Relations Among Verbal Working Memory, Listening Comprehension, and Reading Skills

Joanna S. Futransky

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss
RELATIONS AMONG VERBAL WORKING MEMORY,
LISTENING COMPREHENSION, AND
READING SKILLS

BY

JOANNA S. FUTRANSKY

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND
1992
ABSTRACT

Previous research suggests that verbal working memory deficits contribute to problems experienced in reading and in listening comprehension. The primary purpose of the present study was to explore the role of verbal working memory in listening comprehension for fifth-grade students. Additionally, the present research investigated the association between working memory and the two components of reading: decoding and comprehension. A third goal of the study was to investigate the power of listening comprehension, decoding ability, memory skills, and IQ to predict reading comprehension.

Data from 136 fifth-grade students with average to above average cognitive ability was analyzed for the study. Each student completed three verbal working memory tasks, two listening comprehension measures, one decoding test, and one reading comprehension test. Two listening comprehension measures were used to test the hypothesis that listening measures differing in memory requirements (recall vs. recognition) would produce divergent results. Students were divided into low, middle, and high memory groups based upon their scores on the working memory tasks.

Results of the two listening comprehension measures proposed that the memory demands affected comprehension accuracy. Significant memory group differences were observed on the measure necessitating the recall of specific
factual information but not on the task requiring the recognition of ideas.

Significant memory group differences were observed on both the decoding and reading comprehension measures. Interestingly, listening comprehension scores coupled with working memory scores emerged as the dyad that accounted for the greatest proportion of variance in reading comprehension.

The results called attention to the need to expand educational accommodations used with students with memory problems. Instructional accommodations, as well as direct instruction in metacognitive strategies, were recommended as helpful curriculum modifications.
ACKNOWLEDGEMENTS

Heartfelt appreciation is due to the many people who provided me with hours of expertise and support: to my major professor, Dr. Susan Brady, whose unending patience, warmth, humor, and encouragement, as well as professional excellence, guided me through the complex journey of this project; to Dr. Jerry Cohen, for his statistical expertise; and to Drs. Janet Kulberg and Barbara Culatta, for their valuable input and contributions of time. Special thanks go to Tara Cassidy and Joann Burns, my research assistants who provided invaluable help with data collection.

This project could not have succeeded without the support of public school personnel willing to allow me to work with their students. I extend much gratitude to Superintendent Raymond Spear of the Coventry School Department, Principal Robert Bates of the Narragansett Pier School, Principal Paul Johnson of the West Kingston Elementary School, and the thirteen talented teachers who cheerfully worked with me for six months.

Finally, I send immeasurable bounties of love and appreciation to my special husband and our family, to my parents and sister, to my great uncle and aunt, and to my friends for being my unconditional champions.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION AND OVERVIEW</td>
<td>1</td>
</tr>
<tr>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>Listening comprehension and reading performance</td>
<td>3</td>
</tr>
<tr>
<td>The theoretical role of verbal working memory in reading and listening comprehension</td>
<td>6</td>
</tr>
<tr>
<td>Verbal working memory and reading ability</td>
<td>11</td>
</tr>
<tr>
<td>Evidence that good and poor readers differ in listening comprehension: Relating reading group differences to verbal working memory</td>
<td>16</td>
</tr>
<tr>
<td>Recall and recognition</td>
<td>22</td>
</tr>
<tr>
<td>Present investigation</td>
<td>27</td>
</tr>
<tr>
<td>II. METHOD</td>
<td>29</td>
</tr>
<tr>
<td>Subjects</td>
<td>29</td>
</tr>
<tr>
<td>Selection criteria</td>
<td>29</td>
</tr>
<tr>
<td>Measures</td>
<td>31</td>
</tr>
<tr>
<td>Memory measures</td>
<td>31</td>
</tr>
<tr>
<td>Decoding measure</td>
<td>36</td>
</tr>
<tr>
<td>Listening and reading comprehension measures</td>
<td>36</td>
</tr>
<tr>
<td>Experimental procedures</td>
<td>39</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>41</td>
</tr>
<tr>
<td>Characteristics of the total sample</td>
<td>42</td>
</tr>
<tr>
<td>Characteristics of the three memory groups for the full sample</td>
<td>43</td>
</tr>
<tr>
<td>Correlations between variables</td>
<td>48</td>
</tr>
<tr>
<td>ANOVAs: listening comprehension</td>
<td>50</td>
</tr>
<tr>
<td>ANOVAs: reading measures</td>
<td>57</td>
</tr>
<tr>
<td>The effects of IQ</td>
<td>61</td>
</tr>
<tr>
<td>Predicting reading comprehension</td>
<td>62</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>66</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>82</td>
</tr>
</tbody>
</table>
APPENDICES

A. Familiar Word Span Test................................. 90
B. Pseudoword Repetition Test............................. 91
C. Sentence Span Test....................................... 92
D. Profiles in Listening and Reading (PILAR)........... 95
E. Means, standard deviations, and ranges in raw scores (except where noted) for the sample population (N=136), for the dependent measures................................. 100
F. Means, standard deviations, and ranges in raw scores (except where noted) for the subsample (N=76), for the dependent measures................................. 101
G. Correlation matrix of age, IQ, memory measures, PILAR listening and reading comprehension tasks, Spache listening comprehension, and decoding, (N=76)................................. 103
H. ANCOVA to determine memory group differences when the WISC-R Vocabulary and subtests are used as covariates................................. 104

BIBLIOGRAPHY..................................................... 105
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Means, standard deviations, and ranges of age and IQ for the total experimental sample and the subsample</td>
<td>42</td>
</tr>
<tr>
<td>2.</td>
<td>Correlation matrix for age, IQ, memory measures, PILAR listening and reading comprehension tasks, and decoding, N=136</td>
<td>45</td>
</tr>
<tr>
<td>3.</td>
<td>Means and standard deviations for age and IQ for each memory group, and analyses of group differences with follow-up Tukey results</td>
<td>46</td>
</tr>
<tr>
<td>4.</td>
<td>Means and standard deviations for the working memory tasks for each memory group, and analyses of group differences with follow-up Tukey results, N=136</td>
<td>47</td>
</tr>
<tr>
<td>5.</td>
<td>Means and standard deviations for three memory groups on the three grade levels of passage difficulty of the PILAR listening comprehension task, N=136</td>
<td>51</td>
</tr>
<tr>
<td>6.</td>
<td>Summary table for 3x3 ANOVA for PILAR listening comprehension scores by memory group and levels of passage difficulty, N=136</td>
<td>51</td>
</tr>
<tr>
<td>7.</td>
<td>Means and standard deviations for the two memory groups on the three grade levels of passage difficulty of the PILAR listening comprehension tasks, N=60</td>
<td>53</td>
</tr>
<tr>
<td>8.</td>
<td>Summary table for 2x3 ANOVA for PILAR listening comprehension scores by memory group and levels of passage difficulty, N=60</td>
<td>54</td>
</tr>
<tr>
<td>9.</td>
<td>Means and standard deviations for two memory groups on the three grade levels of passage difficulty of the Spache listening comprehension task, N=60</td>
<td>56</td>
</tr>
</tbody>
</table>
10. Summary table for 2x3 ANOVA for the Spache Diagnostic Scales scores by memory group and levels of passage difficulty, N=60.

11. Means, standard deviations, and ranges for Word Attack subtest for each memory group, and analyses of group differences with follow-up Tukey results, N=136.

12. Means and standard deviations for three groups on the three grade levels of passage difficulty of the PILAR reading comprehension task, N=136.

13. Summary table for 3x3 ANOVA for PILAR reading comprehension scores by memory group and levels of passage difficulty, N=136.

14. Hierarchical multiple regressions with reading comprehension as the dependent variable: N=136 without Spache; N=76 with Spache.
LIST OF FIGURES

Figures                                      PAGE
1. Mean correct responses, in raw scores, for the low, middle, and high memory groups on the PILAR listening comprehension measures, N=136........... 52

2. Mean correct responses, in raw scores, for the low and high memory groups on the PILAR listening comprehension measures, N=60..................... 54

3. Mean correct responses, in raw scores, for the low and high memory groups on the Spache listening comprehension task, N=60............................. 57

4. Mean correct responses, in raw scores, for the low, middle, and high memory groups on the PILAR reading comprehension test, N=136............ 60
INTRODUCTION AND OVERVIEW

Reading disabled children have often been reported to have working memory deficits. These working memory problems have been observed in studies using digit span measures, word and sentence span measures, letter strings, lists of pseudowords, and the recall of sequences of familiar objects (see Brady, 1991, for a review). There is rising controversy, however, about the nature of the association between working memory deficits and reading disabilities. Do working memory deficits impact upon early stages of reading when decoding is the primary task (Perfetti, 1985) or later, when the comprehension processes become more demanding (Pennington, Van Orden, Kirson, & Haith, 1991)?

A second area of research suggests that some reading disabled children experience listening comprehension problems when compared to normal reading peers (Mann, Cowin, & Shoenheimer, 1989, 1990; Shankweiler, 1989; Shankweiler & Crain, 1986; Smith, Mann, & Shankweiler, 1986; Stanovich, 1982b). However, the listening problems appear to be dependent upon the memory demands of the task, with comprehension deficits evident for more complex material and when more demanding recall measures are used (Torgesen, 1988). As a result, several researchers have implicated working memory deficits as strongly associated with the observed listening problems (Daneman & Carpenter, 1980, 1983;
Mann, Cowin, & Shoenheimer, 1990; Shankweiler, 1989; Shankweiler & Crain, 1986). For example, Mann et al. (1990) proposed that "a limited ability to hold linguistic materials in short term memory is one aspect of a phonological processing problem...and this limitation can lead poor readers to misunderstand certain types of phrases and sentences presented orally" (p.60).

The picture emerging suggests that working memory deficits contribute in some way to the problems experienced in reading (decoding and comprehension) and in listening comprehension (Baddeley, 1986; Daneman & Carpenter, 1980). However, only a few studies have formally addressed the relationships between reading skills, listening comprehension, and verbal working memory. While research with college students reported strong correlations between verbal working memory, listening comprehension, and reading ability (Daneman & Carpenter, 1980), studies involving learning-disabled children with known working memory deficits obtained a somewhat different pattern of results. The children could remember the gist of information presented orally as well as students without working memory deficits, but could not demonstrate verbatim recall of phrases and complex sentences (Mann et al., 1990; Torgesen, Rashotte, & Greenstein, 1988).

The purpose of the present study is to examine in greater depth the association of verbal working memory with
listening comprehension for children in the fifth grade. The present research also investigates the correspondence between verbal working memory and reading skills (both decoding and comprehension). The levels of oral and written passage difficulty to be comprehended (third-, fifth-, and seventh-grade material) and the demands of the listening tasks (recall vs. recognition) are manipulated to allow for a more sensitive evaluation of the role of verbal working memory in reading and listening performance.

Before describing the study in full, several background issues will be presented. First, the current literature examining the relationship between reading ability and listening comprehension will be reviewed. Second, the theoretical role of verbal working memory in reading processes will be discussed. And third, the evidence on the verbal working memory and listening comprehension skills of poor readers will be examined.

Listening Comprehension and Reading Performance

In the *Handbook of Reading Research*, Sticht and James (1984) identified three premises regarding the relationship between oral and written language: 1) oral language skills develop to fairly high levels prior to the development of written language skill; 2) oral and written language share the same vocabulary and grammar; and 3) beginning readers draw upon their knowledge of oral language in learning to
read. From these premises, Sticht and James concluded that oral language comprehension establishes a goal or limit for reading comprehension, at least during the first few years of reading instruction. This position was restated by Royer, Kulhavy, Lee, and Peterson (1986) who proposed that reading comprehension is a special case of listening comprehension, and that listening comprehension places an upper bound on reading ability. While students are learning to read, they frequently understand material presented orally at what corresponds to one to two grade-levels higher than their reading comprehension levels (Spache, 1981). As decoding and word recognition skills improve, the gap between the ability to comprehend oral language and the ability to understand the same material in written form narrows.

Research evidence corroborates the initial superiority of listening over reading comprehension, and indicates that the interval closes between the sixth- and eighth-grade reading levels. Higher correlations between listening and reading comprehension are obtained as grade levels increase (Curtis, 1980; Sticht & James, 1984). For example, Curtis (1980) obtained correlations between! reading and listening comprehension of .26 for second grade,.66 for third grade, and .74 for fifth grade. Further, a re-evaluation of the data from the national norming sample for the Durrell Listening-Reading Series indicates that listening and
reading comprehension scores converge in the middle of sixth grade for paragraph comprehension tasks, and at around eighth grade for combined paragraph comprehension and vocabulary scores.

Accordingly, Stanovich (1982b) proposed that "...listening comprehension can account for a proportion of the variance in reading skill that is not accounted for by decoding or other component processes" (p.551). As children's decoding skills improve and they become better readers, the proportion of the variance accounted for by listening comprehension ability increases. Curtis (1980) found that listening comprehension accounted for the most unique variance in reading comprehension for a fifth-grade population. Making a related point, the Report of the Commission on Reading (Anderson, 1985) stated that in a nationwide study involving thousands of students, listening comprehension in the fifth grade was the best predictor of aptitude and achievement in high school.

The hypothesis that reading skill is the product of decoding and listening comprehension was later advanced by Gough and Tunmer (1986). According to this view, decoding enables orthographic input to access or connect with the language system; listening comprehension provides background knowledge and reasoning ability required for comprehension, whether of spoken or written language. Consequently, difficulties in either decoding or listening comprehension
could limit successful reading comprehension.

