), age did not correlate significantly with either rhyming measure used in this study.

Implications.

Word reading performance on lists of words organized by family maximizes poor decoders' abilities to correctly identify words. However, the problem is that poor decoders can simply process the initial letter and join it with the oral rhyme to identify each word. In contrast, word learning exercises involving mixed lists minimizes poor decoders' abilities to identify words and forces them to examine more of the letters to remember how to read the words. When teachers use word families as part of early literacy instruction, they limit their ability to observe poor decoders' difficulty with word reading, since the poor decoders can identify words in word families nearly as well as high decoders. (See Figure 2.) In contrast, using mixed lists maximizes teachers' chances to observe the poor decoders' difficulties with word reading. This carries an important implication for instruction. The findings presented here support the value of teaching students to analyze all the grapheme-phoneme correspondences in words. Other studies have shown that segmentation and blending training involving words with shared rimes can be a successful method of instruction (i.e. Post et al 2001). This study demonstrates that it is best to avoid grouping words into "families" if these type of decoding strategies are not part of reading instruction.

Limitations.

The selection of words for word sets was constrained by the need to create lists that allowed the same words to be grouped by word families as well as by shared initial letters. These constraints on the generation of word lists, as well as the young age of the participants, resulted in word lists containing only four words in each set. In typical classrooms, children are presented with more than four words per family. It is possible that teaching more words would have allowed the rime unit to become solidified in memory which would have allowed children to process other words with the same rime unit on transfer tasks. Higher decoders should be able to develop memory for these rime units, so they should be the ones to exhibit transfer.

Inclusion of “duck” as a training word was particularly problematic, because many children identified the word correctly during pretesting, either recognizing it instantly or through sounding out and blending, though the –uck rime itself was not identified easily by the participants. “Park” was also a word that kindergarten and first grade children are likely to have seen. Although there were no ceiling effects for word learning on the posttest, having a student identify “duck” correctly meant that they did not have the opportunity to learn it during the study, and of all the words in the –uck family (buck, puck, muck, duck), duck was clearly the one vocabulary word that was most accessible for young children.

Based on the word learning outcomes for each item, it is clear that including a difficult rime, -ale, limited participants’ word learning. It is likely that inclusion of another family with a short rather than a long vowel would have been more productive. (See Table 13 for comparisons between words and rimes).

Additionally, using multiple families that shared letters in the rimes (i.e. –ale and –ill; –uck and –ark.) produced a great deal of confusion. This resulted in many reading errors during training and posttesting. Families with less similarity would have reduced confusion among words and discouraged guessing based on partial recognition.

Another limitation was that not all participants in the mixed lists groups were provided with enough trials to master the words during training sessions. It was decided that participants would practice the words for 11 trials. The repetitive nature of the task demanded that the number of trials be limited to some reasonable number in order not to lose young children’s attention. This number was selected because in the pilot study, all the children mastered the word lists in less than 11 trials. However, children in the mixed list groups, particularly the low decoders, were not all able to master the lists during training sessions. Because the words were similar, training to the point of overlearning should have been provided, as explained by Spring et al (1977).

It should be noted that overlearning per se does not guarantee memory for the words. Participants trained with pure word families mastered words much faster and thus engage in many more overlearning trials. Despite this, they did not remember the words well as students trained with mixed lists. If the participants in the mixed lists groups had all had the opportunity to practice the words until they mastered them, it is likely that even stronger differences between the groups would have been observed. The advantage of overlearning for the students trained with mixed lists that shared initial letters might have been especially great. Because the mixed lists with the same initial letters were the most difficult to master, many students did not reach criterion, and very few had the opportunity to engage in overlearning. In contrast, all students in the pure families group, even the low decoders, engaged in overlearning.

Spelling generation tasks were too difficult for many participants, some of whom produced random letter strings, gave no responses, or could not get beyond identifying the initial letters of words. The spelling tasks on the posttests were not useful for distinguishing among the training groups. Results showed that high decoders outperformed low decoders, but there were no differences between the treatment groups and the no-treatment control group that did not even see the words. The spelling recognition tasks and the rhyme oddity task were difficult because multiple choice tasks are not usually presented to young children. These tasks were so difficult that many children performed at chance level, indicating that young children do not appear to possess a strategy for examining the alternative choices to pick the best one. Brown (1984) has found, in studies with adults, that showing attractive misspellings of words in a multiple choice task erodes their memory for the correct spelling Perhaps this happened with these beginning readers as well on the first spelling recognition task.

Directions for Future Research.

This study has shown that simply teaching students words in families will not allow them to spontaneously learn rime units. All three treatment groups practiced the same words. Students saw each word in isolation 13 times before the posttest session, and saw over 50 words containing each of the four rimes. If they did not recognize the word during training, they were told the word while it was displayed in front of them. However, although the trained participants all had many opportunities to read the words, participants trained with mixed lists outperformed students trained with pure word families.

Although low and high decoders in this study were provided with the same amount of practice in reading the training words, the ones who began the study with more advanced skills gained more from training. The high decoders were able to read, on average, about 39% of the words immediately on posttests, while low decoders recognized about 10%. When examining the scores of only the trained groups, high decoders performed about three times as well as low decoders: for training words recognized immediately on posttests, high decoders recalled about 46% of the words, while low decoders recognized about 14%; for words recognized immediately plus those sounded out, high decoders identified about 50% correctly on posttests while low decoders recognized about 15%. This finding lends evidence to the observed “Mathew effects” in reading (Juel, Griffith, & Gough, 1986; Stanovich, 1986). The initial level of word identification skill (as measured by Woodcock Reading Mastery word reading) predicted word learning. This certainly supports the idea that it is necessary for poorer readers to have more instructional time on reading and writing activities than better readers do if we are to see the same proportional gains in reading achievement. Future research is needed to explore whether or not providing more time for literacy instruction is beneficial. Beginning readers should benefit from instructional time devoted to building letter identification skills, letter-sound correspondences, and phonemic segmentation and blending skills. If struggling beginning readers could have these skills strengthened, they too could benefit from classroom reading instruction.

Although there was no statistically significant interaction, it is interesting to note training results for high and low decoders. High decoders demonstrated small benefits from training involving mixed lists (about 47% of training words were recognized immediately on posttests for those trained with mixed lists versus about 29% for those trained with word families). Low decoders trained with mixed lists that shared initial letters recognized more than twice as many words immediately on posttests than did low decoders trained with word families (about 20% of training words were recognized immediately on posttests for those trained with mixed lists that shared initial letters versus about 8% for those trained with word families.) These findings suggest that it was more critical for low decoders to be trained with mixed lists than it was for high decoders. Future research should explore the reasons underlying these types of differences in order to strengthen the arguments presented here.

Several questions have emerged as a result of this research study. It is important to ascertain not just how children process lists of words practiced in isolation from literacy instruction, but also to understand the inter-relationships between spontaneous connection-making and responses triggered through direct instruction and teaching of decoding strategies. More qualitative studies of classroom instruction are needed in order to observe how word families are used as part of instruction. For example, teaching students how to spell words, and giving instruction about rimes and phoneme-grapheme relationships, might lead to different results depending upon how words lists are grouped. Finally, discovering a better method for testing rhyme awareness and spelling recognition would benefit future studies.

