The Dissertation Committee for Robyn Michelle Honig certifies that this is the approved version of the following dissertation:

Central Executive Processing:
Mother, Daughter, or Sister of Suppression?
A Study of Reading Comprehension Ability

Committee:

Clarke Burnham, Co-Supervisor

Philip Gough, Co-Supervisor

Randy Diehl

Wendy Domjan

Diane Schallert

Jacqueline Woolley

Central Executive Processing: Mother, Daughter, or Sister of Suppression? A Study of Reading Comprehension Ability

by

Robyn Michelle Honig, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of
the University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of
Doctor of Philosophy

The University of Texas at Austin
December, 2004

For my grandfather, an avid reader.

Acknowledgements

I am grateful to the people whose efforts and encouragement aided the completion of this project. They include Philip Gough, who introduced me to reading research and helped me convert a passing assumption to the focus of my study; Jayme Brierton, Jessica Cruz, Shahn Daredia, Kelly Engle, Sarah Hartline, Jonathan Hoyt, Naeem Khowja, Layla Lincoln, Meagan McCarley, Jason Murry, Liz Newlin, Paige Peschong, Michael Rosiles, Rachel Vela, John Wilkinson, and Filipp Zlatkin, who helped me implement hours of procedures on 101 undergraduates; and said 101 undergraduates, for enduring several hours of standardized tests and laboratory tasks. I thank Nathan Baldwin and Matt Spera for enduring several hours of complaining amidst the design and implementation of said laboratory tasks, as well as the writing of this dissertation; Erika Lima and Inna Shtakser for taking an interest in my progress; and my always-supportive family, including my mother, who offered to drive 150 miles to copy blank surveys and response forms. This dissertation owes much to the years of mentoring by Jennifer Watts and Sebastian Wren; the thorough explanations of Morton Ann Gernsbacher; and the statistical expertise of Greg Hixon. I am honored to have received feedback from Randy Diehl, Wendy Domjan, Diane Schallert, and Jacqui Woolley, some of the most captivating instructors and researchers I have ever met. I am especially indebted to Clarke Burnham, who, with thoughtful critiques, incredible patience, and a sense of humor, helped ensure that I would have my Ph.D. by twenty-five.

Central Executive Processing:

Mother, Daughter, or Sister of Suppression?

A Study of Reading Comprehenion Ability

Publication No. _________

Robyn Michelle Honig, Ph.D.
The University of Texas at Austin, 2004

Co-Supervisors: Clarke Burnham and Philip Gough

This study examined the relationship between central executive processing and suppression within the realm of reading comprehension ability. Results indicated that the two mechanisms were unrelated, and the suppression of irrelevant information did not account for variance in reading comprehension ability. In addition, one type of suppression measure produced unexpected results, and two types of tasks that supposedly assessed general central executive processing were unrelated. The results might have been due to a number of factors, including homogeneity of the subjects and the different abilities assessed by the central executive, suppression, and reading comprehension measures. Nonetheless, the study confirmed that knowledge, intelligence, and, to a lesser extent, decoding ability, are predictive of reading comprehension ability. Future research on central executive processing, suppression, and their relationship within the realm of reading comprehension ability should concentrate on more specific forms of these three constructs.

Table of Contents

List of Tables viii
List of Figures vix
Chapter One: Introduction 1
Chapter Two: Method 33
Chapter Three: Results 51
Chapter Four: Discussion 87
Appendices 102
References 121
Vita 131

List of Tables

Table 1: *Low and High NDComp Performers' Mean Median Homograph Suppression Times (ms)* 54

Table 2: *Low and High ADRP Performers' Mean Median Homograph Suppression Times (ms)* 54

Table 3: *Low and High Sem Performers' Mean Median Homograph Suppression Times (ms)* 54

Table 4: *Low and High NDComp Performers' Mean Median Homophone Suppression Times (ms)* 55