Considering reading performance from this perspective implies that when a student fails to understand a written passage, it is necessary to determine whether the lack of comprehension stems from difficulty identifying the written words (decoding) or from a more generalized language comprehension problem (Carlisle, 1989a; Royer et al., 1986; Shankweiler, 1989). Could the child comprehend the same passage if it were read out loud? As children master decoding skills, listening comprehension ability may play an increasingly important role in limiting what can be comprehended in print.

The Theoretical Role of Verbal Working Memory In Reading and Listening Comprehension

Verbal working memory is currently defined as a dynamic memory system with both storage and processing functions (Baddeley, 1986; Crain, 1989; Daneman & Carpenter, 1980; Liberman, Shankweiler, & Liberman, 1989; Wagner & Torgesen, 1987). Current theory regards verbal working memory as a limited capacity system with finite attentional resources that stores and processes linguistic information for brief periods of time, whether the source is oral or written (Baddeley, 1986; Wagner & Torgesen, 1987). Consequently, a reading or listening task with heavy processing demands is hypothesized to decrease the resources available for storage. The functional capacity of this limited system
appears to be related to the efficiency of phonological coding. It is proposed that a person who encodes phonological information inefficiently will expend more attentional resources for initial processing and will have less left over for storage (Daneman & Carpenter, 1980; Perfetti, 1985). Conversely, greater efficiency or automaticity of phonological coding will result in more resources available for storage (e.g., Case, Kurland, & Goldburg, 1982; LaBerge & Samuels, 1974; see Brady, 1991, for a review). Thus, individual differences in working memory capacity are thought to be due, at least in part, to variability in efficiency of creating phonological representations.

Within this framework, working memory has been thought to serve central functions in the reading process. First, it has been suggested that working memory may have an important role in the acquisition of decoding. In several studies, students with verbal working memory deficits have consistently shown special difficulty in the rapid and accurate reading of individual words (Perfetti, 1985; Stanovich, 1982b; Torgesen, 1988). Phonological coding deficits in memory may make decoding a more difficult skill to learn (Wagner, 1986; 1988).

Poor decoding skills no doubt have further consequences for comprehension. According to Perfetti and Roth (1981), skilled reading is an interactive process between bottom-up
(data-driven) decoding processes and top-down (conceptually-driven) comprehension processes. Words are concurrently decoded and held in working memory, and their meanings are considered. This implies that an important step in reading skill development is the occurrence of automaticity in decoding ability. When a person can decode automatically, more attentional resources can be directed toward understanding what has been decoded (Curtis, 1980; Samuels, 1987; LaBerge & Samuels, 1974; Perfetti & Roth, 1981).

Within this framework, Perfetti and Roth (1981) propose that difficulties in reading comprehension frequently result from inefficient, nonautomatic decoding skills. Supporting this claim, studies which evaluate the speed and accuracy of pseudoword reading have consistently found a strong relationship between this task and aspects of reading comprehension (Hogaboam & Perfetti, 1978; Perfetti & Hogaboam, 1975; Stanovich, 1988; Stanovich, Cunningham, & Freeman, 1984). Correlations between word recognition skills and reading comprehension have consistently been within the range of .50 to .80 (Stanovich, 1982b).

Curtis (1980) hypothesized that the relationship of comprehension skill to reading achievement should be an inverse function of the attentional demands of the decoding process. For poor readers, phonological deficits in working memory may reduce the automaticity of their decoding skills, which in turn may negatively impact upon their ability to
meaningfully store and integrate what is read for comprehension. That is, if too many attentional resources are allocated to the mechanics of figuring out what the word is, few will be left over for higher level processes that would lead to comprehension. The incoming verbal information becomes 'bottlenecked' at the phonological stage of processing and, because of the limited duration of working memory, fades before it can be meaningfully integrated with preceding and succeeding material (Shankweiler & Crain, 1986).

Further evidence suggests that limitations in verbal working memory capacity may affect reading comprehension even when decoding skills are considered adequate. College students with small working memory capacities demonstrated comprehension deficits when compared to adults with larger verbal memory spans (Baddeley, Logie, & Nimmo-Smith, 1985; Daneman & Carpenter, 1980,1983). Correlations were obtained, ranging from .72 to .86, between memory span tasks and reading comprehension measures in adults (Daneman & Carpenter, 1980). In a study designed to examine how readers integrate successive words into their current understanding of a text, Daneman and Carpenter (1983) also found a significant correlation between memory span size and college students' ability to recall information verbatim from passages just read. Most importantly, they found an inverse relationship between verbatim recall ability and
passage comprehension errors. In a project designed to evaluate the processes involved in normal adult reading, Baddeley, Logie, and Nimmo-Smith (1985) also concluded that verbal working memory made important and independent contributions to reading comprehension.

Verbal working memory may have an additional effect upon reading performance by impacting upon listening or general language comprehension. The Processing Limitation Hypothesis (Shankweiler, 1989) proposes that a modality-free phonological processing deficit in verbal working memory will have consequences for the processing of all verbal material, whether the input is from print or from oral language (Crain, 1989; Perfetti, 1985; Shankweiler, 1989; Shankweiler & Crain, 1986). Listening is a temporal activity that requires simultaneously holding information in memory while integrating new information with that which came earlier. Individuals with memory deficits may have difficulty retaining the context in which to relate new information. Accordingly, Daneman and Carpenter (1980) found that subjects with larger verbal working memory spans were able to retrieve pronoun referents in orally presented passages when the pronouns were at greater distances from their referent than could subjects with smaller spans. Of significance, pronoun retrieval difficulties experienced by subjects with smaller memory spans resulted in impaired understanding of the passages. Thus, verbal working memory
deficits may prevent the full processing of spoken sentences (Mann, Cowin, & Shoenheimer, 1989).

In sum, it has been proposed that phonological processing deficits in working memory impede the development of automatic decoding skills, and reduce the amount of attentional resources available for storage. Reduced storage capacity, in turn, may negatively affect both reading and listening comprehension.

Verbal Working Memory and Reading Ability

In light of the theoretical framework postulating a link between working memory skills and reading/listening abilities, what is the evidence that poor readers are characterized by working memory deficits? Research has repeatedly demonstrated that poor readers have deficits in working memory when compared to good readers (for reviews, see Brady, 1986, 1991; Stanovich, 1982a; Wagner & Torgesen, 1987). Further, the deficits seem to be specific to linguistic material (see Brady, 1991, and Vellutino, 1979, for reviews). For example, in two separate studies, good and poor readers differed in their ability to remember items that can be linguistically coded, but did not differ in their ability to recall "doodle" drawings or unfamiliar faces (Katz, Shankweiler, & Liberman, 1981; Liberman, Mann, Shankweiler, & Werfelman, 1982). In another set of experiments, Torgesen, Rachotte, Greenstein, & Portes (1988)
presented nine different memory tasks to nondisabled students and to learning-disabled students with known memory deficits. Again, the learning-disabled students performed less well than the nondisabled students on those tasks requiring verbal memory skills, but did not differ on the task requiring visual memory for unfamiliar designs.

Following an extensive re-evaluation of the memory research involving children identified as learning disabled, Vellutino (1979) concluded that there were few differences between good and poor readers' abilities to remember visual stimuli that could not be phonologically coded, but noteworthy differences between the two groups when the information to be remembered could be verbally labeled. Thus, it appears that many reading disabled students have recall difficulties specific to linguistic information (Katz et al., 1981).

Further, a pivotal study by Shankweiler, Liberman, Mark, Fowler, and Fisher (1979) demonstrated that poor readers have a verbal memory problem that extends beyond written material, and that appears to have a phonological origin. Good, average, and poor readers were compared on three tasks measuring the recall of letter strings. The strings were either presented auditorily or visually and varied in terms of phonological confusability (i.e., whether the items rhymed or not). Shankweiler et al. (1979) found two noteworthy results: 1) the modality of presentation did
not affect performance for any of the groups, suggesting that coding strategies were similar for oral or written letters; and 2) poor readers recalled less and were less sensitive to the manipulation of rhyme than were the good readers, indicating less efficient phonological coding.

In a supporting study, Brady, Mann, and Schmidt (1987) presented good and poor readers with nonsense syllables for recall which varied systematically in terms of phonological similarity. Both good and poor readers made phonological errors, suggesting the use of the same coding strategies. Again, however, the poor readers made significantly more errors, implying that the strategies are employed less effectively.

Evidence from studies with deaf students strengthen the argument that phonological coding in verbal memory is strongly associated with reading ability. In a review of the literature on the reading skills of the deaf, Hanson (1991) found a strong relationship between deaf readers' use of a phonological coding system and their verbal memory span and reading ability. Those deaf individuals who, despite the lack of auditory input, had discovered the phonological structure of English were able to recall more on verbal memory tasks and to achieve higher reading skills.

Additional support for the role of verbal working memory in reading acquisition has also been obtained in a small number of longitudinal studies following children from
the pre-reading stage to early elementary years. These studies have found verbal working memory capacity to be significantly related to later reading acquisition (Mann, 1984; Mann & Liberman, 1984; Share, Jorm, MacClean, & Mathews, 1984). For example, Share et al. (1984) found a correlation of .40 between sentence memory tasks in kindergarten and reading level at the end of first grade. Similar correlations were also observed by Mann (1984) and Mann et al. (1984) between verbal memory span in kindergarten and first-grade reading ability. In addition, a follow-up review, at age 19, of learning-disabled students who participated in a series of experiments at ages nine to eleven found that those who continued to have reading problems at age 19 also continued to have deficits in verbal working memory (Torgesen, 1991). In contrast, those who had improved in reading ability also had made gains in memory capacity.

The studies discussed above indicate an association between verbal working memory and reading ability. However, there is controversy as to the nature and directionality of that association. Research to date has been unable to determine whether working memory problems are causal, contributory, or a consequence of reading problems (Brady, 1991; Pennington, Van Orden, Kirson & Haith, 1991). According to Pennington et al. (1991), existing data cannot demonstrate the universality of working memory deficits in
people with dyslexia, and thus cannot support memory deficits as a causal factor for reading disabilities. For example, Bradley and Bryant (1985), in a longitudinal study, were unable to consistently find a predictive relationship between working memory and reading skill while such a relationship was obtained between phoneme awareness and reading. In addition, Pennington (1991) challenges the association between memory deficits and decoding problems, arguing instead that individual differences in working memory may be more important for comprehension than for decoding or word recognition. A recent finding that memory was not significantly related to either decoding or reading comprehension when IQ was controlled further contributes to the debate about the unique role of working memory in reading (Evans, 1991).

To review, the bulk of the evidence points to a correspondence between reading ability and verbal memory skills. Numerous studies report that individuals with reading problems have shorter verbal memory spans than those who are better readers, apparently related to less efficient phonological coding. The evidence with deaf readers further supports an important link between verbal working memory processes and reading performance. Yet conflicting findings raise questions, as Pennington et al. (1991) note, about whether verbal working memory deficits are a cause or consequence of reading impairment. Further research is
needed to explore this issue and to clarify which aspects of reading performance (i.e., decoding and/or comprehension) may be affected by memory deficits.

Evidence that Good and Poor Readers Differ in Listening Comprehension: Relating Reading-Group Differences to Verbal Working Memory

As noted earlier, children with reading problems have demonstrated deficits on listening comprehension tasks under certain circumstances. These findings have been observed on tasks measuring the comprehension of single spoken sentences and of short passages read orally to the child. The co-occurrence of reading deficits and listening comprehension problems has been observed in both elementary-aged children and in adolescents (Reidlinger-Ryan & Shewan, 1984; Shankweiler, 1989).

The results of studies comparing the listening abilities of good and poor readers are inconsistent, however. The divergent findings may be the consequence of the different methods which have been used to measure listening comprehension ability. As Shankweiler and Crain (1986) point out, tests that use complex structures are often needed to demonstrate differences in listening comprehension.

Evidence supporting reading group differences in listening comprehension

Evidence for differences between good and poor readers
in listening comprehension has been found in studies using tasks that tax verbal working memory. For example, good readers do better on measures such as the Token Test, which requires the comprehension of sentence information that has minimal semantic content (e.g., Put the green square on the red circle) (Riedlinger-Ryan & Shewan, 1984; Smith, Mann, & Shankweiler, 1986; Torgesen, Rashotte, & Greenstein, 1988). The sections of the Token test that appear to discriminate between good and poor readers are those which place the greatest demands on working memory. In those sections, children are asked to respond to longer sets of directions. Reading group differences have not been found on the sections that are syntactically complex but shorter (Smith, Mann, & Shankweiler, 1986). Similarly, working memory deficits rather than syntactic deficiencies have been implicated in studies examining comprehension difficulties for sentences with relative clause phrases (Mann, Shankweiler, & Smith, 1984; Shankweiler & Crain, 1986).

Further evidence for listening comprehension differences between good and poor readers was found in a study which systematically manipulated the working memory demands of listening tasks by varying prosodic cues in sentences and phrases (e.g., pitch, timing, syllable stress) (Mann, Cowin, & Shoenhimer, 1989, 1990). Prosodic cues affect the memory demands of the task by varying how much of the sentence must be retained in order to recover the
grammatical structure and meaning of the sentence. Mann et al. compared good and poor readers in the second and fourth grades on their ability to comprehend phrases and sentences presented orally when the prosodic cues were varied. Poor readers consistently made more comprehension errors than the good readers even though the sentences used grammatic structures well within the grasp of the children. Mann et al. concluded that, "...poor readers tend to encounter phonological processing problems [in working memory] that lead them to misunderstand certain types of [orally presented] phrases and sentences" (p. 86).