Conclusion.

This study provides data suggesting that two factors contribute to children's skill in remembering how to read words they practice. One is the initial decoding level of beginning readers. Students who possessed word attack skills, greater phonemic awareness (as measured by more letter-sound knowledge, blending skills, and greater ability to produce spellings that were correct or at least phonetic), and a larger sight word vocabulary were able to learn significantly more words during training than were students who lacked or lagged in those critical skills. Additionally, presenting words in mixed lists resulted in greater memory for word learning when compared to simply providing students with word family lists.

Table 1. Training Words from Pilot Study.

Pure Word Families	List 1. LiGHT NiGHT TiGHT RiGHT	List 2. HUMP BUMP LUMP JUMP	List 3. BARK PARK DARK MARK	List 4. NEED SEED WEED DEED
Mixed-Order Lists	List 1. LiGHT HUMP BARK NEED	List 2. NiGHT BUMP PARK SEED	List 3. TiGHT LUMP DARK WEED	List 4. RiGHT JUMP MARK DEED

Table 2: Mean Number of Training Words Read Correctly on Pilot Study Posttests for the Treatment Groups and Control Group

Maximum Score = 16

	Study Groups					
	Word Learning Conditions					
	Pure Family		Mixed, Dif. Initial		Control Group	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
High Decoders	9.83	2.0	9.00	3.3	2.25	2.8
Low Decoders	1.50	0.5	5.66	3.0	0.00	0.0
General Mean	5.66	4.6	7.33	4.8	1.13	1.8

Note: There were 12 participants in each treatment group, and 8 students in the control group, for a total of 32 participants. Of these, 16 were high decoders and 16 were low decoders, with an equal number of high and low decoders in each learning condition.

Table 3. Training Words

Treatment Group: *Pure Word Families*

Training Session:	1	2	3	4
Words:	BiLL	BUCK	BARK	BALE
	DiLL	DUCK	DARK	DALE
	MiLL	MUCK	MARK	MALE
	PiLL	PUCK	PARK	PALE

Treatment Group: *Mixed Lists, Different Initial Letters*

Training Session:	1	2	3	4
Words:	BiLL	DiLL	MiLL	PiLL
	DUCK	BUCK	PUCK	MUCK
	MARK	PARK	BARK	DARK
	PALE	MALE	DALE	BALE

Treatment Group: *Mixed Lists, Same Initial Letters*

Training Session:	1	2	3	4
Words:	BiLL	DiLL	MiLL	PiLL
	BUCK	DUCK	MUCK	PUCK
	BARK	DARK	MARK	PARK
	BALE	DALE	MALE	PALE

Table 4

Mean Performance of Low and High Decoders on Pretest Measures and Test Statistics.

	Low Decoders		*High Decoders*		
Measures	Mean	S.D.	Mean	S.D.	F Statistic
Group Characteristics					
Gender (males: females)	15:20		19:19		
Language (number with two)	28		27		
Age (in months)	68.60	6.54	70.79	6.57	2.03 n.s.
Letter Knowledge					
Names (max 26)	24.94	2.15	25.89	.31	7.26 **
Sounds (max 26)	18.37	5.60	22.60	2.52	17.79 **
Short Vowels (max 5)	2.26	1.15	3.18	1.22	11.07 **
Phonological Awareness					
Rhyme oddity task (max 12)	3.72	2.27	4.66	2.46	2.88 n.s.
Rhyme production (max 10)	3.69	3.39	6.45	3.29	12.47 **
Total Rhyming (max 22)	7.40	4.87	11.12	4.60	11.17 **
Blending (max 30)	9.86	8.50	20.39	7.21	32.80 **

Table 4 (continued).

	Low Decoders		*High Decoders*		
Measures	Mean	S.D.	Mean	S.D.	F Statistic
Word Reading					
Preprimer (max 24)	7.06	5.08	16.0	4.32	65.98 **
Woodcock Words (max 46)	5.49	6.25	20.45	9.24	64.54 **
Range (0-39)					
Grade Equivalent	1.0		1.5		
Word Attack (max 30)	.00	.00	4.26	3.07	67.29 **
(Range 0 –13)					
Grade Equivalent	K.5		1.2		
Training Words					
Immediate (max 16)	.23	.49	.82	.90	11.78 **
Immediate plus decoded (max 16)	.28	.56	2.34	2.77	18.59 **
Spelling					
Points Correct (max 60)	18.74	12.21	33.34	6.70	40.97 **

Note: There were 35 students in the low decoder group and 38 students in the high decoder group. **$p < .05$; n.s. means F value is not statistically significant at $p < .05$.

Table 5: Mean Performance of the Word Family Treatment Groups and the Control Group on the Pretest Measures

	Word Learning Conditions					
	Pure Family (PF)		Mixed, Diff. Initial (MD)		Mixed Same Initial (MS)	
Measures	Mean	S.D.	Mean	S.D.	Mean	S.D.
Group Characteristics						
Gender (males:females)	9:9		10:8		8: 10	
Decoders (high:low)	10: 8		10:8		10: 8	
Language (number with two)	13		16		13	
Age (in months)	68.61	6.22	69.72	5.20	69.22	6.53
Letter Knowledge						
Names (max 26)	25.55	.50	25.61	1.24	25.22	2.60
Sounds (max 26)	21.00	3.18	21.16	3.82	21.05	5.57
Short Vowels (max 5)	2.44	1.14	2.89	1.32	2.94	1.31
Phonological Awareness						
Rhyme oddity task (max 12)	4.50	2.04	4.05	2.31	4.22	2.46
Rhyme production (max 10)	5.61	3.12	6.00	3.77	4.50	3.72
Total Rhyming (max 22)	10.11	3.97	10.05	5.42	8.72	5.26
Blending (max 30)	14.94	7.89	16.33	9.08	16.33	10.14

Table 5 (continued).

	Word Learning Conditions					
	Pure Family (PF)		Mixed, Diff. Initial (MD)		Mixed Same Initial (MS)	
Measures	Mean	S.D.	Mean	S.D.	Mean	S.D.
Word Reading						
Preprimer (max 24)	11.33	5.75	11.39	6.11	13.06	7.89
Woodcock Words (max 46)	14.22	10.77	13.50	11.24	14.77	12.65
Word Attack (max 30)	1.78	2.10	2.16	2.57	2.77	3.96
Training Words						
Immediate (max 16)	.44	.71	.56	.85	.67	.91
Immediate plus decoded (max 16)	.68	.840	1.66	2.74	1.66	2.78
Spelling						
Points Correct (max 60)	25.44	9.79	28.44	12.51	27.22	11.68

Table 5 (continued)

Measures	No Treatment Control Mean	S.D.	F Statistic
Group Characteristics			
Gender (males: females)	7:12		
Decoders (high:low)	8:11		
Language (number with two)	13		
Age (in months)	71.31	8.23	.56 n.s.
Letter Knowledge			
Names (max 26)	25.37	1.10	.22 n.s.
Sounds (max 26)	19.16	5.92	.76 n.s.
Short Vowels (max 5)	2.68	1.34	.57 n.s.
Phonological Awareness			
Rhyme oddity task (max 12)	4.05	2.88	.13 n.s.
Rhyme production (max 10)	4.42	3.76	.87 n.s.
Total Rhyming (max 22)	8.47	5.59	.53 n.s.
Blending (max 30)	13.84	10.84	.30 n.s.