Table 5: *Low and High ADRP Performers' Mean Median Homophone Suppression Times (ms)* 55

Table 6: *Low and High Sem Performers' Mean Median Homophone Suppression Times (ms)* 55

Table 7: *Untransformed Descriptive Statistics for Measures to be Correlated* 65

Table 8: *Score Transformations and Resulting Descriptive Statistics for Measures to be Correlated* 68

Table 9: *Internal Consistencies of Measures and Their Forms* 70

Table 10: *Pearson Correlations for Comprehension Measures and Measures Considered for the Principal Components Factor Analysis* 75

Table 11: *Principal Components Exploratory Factor Analysis for Non-Comprehension Measures* 79

Table 12: *Descriptive Statistics for Independent Variables in Regression Analyses* 84

Table 13: *Regression Coefficients for NDCompSem* 86

Table 14: *Regression Coefficients for Advanced Degrees of Reading Power* 86

List of Figures

Figure 1 2
Figure 2 58
Figure 3 58
Figure 4 58
Figure 5 58
Figure 6 59
Figure 7 59
Figure 8 59

Chapter One: Introduction

Reading comprehension requires the melding of several elements. Common sense and the data tell us that knowledge of the language, general intelligence, general knowledge, and decoding ability are all principal components. In recent decades, psycholinguists have suggested two additional mechanisms that appear to influence reading comprehension ability: suppression and central executive processing. The present research examines possible relationships between the two.

Particularly exciting is the possibility of a causal relationship between central executive processing and the suppression of irrelevant information within the realm of reading comprehension. That is, suppression might be a byproduct of the link between reading comprehension ability and central executive processing. Or central executive processing might be a byproduct of the link between reading comprehension ability and the suppression of irrelevant information. If either notion is true, future research on the cognitive underpinnings of reading comprehension ability could not acknowledge one mechanism without acknowledging the other. Nonetheless, if evidence suggests that suppression and central executive processing are independently related to reading comprehension ability, we have still further decomposed a complex construct. Alternatively, if an ambiguous overlap exists between the suppression of irrelevant information and the central executive, future research could focus on the source of the overlap within reading comprehension. These four respective hypotheses, the Central-Executive-As-Mother hypothesis, the Central-Executive-As-Daughter hypothesis, the Sisters hypothesis, and the Conjoined Sisters hypothesis, are illustrated in Figure 1. They