The comprehension of orally presented material by adults has also been associated with memory capacity (Daneman & Carpenter, 1980). In a study using college students who had a range of reading and listening abilities, Daneman and Carpenter (1980) found significant correlations between memory span and oral language comprehension skills (between .71 and .85). Importantly, the errors made by the students with shorter memory spans reflected qualitative differences in comprehension, suggesting that memory deficits may have limited their ability to understand the meanings of passages.

Listening comprehension deficits in children with reading problems have also been seen on more conventional tasks that require listening to passages and completing follow-up comprehension questions. In one study, good and
poor fifth-grade readers were compared using a reading/listening measure drawn from supplementary reading texts at an "intermediate level" (Berger, 1978). The students were asked to retell stories (either heard or read) in their own words and to answer literal comprehension questions about the passages. Both the good and poor readers performed better on the listening than on the reading comprehension measures. However, the good readers significantly outperformed the poor readers on the listening comprehension tasks, in both the retell and literal question conditions. In a second study (Curtis, 1980), good readers and poor readers in the second, third, and fifth grades were compared on the Diagnostic Reading Scales (Spache, 1981). Again, while all children achieved higher listening than reading levels, the good readers from all grades significantly outperformed poor readers on both the reading and listening comprehension tasks. Similar results were also found in a third study, in which normal and reading-disabled students, aged 9-11, were compared on the Durell Listening-Reading Series, Intermediate level (Wood, Buckhalt, & Tomlin, 1988).

Using a sentence verification task, Carlisle (1989b) obtained further evidence of both listening and reading comprehension differences between good and poor readers. Seventh-grade good and poor readers were presented with passages at third-, fifth-, seventh-, and ninth-grade
levels, and were required to indicate whether the ideas in
the follow-up sentences were contained in the passages. The
overall group effect on both the listening and reading
comprehension subtests was significant.

Evidence against listening comprehension differences

In contrast to the studies discussed above, studies
which have not found differences between good and poor
readers on listening comprehension tasks have often used
measures in which memory demands are deliberately kept to a
minimum, or in which the materials are short and/or highly
structured. For example, a modified version of the Peabody
Individual Achievement Test was used to assess reading and
listening comprehension skills of fifth- and sixth-grade
good and poor readers (Spring & French, 1990). The
investigators reported that they purposefully modified the
test to keep working memory demands low, in order to be able
to evaluate comprehension when memory was not a factor. No
significant group differences were found. In a second
study, short passages were used to evaluate reading and
listening ability in sixth-grade good and poor readers
(Horowitz & Samuels, 1985). Both easy (fourth/fifth-grade
levels) and more difficult (seventh/eighth-grade levels)
passages were utilized. The students were to retell what
they had read or heard in their own words. While expected
reading group differences were found on the reading
comprehension tasks, no significant group differences in listening comprehension were found for either passage. The researchers concluded that the language problems experienced by poor readers are probably unique to reading. However, they proposed that future research should consider using passages of varied lengths.

Between 1979 and 1982, Torgesen and his colleagues administered three listening comprehension tasks as part of a larger study of learning-disabled children (aged 9-11) (Torgesen, 1988, 1990; Torgesen, Rachotte, & Greenstein, 1988). Three groups of subjects were tested: learning-disabled students with known memory difficulties (who were reading disabled), learning-disabled students without memory deficits, and nondisabled students. In the first listening comprehension task, students were required to listen to third-grade narrative passages and to complete cloze and probe memory tasks. In the second measure, students listened to highly structured folktales with a fifth-grade readability level, and were asked to recall the stories in their own words. The investigators stressed that they attempted to avoid material that would place particular or unusual stress on the role of working memory in comprehension. Instead, the first two experiments involved listening to meaningful and very organized materials that were relatively short and highly structured. The two tasks required the students to listen to grade-level material that
may have easily fallen within their expected listening ranges, even for the poor readers. Interestingly, no significant differences between groups were found on either of these two measures.

The third measure employed by Torgesen et al. had students listen and respond to complex directions similar to those of the Token Test, a task chosen to place greater stress on working memory. As in other studies which used the Token Test, the poor readers demonstrated significant deficits in the ability to follow complex directions when compared to the good readers.

In the research reviewed above, it appears that when the working memory demands of listening tasks are reduced or minimized, there are few differences in listening comprehension between good and poor readers. However, when the working memory demands are greater, differences in listening comprehension emerge. Since the association between reading performance and listening comprehension ability seems to depend on the memory requirements of the listening measures, the origin of comprehension problem may stem, at least in part, from memory factors rather than solely from limits in syntactic, inferential, or knowledge factors.

Recall and Recognition

The cognitive demands of memory tasks requiring the
recognition of information are considered different from those requiring the recall of information (Flavell, 1985).

Recognition is defined by Gardiner and Parkin (1990) as what happens when an individual identifies a stimulus as having been encountered previously. This is in contrast to recall memory where the reproduction of previously learned material is required. According to Dyne, Humphries, Bain, and Pike (1990), there are fundamental differences in the memory access processes underlying recognition and recall. They propose that a global matching operation which matches a cue or pair of cues against memories underlies recognition, while a more specific retrieval operation underlies the recovery of information in recall.

In studies comparing learning-disabled and nondisabled students on different memory tasks, disabled students consistently perform better on tasks requiring recognition memory than they do on measures requiring the recall of verbatim information (Torgesen, 1988; Weinberg, McLean, Snider, & Rintelmann, 1989). Recently, Lorsbach and Worman (1990) explored whether learning-disabled and nondisabled students differed on the two types of memory tasks. The first was a cued recall task, while the second used a primed recognition paradigm. The purpose of the study was to tease apart whether difficulties on recall tasks were due to the students' inability to form associations with new material or their inability to access the information. Lorsbach and
Worman (1990) hypothesized that if the two groups were both making associations, they would perform equivalently on the recognition task. If the learning-disabled students were not forming associations, they would do as poorly on the recognition as on the recall task. The results indicated that learning-disabled students performed less well than the nondisabled students only on the recall measure. This outcome suggests "that learning-disabled children may be forming associations at a rate that is comparable to that of nondisabled children. Difficulties only emerge when a task requires the explicit, conscious reinstatement of those newly formed associations" (Lorsbach & Worman, 1990, p.99).

Similar results were found in a study that required good and poor readers to read or to listen to a story read out loud and to answer either open-ended recall or multiple choice recognition questions about the passage (Weinberg, McLean, Snider, & Rintelmann, 1989). In this study, the Gilmore-Oral Reading Test was individually administered to each child in one of four ways: the child read the passage out loud, then answered recall questions asked by the examiner; the child read the passage out loud, then selected a correct answer from a multiple choice provided by the examiner; the child listened as the examiner read the passage, then answered recall questions; the child listened as the examiner read the passage, then selected an answer from the multiple choice provided by the examiner. There
were no differences among the good readers for mode of input (whether the child read or listened to the text) or for questioning method (recognition vs. recall). Nor were there any differences between the good and poor readers who listened to the passage and responded to the multiple choice questions. However, on the listening tasks, the poor readers who answered multiple choice questions significantly outperformed the poor readers who answered recall questions. Thus, poor readers were able to demonstrate adequate comprehension when given tasks that demanded recognition rather than recall memory.

In discussing a series of experiments conducted with learning-disabled students, Torgesen (1988) also concluded that the ability to recognize previously learned material was not impaired in subjects who had shown deficits on recall tasks.

The research reviewed above suggests that recognition and recall memory tasks make different demands upon cognitive processes, and that learning-disabled students tend to perform as well as nondisabled students only on recognition measures. The research implies that while both groups of students can make appropriate semantic associations, the learning-disabled students may have difficulty accessing specific information in memory without prompts provided by a recognition task.
Summary

In sum, the literature points to a strong association between reading ability and performance on working memory tasks. Questions remain about the causal/facilitory nature of the association, and about whether memory deficits impact upon both decoding and comprehension. In addition, poor readers are found to have lower performance on listening comprehension tasks if the material stresses memory (e.g., if it is longer or more complex) and if measures require more of memory (e.g., recall as opposed to recognition). The Processing Limitation Hypothesis proposes that at least one source of individual differences that affects both reading and listening is verbal working memory skills. Yet further work is necessary to explicate the relationship between verbal working memory ability and listening comprehension. As students become older and the listening demands of school increase in length and complexity, the negative affects of working memory deficits on listening may became more pronounced. More research is needed to determine whether students with reduced verbal working memory capacity are at academic risk because of possible listening comprehension problems.
Present Investigation

The previous review indicates that there are several unanswered questions concerning the nature of the association between verbal working memory and reading/listening abilities. Are working memory skills associated with decoding and/or reading comprehension performance? Do children with working memory deficits experience listening comprehension problems? Does listening comprehension accuracy vary according to the memory demands of listening tasks?

Using the Processing Limitation Hypothesis as a conceptual framework, the present study investigates the link between verbal working memory and both written and oral language skills. Specifically, the research asks: given fifth-grade students who are divided into low, middle, and high verbal working memory groups, will their decoding, reading comprehension, and listening comprehension scores correspond to their memory performance? Two listening comprehension measures are used to vary the memory demands of the tasks. The first measure requires the recognition of information, while the second asks students for the verbatim recall of information.

The first hypothesis is that children in the low memory group cannot complete listening comprehension tasks with the same degree of accuracy as children in the middle and high memory groups. The second hypothesis is that students
in the low memory group will show decoding and reading comprehension deficits when compared to children in the middle and high memory groups.

An additional purpose of the present study is to investigate the power of listening comprehension, decoding ability, memory skills, and IQ to predict reading comprehension. Based on a model proposed by Gough and Tunmer that reading comprehension is the product of listening comprehension and decoding, the third hypothesis is that a significant proportion of the variance in reading comprehension scores can be accounted for by listening comprehension and decoding ability.

The predictions are:
1) There will be significantly different performances between the memory groups on the listening comprehension, reading comprehension, and decoding tasks.
2) There will be an interaction between memory ability and levels of passage difficulty on the listening comprehension tasks.
3) There will be an interaction between memory ability and levels of passage difficulty on the reading comprehension task.
4) Listening comprehension and decoding will together account for a greater proportion of the variance in reading comprehension than will all other combinations of listening comprehension, decoding, memory skills, and IQ.
Method

Subjects

A total of 165 fifth-grade students from twelve classes in three school districts voluntarily participated in the study. The districts were in rural/suburban areas of central and southern Rhode Island, and represented mixed socio/economic and educational backgrounds.

Letters of informed consent were distributed to all students in participating classrooms by the researcher and the students' teachers. All students who returned the consent forms with positive responses were included in the study. Out of the 165 participating students, 136 met the eligibility or selection criteria and the data from these 136 students were included in the analyses. Thirteen, or approximately 9%, of these students were identified by their local school systems as Learning Disabled, and were receiving special education resource services.

The selection criteria were:

a. **Enrollment in the fifth grade, and between 9-0 and 11-11 years old.** Fifth graders were chosen for three reasons. First, fifth-grade classwork places greater demands upon listening skills than the material in the earlier grades. Therefore potential deficits in listening comprehension may have more serious consequences at this level. The increased number of children referred for learning disabilities in the fourth and fifth grades may reflect, in
part, problems resulting from poor listening skills. Second, a number of prior studies researching listening comprehension have used children in the fifth grade, who have comparable ages. Therefore, it will be easier to compare the results of the present study with those from previous studies. Third, the range of reading/listening abilities in the fifth grade is broader than in the earlier grades. There should not be a floor effect due to lack of instructional opportunity.

b. **Full-Scale IQ score, as measured by a short form of the Wechsler Intelligence Scales for Children-Revised (WISC-R), (Wechsler, 1974) within the Average to High Average ranges (between 90-119).** An IQ requirement was used to control for intellectual differences among the students. The Vocabulary and Block Design Subtests of the WISC-R were administered to each child. This dyad has a validity coefficient of .91 (N=2200), representing its correlation with the complete WISC-R battery Full-Scale score (Sattler, 1988). The scores from these two subtests were converted to an estimated Full-Scale IQ score using procedures described in Sattler (1988). (While IQ was treated as a composite score, certain analyses examined the effects of the Vocabulary and Block Design subtests separately).
c. Passing a hearing screening and a school vision screening. In order to rule out listening and reading difficulties due to outstanding sensory deficits, all students were screened for uncorrected sensory problems. Each student was given a hearing screening with an audiometer. Right and left ears were tested with tones at five different frequencies: 500 Hz (25dB), 1000 Hz (20dB), 2000 Hz (20dB), 4000 Hz (20dB), and 8000 Hz (20dB). Visual acuity was determined from the results of school administered visual screenings.

d. English used as the primary language spoken at home. In order to attempt a control for the effects of background experience in language performance, students needed to come from homes in which English was the primary language spoken. In addition, a cognitive assessment of children for whom English is a second language would be unreliable. English usage was assessed informally by talking with the students and asking them about languages used in the home.

Measures

Memory Measures

The memory measures were not standardized tests, but were research measures designed to tap hypothesized components of working memory (i.e., storage and processing). Two of the measures (Word Span and Pseudoword Repetition)
were ones that have been used in earlier studies. The third (Sentence Span) was a modification for children of a procedure that had previously been used with adults.