Table 5 (continued)

Measures	No Treatment Control Mean	S.D.	F Statistic
Word Reading			
Preprimer (max 24)	11.11	6.35	.34 n.s.
Woodcock Words (max 46)	10.73	9.25	.50 n.s.
Word Attack (max 30)	2.16	3.46	.32 n.s.
Training Words Immediate (max 16)	.47	.69	.28 n.s.
Immediate plus decoded (max 16)	1.32	2.22	.77 n.s.
Spelling			
Points Correct (max 60)	24.36	14.58	.41 n.s.

Note: There were 18 students in each of the treatment groups and 19 students in the control group. **$p < .05$; n.s. means F value is not statistically significant at $< .05$.

Table 6: Mean Number of Training Words Read Correctly During 11 Trials for the Treatment Groups

Maximum Score = 176

Word Learning Conditions

	Pure Family		Mixed, Dif. Initial		Mixed, Same Initial		All Trained Participants	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
High Decoders	169.40	7.68	156.10	9.04	151.00	9.04	158.83	15.71
Low Decoders	160.63	15.35	96.75	33.56	81.00	32.62	112.79	44.44
General Mean	165.50	12.02	129.72	37.79	119.89	44.31	138.37	39.06

Note: There were 18 participants in each treatment group, for a total of 54 participants. Of these, 30 were high decoders and 24 were low decoders.

Table 7.

ANOVA Test Statistics for Reading Words During Training as a Function of Decoding Level and Treatment Session.

Main Effects & Interactions	Df	Sum of squares	Mean Square	F Value
Between Subjects				
Treatment	2	5925.78	2962.89	25.47**
Decoding Level	1	7066.12	7066.12	60.74**
Treatment x Level	2	2377.69	1188.84	10.22**
Error (Treatment)	48	5583.67	116.33	
Within Subjects				
Session	3	38.79	12.93	.84 n.s.
Session x Level	3	93.80	31.27	2.04 n.s.
Error (Session)	144	2205.51	15.32	
Treatment X Session	6	89.79	14.10	.93 n.s.
Treatment x Session x Level	6	72.10	12.17	.79 n.s.
Error (Treatment x Session)	144	2205.51	15.32	

Note: There were 18 participants in each treatment group, for a total of 54 participants. Of these, 30 were high decoders and 24 were low decoders.
** $p < .01$; * $p < .05$; n.s. means F value is not statistically significant at $p < .05$.

Table 8: Mean Performance of the Word Family Treatment Groups and the Control Group on the Posttest Measures.

	Word Learning Conditions					
	Pure Family (PF)		Mixed, Diff. Initial (MD)		Mixed Same Initial (MS)	
Measures	Mean	S.D.	Mean	S.D.	Mean	S.D.
Word Reading Gain* (immediate) (16 max)	3.00	3.55	4.94	3.72	5.61	4.64
Word reading Gain* (immediate or sounded out) (16 max)	3.28	3.69	4.28	3.64	5.00	3.99
Spelling Production Number of Words Correct (16 max)	.88	1.13	1.67	2.03	1.00	1.37
Spelling Production: Number of Letters Correct (64 max)	24.22	8.23	25.94	10.47	24.22	10.18
Spelling Recognition Task One (24 max)	9.78	4.36	10.50	5.65	11.56	6.14
Spelling Recognition Task Two (8 max)	5.72	1.64	5.78	1.83	5.67	1.88

*Gain scores were calculated by subtracting correct responses on pretests from the posttest scores, because a few participants gave correct responses on one or two words on pretests.

Table 8 (continued)

Measures	No Treatment Control (NT) Mean	S.D.	F Value	Results Tukey Test
Word Reading Gain (immediate) (16 max)	53	1.21	9.67**	MD, MS > NT MS > PF PF> NT MD = MS
Word Reading Gain (immediate or sounded out) (16 max)	.47	1.02	7.34**	PF, MD, MS > NT PF = MD = MS
Spelling Production: Number of Words Correct (16 max)	.68	1.29	1.39 n.s.	
Number of Letters Correct (64 max)	22.42	10.79	.23 n.s.	
Spelling Recognition Task One (24 max)	9.95	5.65	.36 n.s.	
Spelling Recognition Task Two (8 max)	5.68	1.77	.14 n.s.	

Note: There were 18 students in each of the treatment groups and 19 students in the control group.
** $p < .05$; n.s. means F value is not statistically significant at $p < .05$.

Table 9: Mean Performance of Low and High Decoders on Posttest Measures

	Low Decoders		*High Decoders*		
Measures	Mean	S.D.	Mean	S.D.	F Statistic
Word Gain* (immediate) (16 max)	1.40	2.52	5.40	4.11	24.53**
Word Reading Gain * (immediate or sounded out) (16 max)	1.51	2.58	4.79	3.84	17.10**
Spelling Production	.17	.45	1.87	1.68	33.47**
Number of Words (12 max)	.17	.45	1.58	1.31	36.48**
Training Words (8 max)					
Nonwords (4 max)	.00	.00	.37	.75	8.42**
Spelling Production: Number of Letters (48 max)	17.09	8.99	30.71	4.82	66.58**
Spelling Recognition Task One (24 max)	7.51	3.34	13.13	5.61	26.49**
Training Words (16 max)	5.34	2.33	8.92	4.28	19.24**
Transfer Words (8 max)	2.63	1.48	4.21	1.82	16.48**

* Gain scores were calculated by subtracting correct responses on pretests from the posttest scores, because a few participants gave correct responses on one or two words on pretests.

Table 9 (continued)

	Low Decoders		*High Decoders*		
Measures	Mean	S.D.	Mean	S.D.	F Statistic
Spelling Recognition Task Two (8 max)	4.69	1.71	6.66	1.15	33.97**
Selecting from Family (4 max)	2.74	1.20	3.74	.55	21.28**
Selection from Shared Initial Letter (4 max)	1.94	1.08	2.92	.85	18.57**

Note: There were 35 students in the low decoder group and 38 students in the high decoder group.
**$p < .05$; n.s. means F value is not statistically significant at $p < .05$.