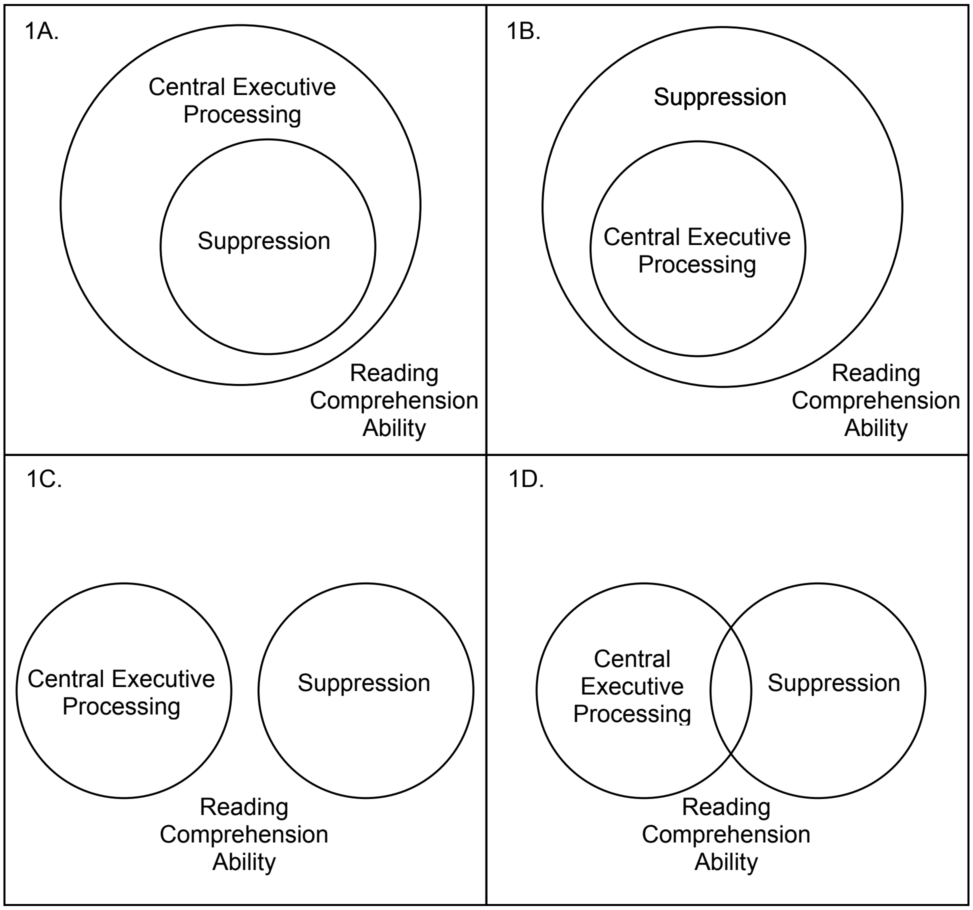

Figure 1. Four hypotheses on the relationship between central executive and suppression mechanisms that might influence reading comprehension ability. *1A.* The Central-Executive-As-Mother hypothesis. *1B.* The Central-Executive-As-Daughter hypothesis. *1C.* The Sisters hypothesis. *1D.* The Conjoined Sisters hypothesis.

all focus on the relationships between central executive processing and suppression within the realm of reading comprehension ability.

The complexity of the reading comprehension process entails that if we wish to isolate one or two of the mechanisms underlying reading comprehension ability, we must do our best to discard the influence of other factors affecting its two doubly-dissociable, major components. These components are *decoding* and *comprehension* (Gough & Tunmer, 1986; Hoover & Gough, 1990; Gough, Hoover, & Peterson, 1996). According to that two-component model, known as the Simple View of Reading, reading comprehension, excluding comprehension, is essentially no more than decoding. The architects of the simple view thus referred to reading comprehension as *reading*. Any future references to reading will refer to the decoding-comprehension combination.

Decoding is the conversion of written symbols to the spoken sounds that combine to form spoken words and phrases. The written form of a spoken language is called an *orthography*. An orthography is essentially a "code" for a spoken language. We translate that code into spoken language via a process like *cryptanalysis*, or code-breaking.

Cryptanalysts are trained to break two types of codes: *codes* and *ciphers*. Codes are arbitrary, meaning that they are *not* constructed according to rules that, once known, facilitate decryption. A current popular code is the system of acronyms used in computerized instant messages. In this code, the acronymous message, "*ty; ttyl,*" means, *Thank you; talk to you later*. Someone wishing to decrypt a discourse composed entirely in the code would have to consult an acronym-English dictionary.

While a code is an arbitrary system, a cipher is a rule-based system that tends to operate on the level of the letter (Calvert, 2001). For example, *Thank you; talk to you later*, could be encrypted as *Uibol zpv; ubml up zpv mbufs*. The rules of that cipher state that each letter of the message should be represented by the letter that follows it alphabetically. A person who has mastered that cipher can decipher any encrypted letter.

The English orthography is essentially a cipher (Gough & Hillinger, 1980)—it is an intricate, rule-based system in which letters correspond to speech sounds, or *phonemes*. Anyone who knows how to apply the letter-phoneme correspondences has mastered the English orthographic cipher. Mastery of the cipher, plus specific knowledge of irregular words, leads to superior decoding ability.