1. **Familiar Word Span Memory Test** (Brady, Shankweiler, & Mann, 1983). This span measure was included to provide a measure of storage capacity. The test contained 10 strings of five monosyllabic words. A trial consisted of hearing a sequence of five words presented at one second intervals, and attempting to repeat back the list of words in the correct order. Responses were scored in two ways: the number of words recalled correctly in the correct position (referred to as Word Span Correct Order), and the number of words correctly recalled regardless of order (referred to as Word Span Free Recall). The students were given as much time as needed to respond. The lists of words were presented on tape to assure consistency of presentation. The students' responses were recorded to insure scoring accuracy. See Appendix A for a copy of the Word Span protocol.

   The split-half reliability coefficient, for this sample, was equal to .84 (N=131).

2. **Pseudoword Repetition Test** (Gathercole & Baddeley, 1989). This measure was given to assess the ability to hold phonological representations in storage. This test contains 40 pseudowords, from two to five syllables in length, which conform to the dominant prosodic constraints of spoken
English (Gathercole & Baddeley, 1989). The words were tape recorded in the order presented on the Pseudoword Repetition Test protocol, with a five second interval between the words.

The students were told that they would hear a list of nonsense words presented one at a time, and were asked to repeat each word as soon as it was heard. The responses were scored as correct or incorrect. All responses were taped to again insure scoring accuracy. While there was no time limit for responses, students were expected to respond within the interval between words. See Appendix B for a copy of the Pseudoword Repetition Test protocol.

Gutman split-half reliability and Equal Length Spearman-Brown reliability coefficients for this sample were both equal to .80 (N=131).

3. Sentence Span Memory Test (Daneman & Carpenter, 1980; Swanson, Cochran, & Ewers, 1990; Turner & Engel, 1989). The sentence span test was originally developed by Daneman and Carpenter (1980) to address both the processing and storage components of verbal working memory. Daneman and Carpenter (1980) found that it correlated significantly with both reading comprehension (between .72 and .86) and listening comprehension (between .71 and .85). Versions of the Sentence Span Test successfully discriminated between
learning-disabled and nondisabled fourth-, fifth-, and sixth-grade students (Swanson, Cochran, & Ewers, 1990), and correlated significantly with the more traditional digit span tasks (Daneman & Carpenter, 1980; Dixon, Le Fevre, & Twilley, 1988).

Following the procedures described in Daneman and Carpenter (1980), Dixon et al. (1988), and Swanson et al. (1990), a variation of the Sentence Span Test was developed for use in the present study. Forty-two unrelated sentences were constructed using third- and fourth-grade social studies and science curricular materials. These curricula were chosen because it was thought that sentences constructed from that content would be conceptually easy for average IQ fifth graders to understand. The sentences ranged in length from five to seven words, with a mean length of six words, and contained between six to nine syllables with a mean of 7.53 syllables. The sentences were arranged in three sets of two sentences, three sets of three sentences, three sets of four sentences, and three sets of five sentences, with twelve sets of sentences in all. Within each set, the sentences were recorded with a three second interval between them. Three seconds after the last sentence in each set, a tone was recorded to signify the end of the set. The children knew in advance how many sentences would be in a set.

The students listened to the sets of sentences read out
loud. After each sentence in a set, they said whether the sentence was true or false. After the students heard all of the sentences in a set and the tone, they repeated back the last word occurring in each of the sentences in the set. Final words did not need to be repeated back in the order they had been presented. Responses were hand-recorded on an answer sheet. The child's score was the total number of words correctly recalled out of 42 items. Because the purpose of the true/false component of the task was to create a processing component, the accuracy of the judgements was not important and incorrect true/false answers were ignored. See Appendix C for a list of the Sentence Span sentences and a copy of the scoring protocol.

The task was piloted in a fifth-grade class from a participating school system one month prior to the beginning of the study. Twenty-five students were tested in the pilot phase. Their scores ranged from fifteen to twenty-nine, with a mean of 26.72 and a standard deviation of 5.90. The results approximated the normal or expected distribution of scores.

The reliability of this task could not be assessed using a split-half method or an odd-even method because of the graduating difficulty of the task and the odd number of sets (i.e., three per number of sentences). The use of a test-retest or alternative forms method was not possible at this time. Therefore, no reliability data are available for
Decoding Measure

1. Woodcock Reading Mastery Test-Revised: Word Attack Subtest (Woodcock, 1987). This standardized test of decoding skills requires students to read phonetically-regular nonsense words (e.g., tiff, jox). There are 45 items, and the students' raw scores equaled the number of items read correctly. Raw scores were then converted into standard scores using the tables provided with the test.

Internal consistency reliability was calculated by Woodcock using split/half procedures and was corrected for length with the Spearman-Brown formula. The split/half coefficients for grade five are between .89 and .94.

Content Validity was determined through logical evidence, by examination of the scope and sequence of the test items, and the use of independent curriculum experts for the test development. The concurrent validity coefficient for grade five, when compared to the word attack subtest of the Woodcock-Johnson Tests is .90.

Listening and Reading Comprehension Measures

1. Profiles in Listening and Reading (PILAR) (Carlisle, 1987). An experimental test, PILAR, was designed by Carlisle (1987) to measure the comprehension through listening and reading of extended discourse. During the
construction of PILAR, the reading and listening passages were matched at grade level on vocabulary difficulty, sentence length, and through the use of four readability formulas. The test uses a sentence verification technique (SVT) in order to measure language processing abilities. The aim of this procedure is to reduce the potential confound of reasoning and test-taking skills. PILAR was designed to place similar memory demands on the reading comprehension tasks as on the listening comprehension tasks. During the listening portion of the test used in the present study, the students listened to three taped passages presented one at a time. Following each passage, the students heard twelve taped statements. Their task was to decide whether the idea in each statement was or was not contained in the passage just heard. The students responded by circling "yes" or "no" on an answer sheet. During the reading portion of the test, the students read three passages silently to themselves, one at a time. After the students completed reading each passage, the passage was removed. The students then read twelve statements about the passage and followed the same procedures as described for the listening portion. The students read and listened to third-, fifth-, and seventh-grade passages. The follow-up statements that the students listened to or read contained paraphrases of sentences from the original stories, verbatim sentences from the stories, information that was true about
the stories' subject but not included in the stories, and false information. See Appendix D for sample PILAR listening and reading stories, accompanying test sentences, and a copy of the scoring protocol for the listening comprehension measures.

The sentence verification technique used in PILAR has been shown to be a reliable and valid measure of comprehension in normal readers (Royer et al., 1986). In one study, corrected split-half reliabilities of $r=.85$ and $r=.71$ were obtained for reading and listening passages, respectively. In a second study, the mean coefficient of internal consistency for both reading and listening passages, was .92 (Carlisle, 1989). Validity studies found that PILAR successfully discriminated between learning-disabled and nondisabled students (Carlisle, 1989).

2. Diagnostic Reading Scales: Listening Comprehension Stories (Spache, 1981). The Diagnostic Reading Scales is a standardized individually administered reading test which includes reading and listening comprehension measures, a graded sight word list, and word attack measures. For the purposes of the study, only the listening comprehension portion of the test was used. The students listened to three stories read aloud on a tape (one third-, one fifth-, and one seventh-grade passage). The stories were taped by the examiner to insure consistency in presentation. Following
each story, the examiner asked the students eight factual recall questions about the story. The students' responses were scored as correct or incorrect following the test-administration directions. According to the test instructions, the students must answer at least five out of the eight questions correctly to pass a given passage.

The Spache technical manual reports that Pearson product-moment correlation coefficients of .89 and .99 were obtained in two studies measuring alternate form reliability. Test-retest reliability coefficients of stability were found to be between .84 and .88. Concurrent validity was established by comparing the Spache to other standardized reading measures. The Spearman rank-difference order correlation coefficients ranged from .90-.92. In order to establish content validity, the reading sections were chosen to be similar to reading materials in use in schools, and the passage grade levels were determined by readability formulas and from the results of student performance.

**Experimental Procedures**

Following the administration of the selection criteria measures, each student was presented with the six tasks. The testing was divided into three testing sessions because of the large number of measures administered to each child.
Each session lasted approximately 20-30 minutes.

During the first session, each student was individually administered the Short Form of the WISC-R, the hearing screening, and the Word Span Memory Test.

During the second session, each subject was individually given the Pseudoword Repetition Task, the Word Attack Subtest, and the Sentence Span Test. These tests were presented in the following order: a memory task, the decoding measure, and then a second memory task. The order in which the two memory measures were presented was alternated to attempt to control for effects due to the order of presentation (i.e., fatigue).

During session three, two subjects at a time were administered the PILAR stories. The two participants sat in a small testing room with their backs to each other.

Following the completion of the three sessions, a subset of 76 students was individually administered the listening comprehension measures from the Diagnostic Reading Scales. These students were chosen because of the geographical convenience of their schools, the willingness of their teachers to continue to participate in the study, and the students' scores on the memory measures. This subset of students was representative of the sample on all of the selection criteria measures.
Results

The test results were analyzed in the following sequence. The data from all 136 students who met the selection criteria were summarized for descriptive statistics. A correlation matrix ascertaining the relationships between the measures was then computed. Since the three memory measures correlated significantly with each other (p<.001), the memory raw scores were converted to z scores and, for each student, were averaged to form one composite z score. The composite memory scores were ranked, and divided into the lowest third, the middle third, and the highest third. From these divisions, three groups were formed: a low memory group (N=45), a middle memory group (N=46), and a high memory group (N=45). Comparisons of means for the PILAR Listening and Reading Comprehension measures and for the decoding test were conducted with these three groups. As noted earlier, a subsample of the total sample (N=76) was given the Spache Listening Comprehension Scales in addition to the other measures. The subsample was subdivided into a low memory group (N=30), a middle memory group (N=16), and high memory group (N=30) according to their status in the total sample. However, comparison of group means were conducted using the low and high memory groups only. Following the completion of the Analyses of Variance, regression analyses were conducted on the entire data set to investigate the extent to which listening
comprehension, memory skills, IQ, and decoding ability predict performance on the reading comprehension measure.

Characteristics of the Total Sample

The descriptive statistics for the two selection criteria, age and IQ, for the sample of 136 subjects and for the subsample of 76 students are described in Table 1. The descriptive characteristics for the subsample of 76 students were similar to those of the larger sample of 136 students on these two variables. (See Appendices E and F for tables displaying means, standard deviations, and ranges of the dependent measures for the total set of 136 subjects and for the subset of 76 students, respectively).

Table 1

Means, Standard Deviations, and Ranges of Age and IQ for the Total Experimental Sample and the Subsample

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th></th>
<th>IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Range</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Experimental Sample,</td>
<td>10.2 (.49)</td>
<td>9.06-11.09</td>
<td>106.64 (7.82)</td>
</tr>
<tr>
<td>N=136</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subsample, N=76</td>
<td>10.3 (.53)</td>
<td>9.10-11.09</td>
<td>106.55 (8.03)</td>
</tr>
</tbody>
</table>
Characteristics of the Three Memory Groups for the Full Sample (N=136)

The Word Span test produced two scores, Correct Order and Free Recall. The Word Span Correct Order scores were chosen in forming the composite memory scores because they correlated more highly with the Pseudoword Repetition and Sentence Span tasks than did the Free Recall Scores (see Table 2). The $z$ scores of the Word Span Correct Order, Pseudoword Repetition, and Sentence Span tasks were averaged to form one composite memory score. As noted earlier, the composite memory scores were ranked and subdivided into memory groups.

As stated in the methods section, thirteen students had been identified by their school systems as learning disabled. Nine of the students fell into the low memory group, three were in the middle memory group, and one student was ranked in the high memory group.

The descriptive statistics for each memory group for the selection criteria are listed in Table 3. Included also are summaries of one-way ANOVAs conducted to evaluate differences between the group means. The groups did not differ significantly in age but did differ in IQ. (The significant group differences on the IQ measures will be further discussed below.)