Table 10. Pearson Correlations Between Pretest Measures and Word Learning for All Participants (N = 73)

	1	2	3	4	5	6	7	8	9	10
1.Letter Sounds	-	.329**	.570**	.509**	.413**	.134	.236*	.746**	.417**	.362**
2. Vowels		-	.455**	.269*	.506**	.168	.354**	.398**	.196	.144
3. Roswell Blending			-	.483**	.602**	.384**	.474**	.742**	.409**	.369**
4. Woodcock				-	.663**	.149	.280*	.648**	.675**	.543**
5. Woodcock Nonword					-	.120	.404**	.580**	.503**	.377**
6. Rhyme Oddity						-	.397**	.185	.207	.208
7. Rhyme Production							-	.470**	.350**	.302**
8. Spelling Production								-	.487**	.404**
9. Gain in Training Words (Immediate)									-	.928**
10. Gain in Training Words (Immediate + Decoded)										-

Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table 10 (continued)

	Mean	Standard Deviation
1. Letter Sounds (max 26)	20.58	4.8
2. Vowels (5 max)	2.74	1.3
3. Roswell Chall Blending (max 30)	15.34	9.4
4. Woodcock (max 46)	13.27	10.9
5. Woodcock Nonword (max 30)	2.22	3.1
6. Rhyme Oddity (max 12)	4.21	2.4
7. Rhyme Production (max 10)	5.12	3.6
8. Spelling Production (max 60)	26.34	12.1
9. Gain in Training Words, Immediate (max 16)	3.48	3.1
10. Gain in Training Words, Immediate + decoded (max 16)	3.22	3.5

Table 11. Pearson Correlations Between Pretest Measures and Word Learning Participants in Treatment Groups (N = 54)

	1	2	3	4	5	6	7	8	9	10
1.Letter Sounds	-	.308*	.570**	.509**	.413**	.134	.236*	.746**	.417**	.362**
2. Vowels		-	.455**	.269*	.506**	.168	.354**	.398**	.196	.144
3. Roswell Blending			-	.483**	.602**	.384**	.474**	.742**	.409**	.369**
4. Woodcock				-	.663**	.149	.280*	.648**	.675**	.543**
5. Woodcock Nonword					-	.120	.404**	.580**	.503**	.377**
6. Rhyme Oddity						-	.397**	.185	.207	.208
7. Rhyme Production							-	.470**	.350**	.302**
8. Spelling Production								-	.487**	.404**
9. Gain in Training Words (Immediate)									-	.928**
10. Gain in Training Words (Immediate + Decoded)										-

Correlation is significant at the .05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

Table 11 (continued)

	Mean	Standard Deviation
1. Letter Sounds (max 26)	21.07	4.2
2. Vowels (5 max)	2.76	1.3
3. Roswell Chall Blending (max 30)	15.87	8.9
4. Woodcock (max 46)	14.17	11.4
5. Woodcock Nonword (max 30)	2.24	3.0
6. Rhyme Oddity (max 12)	4.26	2.2
7. Rhyme Production (max 10)	5.37	3.5
8. Spelling Production (max 60)	27.03	11.2
9. Gain in Training Words, Immediate (max 16)	4.52	4.1
10. Gain in Training Words, Immediate + decoded (max 16)	4.19	3.8

Table 12.

Summary of Regression Analysis for Pretest Variables Predicting Word Learning (N = 54) Gain (Pretest to Posttest) in Words Recognized Immediately

	Initial Model Unstandardized Coefficients			Final Model Unstandardized Coefficients		
Variable[1]	B	SE B	Sig.	B	SE B	Sig.
(Constant)	2.07	1.361	.136			
Decoder Level	1.863	1.558	.239	.183	1.214	.881
Letter Sounds	.314	1.77	.084			
Vowels	-.085	.510	.869			
Woodcock	-.025	.133	.852	.224	.055	.000
Roswell Blending	-.184	.097	.064	-.039	.068	.570
Rhyme Oddity	.573	.329	.089			
Rhyme Production	-.085	.212	.691			
Spelling Production	.058	.096	.547			
Letter Sounds * DL	-.262	.303	.392			
Vowels * DL	-.451	.675	.508			
Woodcock * DL	.222	.152	.154			
Ros. Blending * DL	.292	.132	.032	.205	.099	.044
Rhyme Oddity * DL	-.454	.429	.296			
Rhyme Prod. * DL	.148	.276	.595			
Spelling Prod. * DL	.022	.147	.879			

Note: R^2 for Initial Model = .677 R^2 for Final Model = .575

[1] All pretest variables were centered for this analysis.

Table 13. Children's Reading and Spelling of Rimes.

Reading

Words	Recognized Immediately Pretest	Recognized Immediately Posttest	% of Items Learned Between Pretest and Posttest
Bale	0	9	
Dale	1	9	
Male	1	14	
Pale	0	4	
Total -ale	2	36	13.95%
Bark	0	15	
Dark	0	12	
Mark	0	16	
Park	5	20	
Total -ark	5	63	26.92%
Bill	5	26	
Dill	0	25	
Mill	2	16	
Pill	0	24	
Total -ill	7	91	43.75%
Buck	0	17	
Duck	23	43	
Muck	0	12	
Puck	0	13	
Total -uck	23	82	35.19%

Table 13 (continued).

Spelling Words	Final two letters correct (max 73)	% of Participants who Spelled Final Letters Correctly
Male	3	4%
Dale	1	1%
Lale	3	4%
Bark	12	16%
Park	14	19%
Mark	3	4%
Tark	12	16%
Dill	33	45%
Bill	31	42%
Rill	9	12%
Puck	6	8%
Juck	2	3%

Figure 1. Mean Performance Per Session as a Function of Decoder Level and Treatment Level

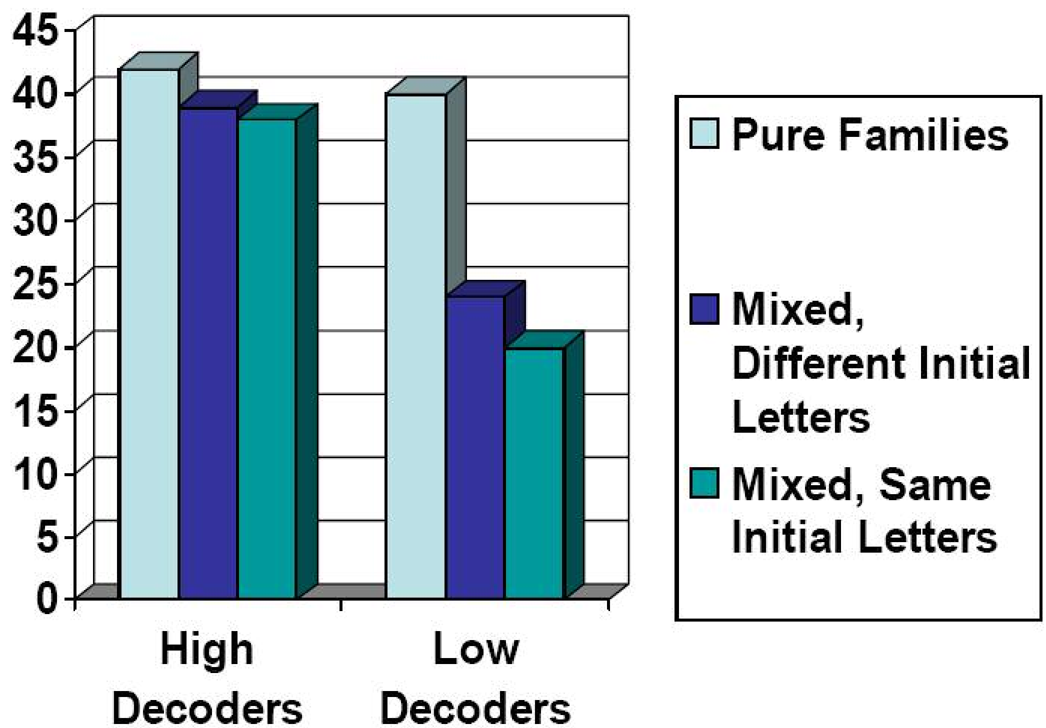
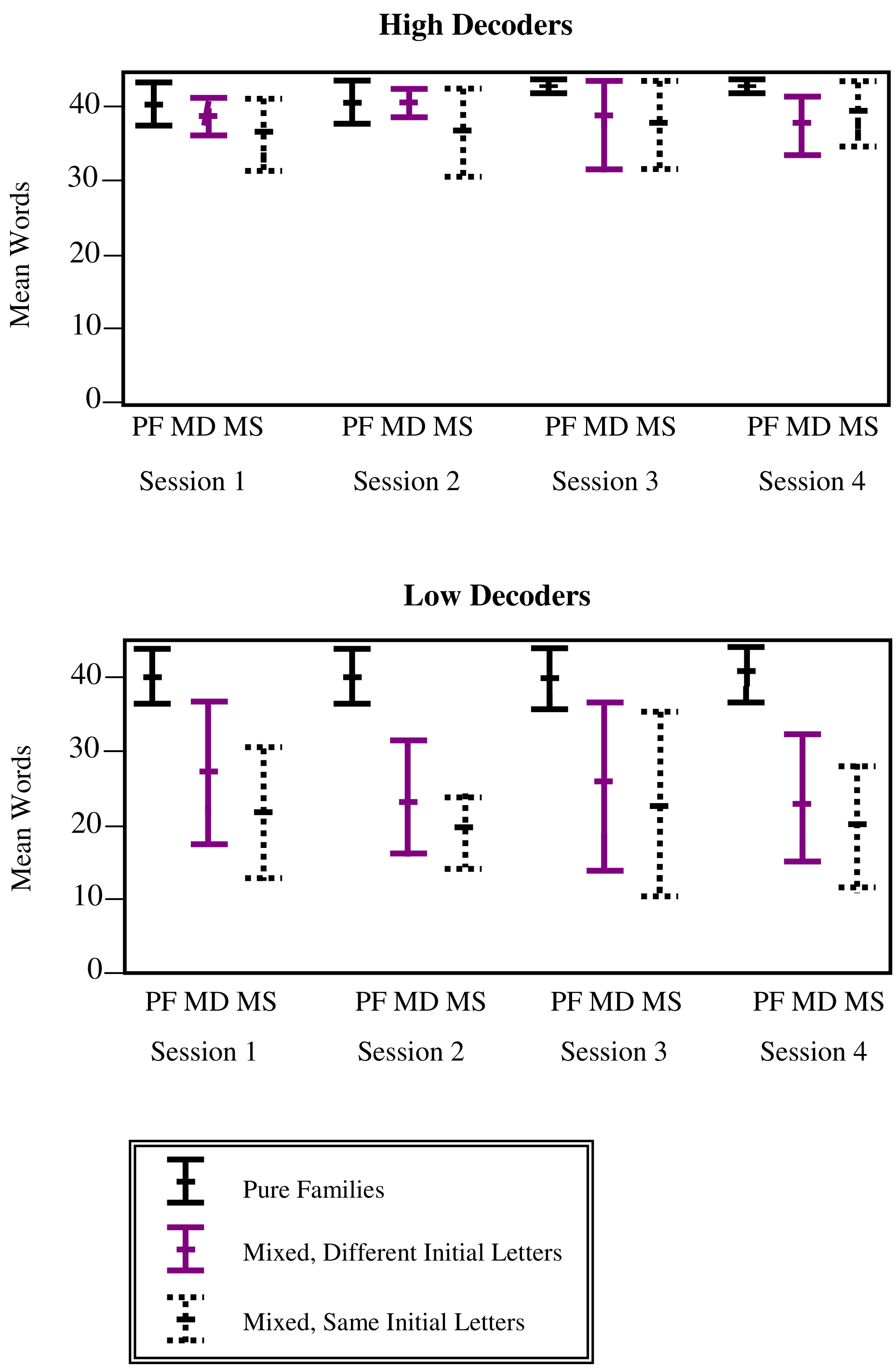


Figure 2.

Mean Number of Words Read Correctly by High and Low Decoders During Training Sessions as a Function of Treatment Group.

Vertical lines depict standard deviations from the means.



Appendixes.

Appendix A.

<u>Alphabet Presented to Participants.</u>

C	R	D	N	J	B	E	H	Q
G	Y	W	S	L	F	U	V	X
A	I	Z	K	P	T	O	M	

Appendix B.

<u>Vowel Task</u>.

Correct responses for each item are underlined.

Practice item: M <u>K</u> R

1. iB AB UB

2. OJ AJ iJ

3. UG AG iG

4. AP iP EP

5. UV EV AV

Appendix C.

Preprimer Words.

a, and, go, you, it, up, help, me, no, in, do, the, play, who, come, can, see, stop, said, ride, down, blue, little, with, green.

Appendix D. Rhyming Tasks

Task 1. Rhyme Oddity.

Script: We are going to play a game with words that rhyme. Let me tell you about rhyming words. They are words that sound alike. I am going to tell you four different words. Three of the words will rhyme and one word will not rhyme with the other words. Here is an example. Listen carefully. “Cat, hat, yup, mat.” I’ll say them again. “Cat, hat, yup, mat.” Did you hear that “yup” was the word that does not rhyme? Listen. “Cat, hat, yup, mat.” Let’s try another one for practice. “Cap, book, tap, lap.” I’ll say it again. “Cap, book, tap, lap.” What word didn’t rhyme? (Correct if wrong and explain, “Cap and tap and lap all sound the same: cap, tap, lap, but *book* sounds different.”)
OK, get ready for the next words.

For each item: OK, get ready for the next one. (Words read three times) I’ll say it again…(Words) Which word did *not* rhyme?

1. Bee, see, tee, car.
2. Sat, red, bed, Ted.
3. Get, met, hot, wet.
4. Pop, jack, hop, cop.
5. Fell, go, no, blow.
6. Reaper, beeper, jumper, deeper.
7. Pet, fat, net, jet.
8. Faking, baking, taking, looking.

9. Clay, play, bride, day.

10. Why, me, pie, cry.

11. Please, cheese, knees, beans.

12. Lap, carp, harp, tarp.

Task 2. Rhyme Production.

Script: Now we are going to play a different rhyming game. I will say a word, and then you think of another word that rhymes with the word I say. For example, if I say, "Cat," you might say, "hat" or you might say, "mat," or you might say a silly word, like "wat." If I say, "Key," then you say… (if child gives rhyme, note and say, "Good." If child does not give rhyming word, say, If I say "key, you could say "me.") OK, now get ready.