Table 4 describes the performance of the memory groups on the memory tasks which contributed to the composite
memory score. Significant differences were found between all three memory groups on the Sentence Span task, $F(2,133)=41.91, p<.001$, and on the Word Span task, $F(2,133)=124.58, p<.001$. While significant group differences were also observed on the Pseudoword Repetition task, $F(2,133)=17.22, p<.001$, the differences existed only between the low and the middle and high groups. Further analysis determined that the Pseudoword Repetition task was negatively skewed ($-4.0528; \text{s.e.skew}=.208$), and significantly different from 0 ($z=21.7692$) (Tabachnick & Fidell, 1983). A ceiling effect may have been responsible for this observed distribution of scores.
Table 2: Correlational Matrix of Age, IQ, Memory Measures, PILAR Listening and Reading Comprehension tasks, and Decoding; N=136

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. IQ</td>
<td>-.28**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Vocab.</td>
<td>-.42**</td>
<td>.57**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Block Design</td>
<td>.002</td>
<td>.75**</td>
<td>-.12</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Word Span Order</td>
<td>-.003</td>
<td>.22*</td>
<td>.16</td>
<td>.14</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Pseudoword Span</td>
<td>.07</td>
<td>.31**</td>
<td>.24*</td>
<td>.17</td>
<td>.36**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Sentence Span</td>
<td>.10</td>
<td>.27**</td>
<td>.24*</td>
<td>.14</td>
<td>.38**</td>
<td>.32**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Composite Mem.</td>
<td>.08</td>
<td>.35**</td>
<td>.28**</td>
<td>.20</td>
<td>.77**</td>
<td>.75**</td>
<td>.75**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. PILAR Total Listening</td>
<td>-.03</td>
<td>.39**</td>
<td>.27**</td>
<td>.26*</td>
<td>-.002</td>
<td>.06</td>
<td>.01</td>
<td>.03</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. PILAR Total Reading</td>
<td>-.16</td>
<td>.42**</td>
<td>.34**</td>
<td>.24*</td>
<td>.19</td>
<td>.32**</td>
<td>.25*</td>
<td>.34**</td>
<td>.42**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>11. Word Attack standard score</td>
<td>-.35**</td>
<td>.31**</td>
<td>.47**</td>
<td>.002</td>
<td>.35**</td>
<td>.31**</td>
<td>.23*</td>
<td>.39**</td>
<td>.01</td>
<td>.22*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*p<.01
**p<.001
Table 3

Means and Standard Deviations for Age and IQ for Each Memory Group, and Analyses of Group Differences with Follow-up Tukey Results, N=136

<table>
<thead>
<tr>
<th></th>
<th>Low Memory Mean (SD) N=45</th>
<th>Middle Memory Mean (SD) N=46</th>
<th>High Memory Mean (SD) N=45</th>
<th>F (df)</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>10.19 (.44)</td>
<td>10.23 (.52)</td>
<td>10.16 (.44)</td>
<td>.21</td>
<td>a</td>
</tr>
<tr>
<td>IQ</td>
<td>103.27 (7.98)</td>
<td>106.52 (7.55)</td>
<td>110.13 (6.44)</td>
<td>9.82</td>
<td>b</td>
</tr>
<tr>
<td>Vocab.</td>
<td>11.36 (1.61)</td>
<td>11.63 (1.64)</td>
<td>12.53 (1.76)</td>
<td>6.11</td>
<td>c</td>
</tr>
<tr>
<td>Block</td>
<td>9.73 (2.32)</td>
<td>10.54 (2.05)</td>
<td>10.84 (1.99)</td>
<td>3.29</td>
<td>b</td>
</tr>
<tr>
<td>Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subtest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a  
no significance between groups

b  
significance between the Low and High groups only, p<.05

c  
significance between the High and Low and the High and Middle groups only, p <.05
Table 4

Means and Standard Deviations for the Working Memory Tasks for Each Memory Group, and Analyses of Group Differences with Follow-up Tukey Results, N=136

<table>
<thead>
<tr>
<th>Task</th>
<th>Low Memory Mean (SD)</th>
<th>Middle Memory Mean (SD)</th>
<th>High Memory Mean (SD)</th>
<th>F (df)</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Span</td>
<td>16.42 (8.17)</td>
<td>26.15 (5.64)</td>
<td>38.29 (5.64)</td>
<td>124.58</td>
<td>a .001</td>
</tr>
<tr>
<td>Pseudo-Word Accuracy</td>
<td>32.58 (5.62)</td>
<td>35.70 (2.52)</td>
<td>37.13 (2.19)</td>
<td>17.22</td>
<td>b</td>
</tr>
<tr>
<td>Sentence Span</td>
<td>22.53 (4.36)</td>
<td>26.37 (3.56)</td>
<td>30.18 (3.98)</td>
<td>41.91</td>
<td>a .001</td>
</tr>
</tbody>
</table>

a  significance between all three groups, p<.05

b  significance between the High and Low, and the Middle and Low groups, p<.05
Correlations Between Variables

Performance on the PILAR listening comprehension test (the listening measure which required recognition memory skills) significantly correlated with scores on the IQ measure \(r = .39, p < .001\), with the PILAR reading comprehension test \(r = .42, p < .001\), and with the Spache listening test \(r = .61, p < .001\). It did not correlate significantly with the decoding measure. Of particular interest, it did not correlate significantly with performance on any of the three working memory tasks. In fact, the composite memory score (the memory measures averaged together) and the PILAR listening scores had a correlation of \(r = .03\). (See Table 2).

As on the PILAR listening task, scores on the Spache listening test (the listening measure which required recall memory skills) correlated significantly with the scores on the IQ estimate \(r = .40, p < .001\) and with the PILAR reading measure \(r = .56, p < .001\). (See Appendix G for a correlation matrix that includes the Spache listening test). Unlike the PILAR, however, performance on the Spache also correlated significantly with the composite memory values \(r = .29, p < .01\). While the correlation is modest, it supports the hypothesis that a listening task more demanding in memory would yield a significant relationship with the memory measures. Further investigation of the relationship between the Spache and the individual memory measures found that the
Spache was significantly correlated with the Sentence Span Task ($r=.29, p<.01$), a task designed to tap both the storage and the processing components of working memory, but was not significantly correlated with either the Word Span ($r=.19$) or the Pseudoword Repetition ($r=.20$) Tests, tasks which mostly reflect memory encoding and storage. As found with the PILAR listening measure, there was no significant relationship between the Spache and the measure of decoding skills ($r=.11$).

The decoding measure correlated significantly with IQ ($r=.31$, $p<.001$), with reading comprehension ($r=.22$, $p<.01$), and with the verbal working memory composite ($r=.39$, $p<.001$). As stated above, no significant correlations were found between the decoding test and either of the two listening comprehension tasks.

The reading comprehension task correlated significantly with IQ ($r=.42$, $p<.001$), as well as with the listening and decoding measures described above. As predicted, reading comprehension also correlated significantly with the composite memory score ($r=.34$, $p<.001$). Further examination of the relationships between the reading comprehension task and the individual memory measures found significant correlations with both the Pseudoword Repetition ($r=.32$, $p<.001$) and the Sentence Span ($r=.25$, $p<.01$) tasks, but not with the Word Span test ($r=.19$).
ANOVA: Listening Comprehension

To review, two different listening comprehension measures were administered in order to vary the memory demands of the listening tasks. One emphasized recognition memory (PILAR) and the other recall memory skills (Spache). Within each measure, different grade levels of material were used to vary the level of passage difficulty.

The descriptive statistics for performance on the PILAR are presented in Table 5. A 3x3 mixed repeated measures ANOVA using the comprehension scores on the PILAR found no main effects for memory group and no significant interaction between the groups and the different grade level passages (see Table 6). All three memory groups performed similarly on the listening passages (see Figure 1). The original prediction of finding memory group differences for the higher grade level material was not confirmed.
Table 5

Means and Standard Deviations for Three Memory Groups on the Three Grade Levels of Passage Difficulty of the PILAR Listening Comprehension Task, N=136

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Low Memory</th>
<th>Middle Memory</th>
<th>High Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td></td>
<td>N=45</td>
<td>N=46</td>
<td>N=45</td>
</tr>
<tr>
<td>Third Grade</td>
<td>10.24 (1.55)</td>
<td>10.61 (1.29)</td>
<td>10.42 (1.59)</td>
</tr>
<tr>
<td>Fifth Grade</td>
<td>10.20 (1.61)</td>
<td>10.43 (1.42)</td>
<td>10.44 (1.24)</td>
</tr>
<tr>
<td>Seventh Grade</td>
<td>9.40 (1.78)</td>
<td>9.30 (1.99)</td>
<td>9.82 (1.42)</td>
</tr>
</tbody>
</table>

Table 6

Summary Table for 3 x 3 ANOVA for PILAR Listening Comprehension Scores by Memory Group and Levels of Passage Difficulty, N=136

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sign. of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Subjects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory Group</td>
<td>5.41</td>
<td>2</td>
<td>2.71</td>
<td>.63</td>
<td>.536</td>
</tr>
<tr>
<td>Within cells</td>
<td>575.33</td>
<td>133</td>
<td>4.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Subjects:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty level</td>
<td>71.07</td>
<td>2</td>
<td>35.53</td>
<td>23.89</td>
<td>.001</td>
</tr>
<tr>
<td>Group x Level</td>
<td>6.21</td>
<td>4</td>
<td>1.55</td>
<td>1.04</td>
<td>.385</td>
</tr>
<tr>
<td>Within cells</td>
<td>395.65</td>
<td>266</td>
<td>1.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A second analysis of the PILAR data was conducted using only the results of the subset of subjects who were chosen to receive the Spache listening comprehension measure (see results below), in order to permit a direct comparison of performance on the Spache and PILAR listening measures. A 2x3 mixed repeated measures ANOVA was completed using the PILAR listening scores for the children in the high and low memory groups who were given the Spache (N=60). The results of this analysis were virtually the same as the analysis of the PILAR listening measures which used scores from the total sample. Again, there was no main effect for memory group membership and no significant interaction between the
groups and levels of passage difficulty. (See Table 7 for the means and standard deviations of the two groups of students on each level of passage difficulty and Table 8 for the results of the 2x3 ANOVA.) The subsets of students from the low and high memory groups were able to complete PILAR listening measures with similar degrees of accuracy (see Figure 2).

Table 7

Means and Standard Deviations for Two Memory Groups on the Three Grade Levels of Passage Difficulty of the PILAR Listening Comprehension Tasks, N=60

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Low Memory Mean (SD) N=30</th>
<th>High Memory Mean (SD) N=30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third Grade</td>
<td>10.53 (1.38)</td>
<td>10.63 (1.56)</td>
</tr>
<tr>
<td>Fifth Grade</td>
<td>10.13 (1.85)</td>
<td>10.46 (1.31)</td>
</tr>
<tr>
<td>Seventh Grade</td>
<td>9.20 (1.85)</td>
<td>9.67 (1.52)</td>
</tr>
</tbody>
</table>
Table 8
Summary table for 2x3 ANOVA for PILAR Listening Comprehension Scores by Memory Group and Levels of Passage Difficulty, N=60

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sign. of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory Group</td>
<td>4.05</td>
<td>1</td>
<td>4.05</td>
<td>.82</td>
<td>.370</td>
</tr>
<tr>
<td>Within Cells</td>
<td>287.61</td>
<td>58</td>
<td>4.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty level</td>
<td>43.08</td>
<td>2</td>
<td>21.54</td>
<td>16.31</td>
<td>.001</td>
</tr>
<tr>
<td>Group x Listen</td>
<td>1.03</td>
<td>2</td>
<td>.52</td>
<td>.39</td>
<td>.677</td>
</tr>
<tr>
<td>Within Cells</td>
<td>153.22</td>
<td>116</td>
<td>1.32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Mean correct responses, in raw scores, for the low and high memory groups on the PILAR listening comprehension measures, N=60
The descriptive statistics for the high and low memory group subsets given the Spache listening comprehension measure are presented in Table 9. A 2x3 mixed repeated measures ANOVA was used to evaluate the performance of students from the low and high memory groups on the three levels of the Spache. On this recall task, a significant main effect for memory group was obtained. As on the PILAR, however, a significant interaction was not observed between the memory groups and level of passage difficulty (see Table 10). While the differences between the groups widened as the passage level increased, demonstrating a trend toward an interaction (see Figure 3), more errors were observed for both groups on the higher grade-level passages. Thus, the results on this measure provided mixed support for the hypothesis that children with poorer memory skills would be less able to comprehend material presented orally than students with stronger memory skills: The more demanding recall task does reveal listening comprehension differences between children who differ in working memory ability. However, manipulations of passage difficulty did not differentiate between the memory groups.
Table 9

Means and Standard Deviations for Two Memory Groups on the Three Grade Levels of Passage Difficulty of the Spache Listening Comprehension Task, N=60

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Low Memory Mean (SD) N=30</th>
<th>High Memory Mean (SD) N=30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third Grade</td>
<td>6.98 (1.54)</td>
<td>7.00 (1.47)</td>
</tr>
<tr>
<td>Fifth Grade</td>
<td>4.58 (2.10)</td>
<td>5.52 (1.96)</td>
</tr>
<tr>
<td>Seventh Grade</td>
<td>4.16 (1.93)</td>
<td>5.30 (1.56)</td>
</tr>
</tbody>
</table>

Table 10

Summary Table for 2x3 ANOVA for the Spache Diagnostic Scales Scores by Memory Groups and Levels of Passage Difficulty, N=60

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sign. of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory Group</td>
<td>21.70</td>
<td>1</td>
<td>21.70</td>
<td>4.01</td>
<td>.05</td>
</tr>
<tr>
<td>Within cells</td>
<td>313.87</td>
<td>58</td>
<td>5.32</td>
<td>2.60</td>
<td></td>
</tr>
<tr>
<td>Within Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty level</td>
<td>179.41</td>
<td>2</td>
<td>89.70</td>
<td>43.79</td>
<td>.001</td>
</tr>
<tr>
<td>Group x Level</td>
<td>10.64</td>
<td>2</td>
<td>5.32</td>
<td>2.60</td>
<td>.079</td>
</tr>
<tr>
<td>Within cells</td>
<td>237.62</td>
<td>116</td>
<td>2.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**ANOVA: Reading Measures**

Word Attack test

One-way Analysis of Variance revealed significant memory group differences on the measure of decoding skills, \( F(2,133)=10.87, p<.001 \) (see Table 11). Follow-up Tukey procedures determined that the significant differences were between the low and high and the middle and high memory groups. There were no significant differences between the low and middle groups. These findings supported the frequent observation in the literature of a link between decoding skill and verbal working memory ability.

---

**Figure 3.** Mean correct responses, in raw scores, for the low and high memory groups on the Spache Listening Comprehension Task, \( N=60 \)
Table 11

Means, Standard Deviations, and Ranges for the Word Attack Subtest for Each Memory Group, and Analyses of Group Differences with Follow-up Tukey results, N=136

<table>
<thead>
<tr>
<th></th>
<th>Low Group</th>
<th>Middle Group</th>
<th>High Group</th>
<th>F</th>
<th>(df)</th>
<th>F prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>F</td>
<td>(df)</td>
<td>F prob.</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
<td>(SD)</td>
<td>(SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>range</td>
<td>range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N=45</td>
<td>N=46</td>
<td>N=45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Attack subtest</td>
<td>96.56</td>
<td>99.80</td>
<td>107.42</td>
<td>10.87</td>
<td>(2,133)</td>
<td>.001</td>
</tr>
<tr>
<td>range</td>
<td>73-123</td>
<td>70-117</td>
<td>82-129</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* a significance between the High and Low, and the High and Middle groups, p<.05

The PILAR reading comprehension test

As described earlier, the students were presented with third-, fifth-, and seventh-grade passages which they read silently to themselves. After they finished reading each passage, it was removed from view and a recognition task was completed. (See Table 12 for the means and standard deviations of each memory group for each level of passage difficulty).

A 3x3 mixed repeated measures ANOVA was conducted to compare performance by the low, middle and high memory groups. For this task, the main effect of memory group was significant. In addition, a significant interaction was observed between memory group membership and level of
reading passage, $F(4, 266) = 2.70, p < .031$ (see Table 13). As noted in Figure 4, the differences between memory groups were most marked for the seventh grade passage.

Table 12

Means and Standard Deviations for Three Memory Groups on the Three Grade Levels of Passage Difficulty of the PILAR Reading Comprehension Task, $N=136$

<table>
<thead>
<tr>
<th>Passage Difficulty</th>
<th>Low Memory Mean (SD)</th>
<th>Middle Memory Mean (SD)</th>
<th>High Memory Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Third Grade</td>
<td>9.64 (1.42)</td>
<td>9.93 (1.64)</td>
<td>10.18 (1.39)</td>
</tr>
<tr>
<td>Fifth Grade</td>
<td>9.33 (2.10)</td>
<td>9.89 (1.80)</td>
<td>9.82 (1.48)</td>
</tr>
<tr>
<td>Seventh Grade</td>
<td>8.93 (2.06)</td>
<td>9.63 (1.90)</td>
<td>10.44 (1.45)</td>
</tr>
</tbody>
</table>
### Table 13

**Summary table for 3x3 ANOVA for PILAR Reading Comprehension scores by Memory Group and Levels of Passage Difficulty, N=136**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>Sign. of F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory Group</td>
<td>48.92</td>
<td>2</td>
<td>24.46</td>
<td>4.36</td>
<td>.015</td>
</tr>
<tr>
<td>Within cells</td>
<td>746.72</td>
<td>133</td>
<td>5.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty level</td>
<td>5.37</td>
<td>2</td>
<td>2.69</td>
<td>1.67</td>
<td>.189</td>
</tr>
<tr>
<td>Group x Level</td>
<td>17.35</td>
<td>4</td>
<td>4.34</td>
<td>2.70</td>
<td>.031</td>
</tr>
<tr>
<td>Within cells</td>
<td>426.63</td>
<td>266</td>
<td>1.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.** Mean correct responses, in raw scores, for the low, middle, and high memory groups on the PILAR Reading comprehension test, N=136
The Effects of IQ

Since a significant difference in IQ between the memory groups was obtained, it was of interest to investigate the influence of IQ upon several of the dependent measures. In the analyses described below, the estimated full-scale IQ scores were used as covariate scores. (See Appendix G for the results of ANCOVAs using the WISC-R Vocabulary and Block Design subtests as separate covariates).

Three one-way ANCOVAs were performed to examine whether group differences on each of the memory tasks would remain significant after the effects of estimated full-scale IQ were controlled. On all three memory tasks, group differences remained significant after the ANCOVAs were completed: Word span, $F(3,132)=117.21, p<.001$; Pseudoword Repetition Task, $F(2,132)=11.78, p<.001$; Sentence Span, $F(2,132)=34.22, p<.001$.

A fourth one-way ANCOVA was conducted to investigate whether significant memory group differences would remain on the decoding task after the effects of estimated full-scale IQ were controlled. When the effects of IQ were co-varied out, the differences between the memory groups again remained significant, $F(3,133)=6.44, p<.002$. This finding provides support for the hypothesis proposed by Stanovich et al. (1984) that decoding ability taps linguistic skills that are relatively independent of general IQ.
The influence of IQ upon reading and listening comprehension was also studied. Using mixed repeated measures ANCOVAs with both the PILAR reading comprehension and the Spache listening comprehension tasks, the effects of IQ were assessed. When the effects of IQ were controlled, the main effect of memory group membership for both tests was no longer significant. The contribution of IQ to reading comprehension was further explored through regression analyses (to be discussed below).

**Predicting Reading Comprehension**

The relative contributions of the different measures to reading comprehension were also examined. Hierarchical multiple regressions were used to investigate how much of the variance in reading comprehension could be accounted for by listening comprehension, decoding, IQ, and memory performance (see Table 14).

In accord with the Gough and Tunmer (1986) prediction that reading comprehension ability is accounted for by decoding and listening comprehension skills, listening comprehension and decoding dyads were examined first. The total Spache listening comprehension scores and the Word Attack test accounted for 31% of the variance in reading comprehension, with the Spache making the only significant contribution. The combination of the total PILAR Listening scores and the Word Attack test accounted for 21% of the
variance, with both variables contributing significantly to reading. Thus, the listening and decoding combination maximally accounted for 31% of the variance. While it is evident that listening comprehension and decoding make an important contribution to reading comprehension, the proportion of the variance accounted for in the present study is less than the Gough and Tunmer theory would predict.

The combination of variables that accounted for the greatest proportion of variance (37%) in reading comprehension was the total Spache and the composite memory scores, with both variables making significant contributions. When IQ and the decoding measure were added to the above combination, the additional proportions of variance were nonsignificant. The composite memory score and the Spache Listening task consistently contributed between 4-7% and 15-31%, respectively, of significant unique variance to reading comprehension regardless of their order of entry into the regressions.

1 The Gough and Tunmer theory is based on a multiplicative model, while multiple regressions are additive processes. Therefore, analyses completed here were unable to fully test their model and can be used only to suggest sets of variables for further evaluation with multiplicative analyses.
IQ only contributed significantly to reading comprehension when entered without the total Spache scores. The Word Attack measure made significant contributions only when entered with the PILAR listening task.

In sum, the combination of listening comprehension and memory skills was more predictive of reading comprehension for this normal fifth-grade sample than was the combination of listening comprehension and decoding ability. Perhaps the Gough and Tunmer model may be more appropriately applied to students at the beginning stages of reading, when decoding skills have the strongest correlations with reading comprehension (Stanovich et al., 1984).
Table 14

Hierarchical Multiple Regressions with Reading Comprehension as the Dependent Variable: N=136 without Spache; N=76 with Spache

<table>
<thead>
<tr>
<th>Combinations in the order entered</th>
<th>Adjusted $R^2$</th>
<th>Change $R^2$</th>
<th>Signif.</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Spache Word Attack</td>
<td>.306</td>
<td>.316</td>
<td>s</td>
<td>.548***</td>
</tr>
<tr>
<td>b. PILAR Word Attack</td>
<td>.313</td>
<td>.016</td>
<td>ns</td>
<td>.127</td>
</tr>
<tr>
<td>c. Spache Memory IQ Word Attack</td>
<td>.168</td>
<td>.175</td>
<td>s</td>
<td>.415***</td>
</tr>
<tr>
<td>d. Word Attack IQ Memory Spache</td>
<td>.211</td>
<td>.048</td>
<td>s</td>
<td>.219**</td>
</tr>
<tr>
<td>e. IQ Memory Word Attack</td>
<td>.306</td>
<td>.316</td>
<td>s</td>
<td>.450***</td>
</tr>
</tbody>
</table>

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

The primary goal of the present study was to investigate whether verbal working memory ability relates to performance on both written and aural language tasks. Using the Processing Limitation Hypothesis as a conceptual framework, it was predicted that children who demonstrate poor performance on verbal working memory measures would also do less well on listening comprehension, reading comprehension, and decoding measures when compared to children with better verbal working memory skills.

An additional purpose of the present study was to examine which variables best predict reading comprehension. Based on the model proposed by Gough and Tunmer (1986), it was hypothesized that the combination of listening comprehension and decoding scores would account for the greatest proportion of variance in reading comprehension.

While a central focus of the study was on listening comprehension, the link between memory and reading skills will be commented upon first because the outcome of this portion of this investigation is more straightforward. The results supported the hypothesis that students in the low memory group would demonstrate decoding and reading comprehension deficits when compared to children in the other memory groups. Significant memory group differences were observed on both the decoding and reading comprehension
measures. This finding is consistent with research proposing that verbal working memory ability affects the acquisition of word attack skills, and is associated with poor reading comprehension (Baddeley, 1986; Perfetti, 1985; Stanovich, 1982a; Wagner, 1988).

In a meta-analysis involving both longitudinal and training studies, Wagner (1988) identified phonological coding in working memory as playing an independent and significant role in word analysis. The finding in the present study that working memory is associated with decoding for this average-achieving, fifth-grade sample (110 students out of 136 achieved a standard score of 100 or above on the decoding measure) argues that some aspect of working memory is important even for adequate decoders.

The fact that reading comprehension was assessed with a recognition task, which minimizes the memory requirements, strengthens the argument that working memory and reading comprehension are associated. Indeed, the way in which comprehension was measured excluded the need to make inferences, draw conclusions, etc., and tapped only minimal comprehension skills. Would the link between memory and comprehension have been even stronger if more complex comprehension demands had been made? Regardless, the results with the reading measures support Pennington's (1991) claim that working memory is associated with reading comprehension, but argued against his stance that it is not
related to decoding skill.

These results, like those of Torgesen (1990) and Daneman and Carpenter (1980), demonstrate that the relevancy of verbal working memory skills to reading ability can be observed when the research design selects groups according to memory ability as well as by the more frequent design of grouping subjects according to reading levels.

The connection between verbal working memory skills and listening comprehension ability was less pronounced than it was for reading comprehension. The hypothesis that children in the low memory group could not complete listening comprehension tasks with the same degree of accuracy as children in the middle and high memory groups was not consistently supported.

On the first listening measure (PILAR), which used the Sentence Verification Technique (SVT), the memory groups demonstrated equivalent oral language comprehension skills. The SVT requires students to recognize previously learned material. As discussed earlier, students with poor memories can often perform recognition tasks as well as their counterparts with better memory skills (Torgesen, 1988). Thus, significant memory group differences were not observed on the SVT task for listening comprehension, although significant group differences were observed on the PILAR reading comprehension task. The results of the two PILAR measures converge with those of Horowitz and Samuels (1985),
who also found that good and poor sixth-grade readers did not differ on listening comprehension tasks but did differ significantly on reading comprehension tasks. In their study, students were asked to freely recall as many ideas as they could from the stories that they listened to or read. Horowitz and Samuels (1985) argued that the difference between the reading groups on listening and reading comprehension scores was due to discrepancies in decoding skill rather than in general language comprehension ability. If the poor readers were experiencing global language deficits, they maintained, then their problems would have surfaced on the listening task as well as on the reading comprehension measure. In the present study, if one looked only at the results of the PILAR measures, the findings are comparable with the Horowitz and Samuels hypothesis and suggest that having limited memory capacity is not associated with reduced listening comprehension.

However, the outcome changed in this study when the more demanding listening comprehension task requiring factual recall (Spache) was given to the students. For this second listening measure, memory ability was significantly correlated with listening comprehension, and a significant main effect for memory group membership (between the low and high groups) was observed. Further, a trend was observed such that as the stories became longer and more complex, children with weaker verbal working memory skills
experienced greater difficulty recalling specific information from the stories.

This pattern is interesting for two reasons. First, it was observed under optimal testing conditions. Children with subtle language deficits often function poorly under the stresses and competing demands within the classroom, yet perform well on individually administered tests (Levine & Bernstein, 1991). Second, listening comprehension, like reading comprehension, was measured with tasks requiring only the retrieval of facts or simple ideas, again excluding the need to make inferences or draw conclusions. One might speculate that in a complex classroom setting with more extensive comprehension tasks, such as those entailed in many school activities, that the consequences of memory limitations would be even more evident.

In addition to the experimental results, student comments corroborate the differences in task difficulty found between the PILAR (recognition task) and the Spache (recall task). When asked to compare the two listening measures, one student replied, "The first one [PILAR] is like hearing the story twice. It's easier. You don't have to remember." Another student said of the recall task, "This is hard. You have to remember this!"

The results help to make sense out of the contradictions found in the listening comprehension literature. In previous research, differences between good
and poor readers are reported on listening tasks that place
greater demands on verbal working memory, but are not seen
in studies that keep memory requirements to a minimum (e.g.,
Mann et al., 1989; Spring & French, 1990). In the present
study, students of average to high average cognitive ability
but differing in memory skills demonstrated comparable
listening comprehension given a listening task that tapped
recognition memory. In that condition, all the children
were able to correctly identify ideas from the passages,
suggesting, as Torgesen (1988) proposed, that even students
with poor verbal working memories can comprehend the gist of
information. The effects of individual differences in
memory ability were more apparent when children were asked
to demonstrate comprehension by recalling specific facts
from stories. These results seem to point to memory
difficulties rather than to more extensive comprehension
problems as contributory to the larger number of errors made
on the Spache by the subjects in the low memory group.