For each item. "Tell me a rhyme for ____________

Word list:

Bean, blue, pick, wet, door, bud, got, pop, deep, look.

Appendix E.

Spelling Production Word List. Adapted from Santa, C.M., & Hoein, T. (1999)

Practice Item:

Say the word nose. (gesture to nose). What sound do you hear at the beginning? What letter makes that sound? Can you write that letter? Do you hear any more sounds? (etc).

Test Items. Word was read to child once. The child was asked to repeat the word aloud ("Now you say it"). Sentences were provided for all children, though it was intended by the original authors that sentences be given only if there were confusions about what the word was. Each word was repeated following the sentence.

1.	back	I hurt my back (gesture to back).
2.	feet	These are my feet (gesture).
3.	step	Don't step on me.
4.	junk	That old chair is a piece of junk.
5.	picking	My mom is picking me up after school.
6.	mail	I got a letter in the mail.
7.	side	The side of the building is painted red.
8.	chin	He rubbed his chin. (gesture to chin)
9.	dress	She has on a pretty pink dress.
10.	peeked	The girl peeked through her fingers (gesture)
11.	lamp	It's dark in here, turn on the lamp.
12.	road	Don't go in the road, you'll get hit by a car.

Appendix F.

Scoring System for Pretest Spelling Task, Adapted from Santa and Hoein (1999).

Note: The general rubric called for awarding one point for providing an initial letter or letters for the first sound in the word that was deemed to be correct or phonetically appropriate. Two or three points were given for providing, in addition, a letter or letters that represented a second sound in the word, either correctly or phonetically. Three points were given when a third and/or fourth sound was represented in this way, and four points were awarded for representing all the major sounds in the word in way that demonstrated knowledge of the orthographic system, even though the words were not spelled correctly (i.e. adding a final letter – dresse for dress, or representing a vowel with a phonetically correct pair of letters that was not the correct spelling – roed for road). Correct spellings of words were given 5 points.

The following chart illustrates common examples encountered.

Word	1 point	2 points	3 points	4 points
Back	b	bk, bc, ba,	bek,bik, baok, beak	bak, backe, bakk
Feet	v, f	ft, fait, fat, fe, foot	fyt, fet, fit	feat, fete

Appendix F (continued)

Word	1 point	2 points	3 points	4 points
Step	c, s	sp, se, sta, st, setp	sap, set, cap, stap	stepe, stepp
Junk	g, j	gc, jk, jc, gk, jo jeke, jak, jek, joc,	guc, juk, , juke, gok, juc, guk, gnk, gonk,	gunk, gunc, junc, juncke
Picking	p	pi, pg, pn, pcn, peke Pk, pin, pac, pig, pik	peking, pekn, pekin, picg, pikn, picin, pick	piking, picing
Mail	m	ml, ma, mial, mlae,	mal, mall, mel, mael	male, maill, maile
Side	s, c	sd, sed, si, sud	seid, sid, sod, cid, sade, sad	sied
Chin	c, j, g, h, t, n	ch, cn, ci, hn, ce,	cind, shen, thon, chen	chine
Dress	d, g, j,	ds, js, drs, gs, gras,	des, dais, drres, dras, jres	dres, dresse

Appendix F (continued)

Word	1 point	2 points	3 points	4 points
Peeked	p	pk, pikt, pen, peet, pct	pect, pekd, pek,	peaked
Lamp	l	lp, la, lm, lip	Lel, lap, lape, lam, lmp	lampe
Road	r, w	rd, ro, rodt, wd, rud, rde	rod, wod, raod, roweid	rode, roed, rowed

Appendix G

Training Words.

Script: I am going to teach you to read some words on cards.

This word says, __________. (Sentence. Word.) Now you look at it and read it to me."). (Child responds). Examiner Response, if correct: Good. If child makes an error: This word says _______.

BiLL I have one dollar bill.

DiLL I put a green vegetable in my soup; it's called dill.

MiLL A farmer grinds up wheat in a mill.

PiLL When I'm sick the doctor gives me medicine in a pill.

DUCK "Quack quack" says the yellow duck.

BUCK A boy deer is called a buck.

PUCK When I play the game hockey, I use my stick to hit the puck.

MUCK The pig plays in the brown muck.

MARK I did well on my test; I got a good mark.

PARK I like to go on the swings in the park.

BARK A dog has a very loud bark.

DARK I turn on the light when it gets dark.

PALE A red apple is shiny and bright, but a yellow apple is light and pale.

MALE A girl is a female, a boy is a male.

DALE A girl or a boy could have the name Dale.

BALE The farmer makes a big pile of hay and ties it together and it is called a bale.

Appendix H

Posttest Spelling Production

Script:

I am going to tell you some words and ask you to write them on this paper.

Remember the words that we did together?

If you are not sure how to spell the word, make your best guess. If you don't know all of the letters in the word, you can write some of them.

Words:

dill, male, bark, puck

duck, mark, bill, park

Script: Now I am going to ask you to write some silly words. If I ask you to write "yad," you can write it like this "/y/ /a/ /d/" (Show child y-a-d.) "Yad" doesn't really mean anything. It's a silly word, but I can say it and I can write it down. Now you try it.

Nonwords (transfer words): *juck, sark, rill, lale*

No feedback given.

Appendix I.

Posttest Spelling Recognition Task 1.

Script: I am going to read you some words, and show you a few choices. Remember the words that we did together? (Show one word and foils at a time.) For each word: These words say, ________. Which spelling do you think is right? Circle the one you think is right.

1.	BiLL	BULL	BiL
2.	MUK	MUNK	MUCK
3.	PACK	PARK	PRK
4.	DAL	DALE	DAKE
5.	MARK	MACK	MRK
6.	DiL	DULL	DiLL
7.	BAKE	BALE	BAL
8.	PUNK	PUK	PUCK
9.	PALE	PAKE	PAL
10.	BRK	BACK	BARK
11.	DUNK	DUCK	DUK
12.	MULL	MiL	MiLL
13.	DARK	DACK	DRK
14.	PiL	PULL	PiLL
15.	MAKE	MALE	MAL
16.	BUCK	BUK	BUNK

17. <u>LUCK</u> LUK LUNK

18. LRK <u>LARK</u> LACK

19. <u>SiLL</u> SULL SiL

20. TAL <u>TALE</u> TAKE

21. <u>CLUCK</u> CLUNK CLUK

22. SHRK <u>SHARK</u> SHANK

23. <u>SPiLL</u> SPULL SPiL

24. STAL STAKE <u>STALE</u>

<u>Spelling Recognition Task 2</u>

Items:

1.	<u>DiLL</u>	DARK	DALE	DUCK
2.	MALE	MiLL	<u>MARK</u>	MUCK
3.	PUCK	PARK	PiLL	<u>PALE</u>
4.	BALE	<u>BUCK</u>	BARK	BiLL
5.	PARK	BARK	<u>MARK</u>	DARK
6.	BiLL	<u>DiLL</u>	PiLL	MiLL
7.	<u>PALE</u>	MALE	BALE	DALE
8.	DUCK	MUCK	PUCK	<u>BUCK</u>

References

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. (2000). Washington, DC: American Psychiatric Association.