The second purpose of the study was to evaluate the
relative strength of the dependent variables to predict
reading performance. Confirming the results from previous
studies with older children, listening comprehension emerged
as the single variable that contributed the largest
proportion of unique variance to reading comprehension.
Curtis (1980) also found that listening comprehension made
the most unique contribution to reading comprehension for a
fifth-grade population. In the Report of the Commission on Reading, Anderson (1985) reported that listening comprehension in kindergarten and in the first grade was a strong predictor of later reading achievement. In addition, he reported that in a nationwide study involving thousands of students, listening comprehension in the fifth grade was the best predictor of aptitude and achievement in high school.

Each of the listening comprehension measures contributed more individual variance to reading comprehension than did the other variables. However, the Spache which required the recall of information contributed 31%, while the PILAR which involved the recognition of ideas contributed 17% unique variance. This finding suggests, as Daneman and Carpenter proposed (1980), that accurate recall of specific information is important for adequate comprehension to occur.

In order to evaluate the Gough and Tunmer model, both the PILAR Listening and the Spache tasks were tested in conjunction with the decoding measure. Contrary to what was predicted, neither combination emerged as the dyad that contributed the most variance to reading comprehension. Instead, the combination of listening comprehension as measured by the Spache and the composite memory score contributed the greatest proportion of the variance to reading. Importantly, memory consistently accounted for a
significant and greater proportion of unique variance in reading comprehension than did decoding skills.

There are at least two possible explanations for not finding support for the Gough and Tunmer prediction. First, the formula may be more appropriately applied to beginning or less skilled readers, when individual differences in decoding ability are most strongly associated with the development of reading skills (Stanovich, 1982a). For example, Curtis (1980) found that in the fifth grade, decoding measures were more related to reading achievement for less skilled readers than for their skilled peers. For Curtis' average achieving fifth graders, variations in decoding skills did not seem to affect comprehension levels. Anecdotal evidence from this study supports Curtis' findings and provides an alternative hypothesis as to the effect of decoding ability in the fifth-grade. One teacher reported that his adequately reading students who performed most poorly on the decoding task demonstrated very poor spelling ability.

The second explanation relates to how decoding was measured. According to Stanovich, Cunningham, and Feeman (1984), "...speed of decoding, in addition to accuracy, is important due to short-term memory and other processing limitations that place constraints on how rapidly word meanings must be identified in order to sustain adequate comprehension" (p.287). In fact, in some studies, good and
poor readers have differed in the speed rather than accuracy of decoding (Perfetti & Lesgold, 1979). In the current study, decoding was assessed with an accuracy measure only. Perhaps if the speed had been measured as well, the decoding measure would have been more sensitive to the relationship between decoding and comprehension at the upper elementary grades.

The Question of IQ

An important question in cognitive research concerns the interaction between IQ and verbal working memory skills. If, as in the present study, those with better verbal memory scores also have higher IQ's, is it possible to dissociate the contribution of IQ to memory performance? Do the verbal measures on IQ tests tap many of the same cognitive processes that verbal memory tasks require? Thus, will people with higher IQ's necessarily have better working memories, or will people with good memories score higher on IQ tests because they are better at learning and remembering vocabulary and world knowledge, areas commonly tested on IQ tests?

The literature has mixed responses to the question of the interdependence between IQ and verbal working memory, and how they relate to reading skill. Bowers, Steffy, and Tate (1988) found that statistical control of verbal IQ considerably reduced the contribution of memory to word
They argued that controlling for IQ may be difficult because of shared variance between memory and IQ. However, in another study designed to clarify the role of IQ in the definition of reading disability, Siegel (1988) found memory processes to exist independently of IQ. In her study, children were considered reading disabled if they scored less than or equal to the twenty-fifth percentile on the reading subtest of the Wide Range Achievement Test. Siegel observed similar language and short-term memory deficits in reading-disabled students aged 9-10 years old who were in three IQ groups: 80-90; 90-109; and greater than 110. In spite of a wide IQ range, all of the reading-disabled students demonstrated similar deficits on reading and on working memory measures. The normal reading children, also divided into the three IQ levels, demonstrated no problems on the reading and memory measures despite the different IQ levels. Using a different design, Ellis and Large (1987) matched good and poor readers of equally high IQ scores, and found significant performance differences on the verbal short-term memory tasks between the two groups. Further evidence for the dissociation between IQ and verbal working memory skills emerged from the study by Healy, Aram, and Horowitz (1982) that examined the condition of hyperlexia. Healy et al. (1982) studied 12 hyperlexic children who were good decoders despite overall low cognitive functioning. These children demonstrated a
relatively isolated strength in verbal working memory abilities.

How responsible was IQ for the results observed in the present study? The pattern of low and high memory group differences on the memory, listening, and reading comprehension tasks paralleled a pattern of group differences on the IQ measure. That is, subjects with higher IQ scores tend to have better working memory, decoding, and comprehension abilities. In this study, IQ was first controlled by limiting participation to those children with estimated full-scale scores in the average to high average ranges (Weschler, 1974) and later (for some variables) by statistical means. The mean IQ scores for the low and high groups differed by seven points, and the ranges were identical. While the seven point difference was found to be statistically significant, the important question is whether the difference was functionally meaningful.

In the present research, memory group differences on the decoding and memory tasks withstood the statistical control of IQ, remaining significant. On the comprehension measures, overall levels of performance were no longer statistically significant when IQ differences were controlled. However, in the multiple regression analyses, working memory was found to make a significant contribution to reading comprehension even if IQ was entered first. While the important debate continues as to the
interdependence and shared variance between IQ and working memory abilities, these results point to at least some separate factors contributing to verbal memory and IQ scores. More importantly, the association between verbal memory ability and reading performance does not appear to be merely a by-product of IQ.

Practical Implications

Historically, reading disability has been referred to as the invisible handicap. Subtle oral language deficits may, in fact, be even less likely to be detected. Comparing the results of the two listening tasks suggests that students with poor working memories comprehend ideas as well as those with good working memory skills but have trouble recalling specific details. The present results suggest that a child with known memory deficits may experience listening comprehension difficulties if the comprehension task stresses working memory. Recall that Daneman and Carpenter (1980) found that college students with shorter memory spans made more factual recall errors than students with larger memory spans, and subsequently demonstrated misunderstandings of the main ideas and themes of passages to which they listened.

A purpose of educational research is to develop more appropriate interventions for children experiencing learning difficulties. The applied implications from this study fall
into the following categories: instructional accommodations and metacognitive instruction. Instructional accommodations are variations in teaching techniques that help students achieve success with the regular curriculum. Metacognitive strategies engage the learner in the regulation of cognitive activity.

Instructional accommodations include advance organizers to help students develop a context for new material. Advance organizers provide students with new vocabulary words and an outline of themes that will be taught in a new unit in advance of when the words and themes are presented in class. This provides children with the opportunity to become familiar with new material prior to hearing it in a class presentation. Text schema research has consistently demonstrated that when students have a familiar context in which to frame new material, they retain and recall the new information better (Lyman & Collins, 1990; Torgesen, 1985). Since new information is more easily remembered if it can be integrated with existing knowledge, advance organizers may be particularly helpful to children with working memory deficits.

Direct instruction in metacognitive strategies is the second type of implication to be discussed. Metacognition "refers to readers' abilities to reflect on, monitor, and evaluate their understanding of material as they read, and to apply [cognitive] strategies when comprehension is
impaired," (Schmitt & Baumann, 1990, p.1). Consistently, research has shown that children with poor memories demonstrate improvements in both memory and comprehension when metacognitive strategies are taught and utilized (Schmitt & Baumann, 1990; Tarver, Hallahan, Kauffman, & Ball, 1976; Zabrucky & Ratner, 1990). For example, Zabrucky and Ratner (1990) found that after learning-disabled children with poor metacognitive abilities were taught monitoring strategies, they demonstrated improvements in recall of facts and in comprehension. Although the working memory difficulties of reading-disabled individuals do not appear to only arise from metacognitive deficiencies, it is likely that instruction in metacognitive strategies would help reading-disabled children comprehend and retain what they are learning.

\[2\] Some have argued that the memory problems experienced by learning-disabled students are the consequence of poor metacognitive abilities, because their memory scores show improvement following metacognitive instruction. However, the relative differences in memory scores between reading-disabled students and non-disabled children remain following instruction in metacognitive strategies: both groups show comparative improvements (See Brady, 1991, for a review).
Summary

What emerged from this investigation are two findings: evidence of a link between verbal working memory and each of the major components of reading (decoding and comprehension), and evidence that the memory demands of listening tasks may impact upon accurate listening comprehension. More research is needed to determine how to identify those students who may be at academic risk because of memory difficulties, and the relative point (of task difficulty) at which memory deficits begin to interfere with listening comprehension accuracy. Such research should evaluate measures of listening comprehension that assess a broad range of comprehension skills.

In addition, future research is needed to clarify the role of memory in decoding and in reading comprehension. Perhaps different aspects of memory processing are being utilized in these two tasks. The fact that the removal of the effects of IQ significantly impacted upon memory group differences on the reading comprehension task but not on the decoding measure underscores the different processing requirements of these components of reading.

Contrary to what was expected by the Gough and Tunmer (1986) model, listening comprehension combined with the total memory score emerged as the dyad that accounted for the largest proportion of the variance in reading comprehension for this fifth-grade sample. Future
developmental research is needed to investigate whether the listening/decoding combination proposed by Gough and Tunmer (1986) is more predictive at earlier grades and/or for less skilled readers.

Finally, the results call attention to the need to expand educational accommodations used with students with memory problems. Instructional accommodations, as well as direct instruction in metacognitive strategies, were recommended as helpful curricular modifications.
References


Appendix A

Date ___________________  # __________________
Examiner __________________ Name __________________
School ___________________ Teacher __________________

FAMILIAR WORD SPAN MEMORY TEST
(Use tape recording)

Reliability

<table>
<thead>
<tr>
<th>No.</th>
<th>Words</th>
<th>Order/Recall</th>
<th># Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>cat, fly, score, meat, scale</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td></td>
<td>(Response and # of position)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>roar, wheat, fat, tail, sky</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>3.</td>
<td>tie, hat, nail, floor, sheet</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>4.</td>
<td>mail, pie, store, cap, feet</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>5.</td>
<td>treat, door, eye, sail, map</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>6.</td>
<td>bell, state, knee, pain, chair</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>7.</td>
<td>bee, cell, train, air, plate</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>8.</td>
<td>gate, brain, pair, tea, well</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>9.</td>
<td>bear, key, weight, shell, chair</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
<tr>
<td>10.</td>
<td>rain, hair, spell, fate, tree</td>
<td><em><strong>///</strong></em>___</td>
<td>___</td>
</tr>
</tbody>
</table>

R/S  ___/50  %  ___
(order)  (order)  (recall)  (recall)  (order)  (recall)

Raw Score  ___/50  ___/50  ___/50  ___/50
Appendix B

Pseudoword Repetition Test

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>School</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Dopelate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Glistering</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Pennel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Defermication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Contramponist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Hampent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Reutterpation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Perplisteronk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Blonterstaping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Sepretennial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Detratapillic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Glistow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Frescovent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Bannifer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Stopogratnic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Woogalamic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Ballop</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Confrantually</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Fenneriser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Altupatory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Pristoractional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Underbrantuand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Trumpetine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Sladding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Commecitate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Tafflest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Loddenapish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Barrazon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Commerine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Empliforvent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Thickery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Voltularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. Versatrationist</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Rubid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Brasterer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Diller</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. Penneriful</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. Bannow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39. Prindle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40. Skiticult</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Sentence Span Test: # of words/# of syllables

Sets of two:
1. People buy many things in stores. 6/8
2. Wax candles grow in gardens. 5/7
   mean words=5.5
   mean syllables= 7.5
1. We see colors in a rainbow. 6/8
2. Owls are a kind of insect. 6/8
   mean words=6
   mean syllables= 8
1. As plants grow, they get longer roots. 7/8
2. At night, we sleep in school. 6/6
   mean words= 6.5
   mean syllables= 7

Sets of three:
1. Magnets attract things made of metal. 6/9
2. A beach is covered with sand. 6/7
3. Good friends never talk to each other. 7/9
   mean words= 6.33
   mean syllables= 8.33
1. We milk cows to get eggs. 6/6
2. The weather is always the same. 6/8
3. We see ourselves in a mirror. 6/8
   mean words= 6
   mean syllables= 7.33
1. We learn to read in school. 6/6
2. Eagles are protected by laws. 5/8
3. People drink out of forks. 5/6
   mean words= 5.33
   mean syllables= 6.66

Sets of four:
1. Glaciers are found in the desert. 6/8
2. New York is a large city. 6/7
3. It doesn't rain in the summer. 6/8
4. Food helps your body stay healthy. 6/8
   mean words= 6
   mean syllables= 7.75
Sentence Span test continued:

1. The pilgrims sailed the ocean on ships. 7/9
2. Animals do not drink water. 5/8
3. A fire always gives off heat. 6/8
4. Apples can be picked up with magnets. 7/9
   mean words = 6.25
   mean syllables = 8.5

1. Earthquakes can be very dangerous. 5/9
2. Soap is used for getting dirty. 6/8
3. In autumn trees grow new leaves. 6/7
4. Bears like to eat honey. 5/6
   mean words = 5.5
   mean syllables = 7.5

Sets of five:
1. Sun can burn your skin. 5/5
2. We have snowstorms in the summer. 6/8
3. There are seven days in a week. 7/8
4. Elephants are bigger than flies. 5/8
5. The world is flat like a plate. 7/7
   mean words = 6
   mean syllables = 7.2

1. Most fish can fly in the air. 7/7
2. Bicycles are faster than planes. 5/8
3. A clock is used to tell time. 7/7
4. There are 12 months in a year. 7/7
5. Ice cubes are made in an oven. 7/8
   mean words = 6.6
   mean syllables = 7.4

1. A turtle has a hard shell. 6/7
2. Pencils have erasers on the end. 6/9
3. The moon is a ball of fire. 7/7
4. Rocks can float on water. 5/6
5. Plants and trees are living things. 6/7
   mean words = 6
   mean syllables = 7.2
Score Sheet for Sentence Span test

Date: _______________  # __________
Examiner: _______________  Name: _______________
School: _______________  Teacher: _______________

Check the words that are correctly recalled, regardless of order.

SETS OF TWO

A. stores___; gardens___
   ___
B. rainbow____; insect___
   ___
C. roots____; school____
   ___

SETS OF THREE

A. metal____; sand____; other____
   ___
B. eggs____; same____; mirror____
   ___
C. school____; laws____; forks____
   ___

SETS OF FOUR

A. desert____; city____; summer____; healthy____
   ___
B. ships____; water____; heat____; magnets____
   ___
C. dangerous____; dirty____; leaves____; honey____
   ___

SETS OF FIVE

A. skin____; summer____; week____; flies____; plate____
   ___
B. air____; planes____; time____; year____; oven____
   ___
C. shell____; end____; fire____; water____; things____
   ___

Total ________
Appendix D
Profiles in Listening and Reading (PILAR)

SLED DOGS
(Fifth-grade listening comprehension passage)

Eskimo sled dogs are raised to work. They are used to pull sleds across the ice and snow in northern lands. First, the dogs are teamed together in groups of four to twelve dogs. Then they are harnessed to the sled after it has been loaded. On the master's command, they start pulling the sled across the snow. This is a hard job because the load on the sled can weigh more than 2000 pounds. The master drives his team as far as 20 miles across the cold land. Though the sled dogs become tired, their master does not show them much sympathy. Dogs who are not doing their part of the work may feel the sting of the master's whip. And one dog may snap at another who is not pulling his share. Still, when the work is done, the master makes sure that his dogs are well fed. Understanding how hard they work, he takes care that they stay healthy and strong.
Statements for "Sled Dog" Listening Passage

Directions to be read to students: Listen carefully to each sentence. Is the idea of the sentence found in the passage? If it is, circle YES. If it is not, circle NO.

1. First, the dogs are separated into teams of between four and twelve dogs.

2. They are used to pull the sleds across the ice and snow in northern lands.

3. Eskimo sled dogs are raised to work.

4. Sometimes the master leaves his team to go hunting or fishing.

5. Since the sled and its contents may weigh more than 2,000 pounds, hauling it is a tough job.

6. Some of the dogs live to be fifteen years old.

7. When the sled dogs become tired, the master shows them sympathy.

8. The master makes sure that his dogs are given the best care because he is aware of how hard they work.

9. The master drives his team less than 20 miles across the cold land.

10. One dog may snap at another who is not pulling his share.

11. Each year a big race for teams of sled dogs is held in Alaska.

12. Dogs who are doing their part of the work may feel the sting of the master's whip.
PILAR Listening Passages Answer Sheet

<table>
<thead>
<tr>
<th>Passage 1:</th>
<th>Passage 2:</th>
<th>Passage 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
<tr>
<td>YES NO</td>
<td>YES NO</td>
<td>YES NO</td>
</tr>
</tbody>
</table>
ICE BOXES
(Fifth-grade reading comprehension passage)

How did people keep fresh food from spoiling before refrigerators were invented? Many people kept their fresh food in ice boxes. An ice box was a large chest with one space on top for a block of ice and a second space for the food. The ice came from frozen lakes and rivers in the winter. This ice was cut up with large saws and stored in ice houses. It was covered with saw dust to keep it from melting. For as long as the ice lasted, blocks of ice were brought to the homes of people who had ice boxes. But there was little ice for use in ice boxes in warm winters. And there was no ice for ice boxes in hot climates. Food spoiled quickly in such places, too. Where ice boxes could not be used, people kept potatoes and carrots in the cellar of their houses. Even with ice boxes and cool cellars, food could not be kept fresh for very long!
Questions for the Reading Passage

Read each sentence carefully. Is the idea of the sentence found in the passage? If it is, circle YES. If it is not, circle NO.

YES NO 1. Ice boxes were used by lots of people to prevent the food from going bad.

YES NO 2. This ice was cut up with large saws and stored in ice houses.

YES NO 3. In the winter, ice was taken from frozen lakes and rivers.

YES NO 4. It was covered with saw dust to keep it from freezing.

YES NO 5. An ice box was a large chest with one space on top for a block of ice and a second space for the food.

YES NO 6. Large sleds or wagons were used to move the blocks of ice.

YES NO 7. Where ice boxes could not be used, people kept potatoes and carrots in the cellar of their houses.

YES NO 8. Before refrigerators were invented, people often became quite ill from eating spoiled meat.

YES NO 9. But there was little ice for use in ice boxes in cold winters.

YES NO 10. Blocks of ice were delivered to the homes of people for use in the ice boxes until the supply of ice was gone.

YES NO 11. Men who brought the blocks of ice to the peoples' homes were called ice men.

YES NO 12. With ice boxes and cool cellars, food could be kept fresh for a very long time.
Appendix E

Means, Standard Deviations, and Ranges in Raw Scores (except where noted) for the Sample Population (N=136), for the Dependent Measures

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Span:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct Order (50)</td>
<td>26.95</td>
<td>11.08</td>
<td>2-48</td>
</tr>
<tr>
<td>Word Span:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free Recall (50)</td>
<td>35.48</td>
<td>6.93</td>
<td>18-48</td>
</tr>
<tr>
<td>Pseudoword Repetition (40)</td>
<td>35.14</td>
<td>4.19</td>
<td>0-40</td>
</tr>
<tr>
<td>Sentence Span (42)</td>
<td>26.36</td>
<td>5.02</td>
<td>10-39</td>
</tr>
<tr>
<td>Word Attack*</td>
<td>101.25</td>
<td>12.153</td>
<td>70-129</td>
</tr>
<tr>
<td>PILAR Listening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third grade (12)</td>
<td>10.43</td>
<td>1.48</td>
<td>7-12</td>
</tr>
<tr>
<td>Fifth grade (12)</td>
<td>10.36</td>
<td>1.43</td>
<td>5-12</td>
</tr>
<tr>
<td>Seventh grade (12)</td>
<td>9.51</td>
<td>1.47</td>
<td>2-12</td>
</tr>
<tr>
<td>Total: 3rd, 5th, 7th (36)</td>
<td>30.28</td>
<td>3.61</td>
<td>20-36</td>
</tr>
<tr>
<td>PILAR Reading</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third grade (12)</td>
<td>9.92</td>
<td>1.49</td>
<td>5-12</td>
</tr>
<tr>
<td>Fifth grade (12)</td>
<td>9.68</td>
<td>1.82</td>
<td>4-12</td>
</tr>
<tr>
<td>Seventh grade (12)</td>
<td>9.67</td>
<td>1.91</td>
<td>4-12</td>
</tr>
<tr>
<td>Total: 3rd, 5th, 7th (36)</td>
<td>30.28</td>
<td>3.61</td>
<td>20-36</td>
</tr>
</tbody>
</table>

*standard scores with mean=100, SD=15
Appendix F

Means, Standard Deviations, and Ranges in Raw Scores (except where noted) for the Subsample, N=76, for Dependent Measures

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Span Order (50)</td>
<td>26.37</td>
<td>11.99</td>
<td>2-48</td>
</tr>
<tr>
<td>Pseudoword Repetition (40)</td>
<td>35.8</td>
<td>5.09</td>
<td>0-40</td>
</tr>
<tr>
<td>Sentence Span (42)</td>
<td>26.58</td>
<td>5.62</td>
<td>10-39</td>
</tr>
<tr>
<td>Word Attack subtest*</td>
<td>100.58</td>
<td>12.55</td>
<td>73-129</td>
</tr>
<tr>
<td>PILAR Listening Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third grade (12)</td>
<td>10.63</td>
<td>1.45</td>
<td>7-12</td>
</tr>
<tr>
<td>Fifth grade (12)</td>
<td>10.37</td>
<td>1.65</td>
<td>5-12</td>
</tr>
<tr>
<td>Seventh grade (12)</td>
<td>9.49</td>
<td>1.65</td>
<td>5-12</td>
</tr>
<tr>
<td>Total: 3rd, 5th, 7th (36)</td>
<td>30.46</td>
<td>3.75</td>
<td>20-35</td>
</tr>
<tr>
<td>PILAR Reading Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third grade (12)</td>
<td>10.03</td>
<td>1.46</td>
<td>5-12</td>
</tr>
<tr>
<td>Fifth grade (12)</td>
<td>9.55</td>
<td>1.94</td>
<td>4-12</td>
</tr>
<tr>
<td>Seventh grade (12)</td>
<td>9.74</td>
<td>1.96</td>
<td>5-12</td>
</tr>
<tr>
<td>Total: 3rd, 5th, 7th (36)</td>
<td>29.42</td>
<td>4.35</td>
<td>17-36</td>
</tr>
</tbody>
</table>

*Standard scores with mean=100, SD=15
<table>
<thead>
<tr>
<th>Test</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spache Listening Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third grade</td>
<td>7.05</td>
<td>1.47</td>
<td>0-8</td>
</tr>
<tr>
<td>(8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fifth grade</td>
<td>5.20</td>
<td>1.99</td>
<td>0-8</td>
</tr>
<tr>
<td>(8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seventh grade</td>
<td>4.88</td>
<td>1.95</td>
<td>0-8</td>
</tr>
<tr>
<td>(8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total: 3rd, 5th, 7th</td>
<td>17.13</td>
<td>4.12</td>
<td>4-23</td>
</tr>
<tr>
<td>(24)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix G

Correlation Matrix of Age, IQ, Memory Measures, PILAR Listening and Reading Comprehension Tasks, Spache Listening Comprehension, and Decoding, N=76

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. IQ</td>
<td>-.33*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Vocabulary</td>
<td>-.42**</td>
<td>.56**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Block Design</td>
<td>-.07</td>
<td>.77**</td>
<td>-.10</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Word Span Order</td>
<td>.07</td>
<td>.27*</td>
<td>.16</td>
<td>.20</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Gathercole</td>
<td>.21</td>
<td>.33*</td>
<td>.16</td>
<td>.28*</td>
<td>.43**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Sentence Span</td>
<td>.08</td>
<td>.29*</td>
<td>.32*</td>
<td>.11</td>
<td>.53**</td>
<td>.51**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Composite Memory</td>
<td>.16</td>
<td>.37**</td>
<td>.26</td>
<td>.25</td>
<td>.78**</td>
<td>.82**</td>
<td>.83**</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Spache Total</td>
<td>-.03</td>
<td>.40**</td>
<td>.42**</td>
<td>.16</td>
<td>.20</td>
<td>.20</td>
<td>.29*</td>
<td>.29*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. PILAR Total Listening</td>
<td>-.11</td>
<td>.39**</td>
<td>.21</td>
<td>.31*</td>
<td>.03</td>
<td>.01</td>
<td>.02</td>
<td>.02</td>
<td>.61**</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. PILAR Total Reading</td>
<td>-.19</td>
<td>.37**</td>
<td>.37**</td>
<td>.16</td>
<td>.26</td>
<td>.36**</td>
<td>.38**</td>
<td>.41**</td>
<td>.56**</td>
<td>.48**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>12. Word Attack standard score</td>
<td>-.23</td>
<td>.37**</td>
<td>.40**</td>
<td>.14</td>
<td>.42**</td>
<td>.28*</td>
<td>.34*</td>
<td>.43**</td>
<td>.11</td>
<td>.02</td>
<td>.19</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*p<.01

**p<.001
Appendix H

**ANCOVAs to Determine Memory Group Differences when the WISC-R Vocabulary and Block Design Subtests are used as Covariates.**

<table>
<thead>
<tr>
<th>Test</th>
<th>Vocabulary</th>
<th></th>
<th>Block Design</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>(df)</td>
<td>F</td>
<td>(df)</td>
</tr>
<tr>
<td></td>
<td>F prob.</td>
<td></td>
<td>F prob.</td>
<td></td>
</tr>
<tr>
<td>Word Span:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct Order</td>
<td>122.29</td>
<td>(3,132)</td>
<td>120.12</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td></td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Pseudoword Repetition</td>
<td>13.99</td>
<td>(3,132)</td>
<td>15.21</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td></td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Sentence Span</td>
<td>36.35</td>
<td>(3,132)</td>
<td>39.64</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td></td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Word Attack Subtest</td>
<td>5.76</td>
<td>(3,132)</td>
<td>11.25</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.004</td>
<td></td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>2.13</td>
<td>(3,132)</td>
<td>2.91</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.122</td>
<td></td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>Spache Listening Comprehension</td>
<td>1.90</td>
<td>(3,132)</td>
<td>2.85</td>
<td>(3,132)</td>
</tr>
<tr>
<td></td>
<td>.174</td>
<td></td>
<td>.09</td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