Bangerter, A. (2002). Teaching of word level work in the literacy hour and its implications for educational psychologists' assessment and intervention. Educational Psychology in Practice, 18, 5-19.

Bowey, J.A. & Francis, J. (1991). Phonological analysis as a function of age and exposure to reading instruction. Applied Psycholinguistics, 12, 91-123.

Bowey, J.A. & Underwood, N. (1996). Further evidence that orthographic rime usage in nonword reading increases with word-level reading proficiency. Journal of Experimental Child Psychology, 63, 526-562.

Bradley, L. & Bryant, P.E. (1983). Catergorising sounds and learning to read – a causal connection. Nature, 301, 419-521.

Bradley,L. & Bryant, P.E. (1985). Rhyme and reason in reading and spelling. I.A.R.L.D. Monographs No.1, Ann Arbor: University of Michigan.

Bruck, M. & Treiman, R. (1992). Learning to pronounce words: The limitations of analogies. Reading Research Quarterly, 27, 374-389.

Bryant, P., MacLean, M., & Bradley,L. (1990). Rhyme, language and children's reading. Applied Psycholinguistics, 2, 237-252.

Christensen, C.A. (1997). Onsets, rimes, and phonemes in learning to read. Scientific Studies of Reading, 1, 341-358.

Cohen, J. A. (1992) A Power Primer. Psychological Bulletin, 112, 155-159.

Bryant, P.E. & Bradley, L. (1985). Children's reading problems. NY: Basil Blackwell, Inc.

Byrne, B. (1998). The foundation of literacy: The child's acquisition of the alphabetic principle. Hove, East Sussex: Taylor & Francis Group.

Deavers, R. P., Solity, J.E., & Kerfoot, S. (2000). The effect of instruction on early nonword reading strategies. Journal of Research in Reading, 23, 267-286.

Duncan, L.G., Seymour, P.H.K., & Hill, S. (2000). A small-to-large unit progression in metophonological awareness and reading? The Quarterly Journal of Experimental Psychology, 53A, 1081-1104.

Edelen Smith, P.J. (1997) How now brown cow: phoneme awareness activities for collaborative classrooms. Intervention in School and Clinic, 33, 103-11.

Ehri, L. (1991). Development of the ability to read words. In R. Barr, M. Kamil, P. Mosenthal and P. Pearson (Eds.), Handbook of Reading Research (Vol II, pp.383-417). NY: Longman.

Ehri, L. (1994). Development of the ability to read words: Update. In R. Ruddell, M.Ruddell, and H.Singer (Eds.), Theoretical models and processes of reading (4th edition pp.323-358). Newark, DE: International Reading Association.

Ehri, L. (1995). Phases of development in learning to read words by sight. Journal of Research in Reading, 18, 116-125.

Ehri, L. (1998a). Learning to read and learning to spell are one and the same, almost. In C.A. Perfetti, L. Rieben, & M. Fayol. (Eds.) Learning to spell: Research, theory, and practice across languages (pp. 237-270). Mahwah, NJ: Lawrence Erlbaum Associates.

Ehri, L. (1998b). Word reading by sight and by analogy in beginning readers. In C. Hulme & R.M. Joshi (Eds.) Reading and spelling: Development and disorders. Mahwah, NJ: Lawrence Erlbaum Associates.

Ehri, L. & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. Reading and Writing Quarterly: Overcoming learning difficulties, 14, 135-163.

Ehri, L. & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. Reading Research Quarterly, 27, 12-26.

Ellis, N.C. & Large, B. (1987). The development of reading: As you seek so you shall find. British Journal of Developmental Psychology, 78, 1-28.

Gaskins, I.W., Ehri, L.C., Cress, C., O'Hara, C. & Donnelly, K. (1996-97). Procedures for word learning: Making discoveries about words. The Reading Teacher, 50, 312-327.

Gilbert, N., Spring, C., & Sassenrath, J. (1977). Effects of overlearning and similarity on transfer in word recognition. Perceptual and Motor Skills, 44, 591-598.

Glushko, R. (1979a). Cognitive and pedagogical implications of orthography. Quarterly Newsletter of the Laboratory of Comparative Human Cognition, 1, 22-26.

Glushko, R. (1979b). The organization and activation of orthographic knowledge in reading aloud. Journal of Experimental Psychology: Human Perception and Performance, 5, 674-691.

Glushko, R.J. (1981). Principles for pronouncing print: The psychology of phonography. In A.M. Lesgold & C.A. Perfetti, (Eds.) Interactive processes in reading, NJ: Erlbaum.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. Journal of Experimental Child Psychology, 42, 73-83.

Goswami, U. (1988a). Children's use of analogy in learning to spell. British Journal of Developmental Psychology, 6, 21-33.

Goswami, U. (1988b). Orthographic analogies and reading development. Quarterly Journal of Experimental Psychology, 40A, 239-268.

Goswami, U. (1990a). Phonological priming and orthographic analogies in reading. Journal of Experimental Child Psychology, 49, 323-340.

Goswami, U. (1990b). A special link between rhyming skills and the use of orthographic analogies by beginning readers. Journal of Child Psychology and Psychiatry, 31, 301-311.

Goswami, U. (1991). Learning about spelling sequences: The role of onsets and rimes in analogies in reading. Child Development, 62, 1110-1123.

Goswami, U. (1992). Analogical reasoning in children. East Sussex, UK: Erlbaum.

Goswami, U. (1994). Reading by analogy: Theoretical and practical perspectives. In C. Hulme & M. Snowling (Eds.), Reading development and dyslexia. San Diego, CA: Singular Publishing Group.

Goswami, U. (1998). Rime-based coding in early reading development in English: Orthographic analogies and rime neighborhoods. In C. Hulme & R.M. Joshi (Eds.) Reading and spelling: Development and disorders. Mahwah, NJ: Lawrence Erlbaum Associates.

Goswami, U. (1999a). Phonological development and reading by analogy: Epilinguistic and metalinguistic issues. In J. Oakhill & R. Beard (Eds.) Reading development and the teaching of reading, Malden, MA: Blackwell Publishers.

Goswami, U. (1999b). Causal connections in beginning reading: The importance of rhyme. Journal of Research in Reading, 22, 217-240.

Goswami, U. & Bryant, P. (1990). Phonological skills and learning to read. East Sussex, U.K.: Erlbaum.

Goswami, U. & East, M. (2000). Rhyme and analogy in beginning reading: Conceptual and methodological issues. Applied Psycholinguistics, 21, 63-93.

Goswami, U. & Mead, F. (1992). Onset and rime awareness and analogies in reading. Reading Research Quarterly, 27, 152-162.

Greaney, K.T., Tunmer, W.E., & Chapman, J.W. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. Journal of Educational Psychology, 89, 645-651.

Hanley, J.R., Reynolds, C.J., & Thornton, A. (1997). Orthographic analogies and developmental dyslexia. British Journal of Psychology, 88, 423-440.

Insideschools, (n.d.) School profile. Retrieved December 15, 2004, from http://www.insideschools.org

Juel, C. & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. Reading Research Quarterly, 35, 458-492.

Juel, C., Griffith, P.L., & Gough, P.B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. Journal of Educational Psychology, 78, 243-255.

Knafle, J. D. (1974). Children's discrimination of rhyme. Journal of Speech and Hearing Research, 17, 367-372.

Leslie, L. & Calhoon, A. (1995). Factors affecting children's reading of rimes: reading ability, word frequency, and rime-neighborhood size. Journal of Educational Psychology, 87, 576-586.

Levy, B.A., Bourassa, D.C., & Horn, C. (1999). Fast and slow namers: Benefits of segmentation and whole word training. Journal of Experimental Child Psychology, 73, 115-138.

Lovett, M.W., Warren-Chaplin, P.M., Ransby, M.J., & Borden, S.L. (1990). Training the word recognition skills of reading disabled children: Treatment and transfer effects. Journal of Educational Psychology, 82, 769-780.

Macmillan, B.M. (2002). Rhyme and reading: a critical review of the research methodology. Journal of Research in Reading, 25, 4-42.

Marchbanks, G. & Levin, H. (1965). Cues by which children recognize words. Journal of Educational Psychology, 56, 57-61.

Marsh,G., Desberg, P., & Cooper, J. (1977). Developmental strategies in reading. Journal of Reading Behavior, 9, 391-394.

Marsh, G. Friedman, M.P., Welch, V., & Desberg,P. (1980a). A cognitive-developmental approach to reading acquisition. In G.E. MacKinnon and T.G. Waller, (Eds), Reading research. Advances in theory and practice. Vol 3. NY: Academic Press.

Marsh,G. Friedman, M.P., Welch, V., & Desberg,P. (1980b). The development of strategies in spelling. In U. Frith (ed.) Cognitive processes in spelling. London: Academic Press.

Marsh, G., Friedman, M.P., Desberg, P. & Saterdahl, K. (1981). Comparison of reading and spelling strategies in normal and reading disabled children. In M.P. Friedman, J.P. Das and N. O'Conner (Eds). Intelligence and learning (pp.363-367). NY: Plenum.

Masonheimer, P.E., Drum, P.A., & Ehri, L.C. (1984). Does environmental print lead children into word reading? Journal of Reading Behavior, 16, 257-71.

Masterson, J. (1983). Sublexical and operational routes in dyslexia. Unpulished PhD dissertation, University of London.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. Journal of Experimental Child Psychology, 71, 3-27.

Muter, V., Snowling, M.J., & Taylor, S. (1994) Orthographic analogies and phonological awareness: Their role and significance in early reading development. Journal of Child Psychology and Psychiatry, 35, 293-310.

Nation, K. & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. Reading Research Quarterly, 32, 154-167.

O'Shaugnessy, T.E. & Swanson, H.L. (2000). A comparison of two reading interventions for children with reading disabilities. Journal of Learning Disabilities, 33, 257-278.

Peterson, M.E. & Haines, L.P. (1992). Orthographic analogy training with kindergarten children: Effects on analogy use, phonemic segmentation, and letter-sound knowledge. Journal of Reading Behavior, 24, 109-127.

Post, Y.V., Carreker, S. & Holland, G. (2001). The spelling of final letter patterns: A comparison of instruction at the level of the phoneme and the rime. Annals of Dyslexia, 51, 121-146.

Rayner, K. (1976). Developmental changes in word recognition strategies. Journal of Educational Psychology, 68, 323-329.

Read, C. (1971). Preschool children's knowledge of English phonology, Harvard Educational Review, 41, 1-34.

Read, C. (1978). Children's awareness of language, with emphasis on sound systems. In A.Sinclair, R.J. Jarvella & W.J.M. Levelt (Eds.), The child's conception of language. Berlin: Springer.

Roswell, F. G. & Chall, J. S. (1992). Roswell Chall Screening Tests, Auditory Blending Test. Cambridge, MA: Educators Publishing Service.

Santa, C.M., & Hoein, T. (1999). An assessment of Early Steps: A program for early intervention of reading problems. Reading Research Quarterly, (34) 1, 54-79.

Savage, R. (1999). A re-evaluation of the evidence for orthographic analogies: A reply to Goswami. Journal of Research in Reading, 24, 1-18.

Savage, R. & Stuart, M. (1998). Sublexical inferences in beginning reading: Medial vowel digraphs as functional units of transfer. Journal of Experimental Child Psychology, 69, 85-108.

Scarborough, H.S. (1990). Very early language deficits in dyslexic children. Child Development, 61, 1728-1743.

Scott, J.A., & Ehri, L.C. (1990). Sight word reading in prereaders: Use of logographic vs. alphabetic access routes. Journal of Reading Behavior, 2, 149- 166.

Snowling, M., Goulandris, N., & Defty, N. (1998). Development and variation in developmental dyslexia. In C. Hulme & R.M. Joshi (Eds.) Reading and spelling: Development and disorders. Mahwah, NJ: Lawrence Erlbaum Associates.

Solity, J.E., Deavers, R.P., Kerfoot, S.R., Crane,G., & Cannon, K. (1999). Raising literacy attainment in the early years: the impact of instructional psychology. Educational Psychology, 19 (4), 373-397.

Spring, C., Gilbert, N. & Sassenrath, J. (1979). Learning to read words: Effects of overlearning and similarity on stimulus selection. Journal of Reading Behavior, XI, 1, 69-71.

Stahl, S.A., & Murray, B.A. (1994). Defining phonological awareness and its relationship to early reading. Journal of Educational Psychology, 86, 221- 234.

Stanback, M.L. (1992). Syllable and rime patterns for teaching reading: Analysis of a frequency-based vocabulary of 17, 602 words. Annals of Dyslexia, 42, 196-221.

Stanovich, K.E. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-407.

Treiman, R. (1985). Onset and rimes as units of spoken syllables: Evidence from children. Journal of Experimental Child Psychiatry, 39, 161-181.

Treiman, R. (1992). The role of intrasyllabic units in learning to read and spell. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.) Reading Acquisition (p.65-106). Hillsdale, NJ: Earlbaum.

Treiman, R. (1993). Beginning to spell: A study of first-grade children. NY: Oxford University Press.

Vellutino, F. (1991) Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches in reading instruction. Journal of Educational Psychology, 83, 437-443.

Walton, P.D., Walton, L.M., & Felton, K. (2001). Teaching rime analogy or letter recoding reading strategies to prereaders: effects on prereading skills and word reading. Journal of Educational Psychology, 93, 160-180.

White, T.G. & Cunningham, P.M. (1990). Teaching disadvantaged students to decode by analogy. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA April.

Wise, B.W., Olson, R.K. & Treiman, R. (1990). Subsyllabic units in computerized reading instruction: onset-rime vs. post vowel segmentation. Journal of Experimental Child Psychology, 49, 1-19.

Wood, C. (2000). Rhyme awareness, orthographic analogy use, phonemic awareness and reading: An examination of relationships. Educational Psychology, 20, 5-15.

Woodcock, R. (1987). Woodcock Reading Mastery Tests – Revised. Circle Pines, MN: American Guidance Service.