Comprehension, the second major component of reading comprehension, is the understanding of the meanings of the words and phrases that we encounter. Gough and colleagues argued that we comprehend written and spoken words in a very similar fashion (Gough & Tunmer, 1986; Hoover & Gough, 1990; Gough et al., 1996). Factors influencing the comprehension abilities of people who speak and read English include general intelligence, general knowledge, and vocabulary knowledge. Because these factors are necessary for comprehending written English but are not important to the research question, this paper will largely acknowledge them under the collective moniker, *Given*. Independently of decoding ability and the Given, the present research analyzes the relationships among reading comprehension ability, central executive processing, and suppression.

Known Relatives: Decoding Ability and the Given

As aforementioned, reading comprehension would be implausible without the ability to decode text and the capacity to comprehend it. Here, I will briefly discuss how decoding ability and the elements of the Given relate to reading comprehension ability.

Decoding Ability

A wealth of literature reveals a relationship between decoding ability and reading comprehension ability (Stanovich, 2000, pp. 208-209).

McCormick and Samuels (1979) found moderate correlations between first and second-graders' reading comprehension ability and word recognition speed and accuracy (also see Hess, 1982; Levy & Hinchley, 1990; Perfetti & Hogaboam, 1975; Singer & Crouse, 1981). Others have associated children's ability to rapidly decode *pseudowords* (pronounceable nonwords, e.g., *plark*) with relatively skillful reading comprehension (e.g., Curtis, 1980; Hess, 1982; Perfetti & Hogaboam, 1975). Nonetheless, a child is rarely asked to read text that is difficult to comprehend. Therefore, a child's reading comprehension ability is related to decoding ability more than comprehension ability. But as the child ages, the link between decoding ability and reading comprehension ability weakens, and the link between listening comprehension ability and reading comprehension ability strengthens (Chen & Vellutino, 1997). That occurs, in part, because we are more likely to read less comprehensible material as we age.

Using college students, Cunningham, Stanovich, and Wilson (1990) found a moderate correlation between reading comprehension ability and the decoding speed of words and pseudowords. A multiple regression also indicated that decoding ability accounted for reading comprehension variance independent of general intelligence and vocabulary knowledge. Furthermore, Bell and Perfetti (1994) have noted that word recognition speed differentiates good and poor adult readers.

General Intelligence

A number of studies have suggested a positive, though not necessarily strong, relationship between general intelligence and reading comprehension ability. Specifically, reading comprehension ability is correlated with nonverbal and performance measures of general intelligence. Correlations have been found in research with children (Naglieri & Ronning, 2000; Oakhill, Cain, & Bryant, 2003; Singer & Crouse, 1981; Stanovich, Cunningham, & Feeman, 1984; Vellutino, Scanlon, & Lyon, 2000), teens (Naglieri & Ronning, 2000), and adults (Cantwell, 1966; Palmer, Kyllonen, & Christal, 1990; Palmer, MacLeod, Hunt, & Davidson, 1985).

Vocabulary Knowledge

We cannot comprehend text without understanding the individual words we encounter. In fact, several studies have causally linked vocabulary instruction to children's reading comprehension ability (Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson, & Perfetti, 1983; Stahl, 1983). Others have found positive correlations between children's vocabulary knowledge and reading comprehension ability (Naglieri & Ronning, 2000; Oakhill et al., 2003; Singer & Crouse, 1981). Positive relationships have additionally been obtained in studies of teens (Naglieri & Ronning, 2000) and college students (Burt & Fury, 2000; Dixon, LeFevre, & Twilley, 1988).

General Knowledge

Our ability to comprehend a text can depend on whether we approach it with applicable, prior knowledge. While reading a Jane Austen novel, I discovered the benefit of footnotes, which enhanced my understanding of the dialogue by explaining the unfamiliar cultural norms that Austen subtly mocked. In addition to my experience, research supports the notion that knowledgeable readers are better at comprehending text (Perfetti, 1985, pp. 72-78). For example, adults knowledgeable about baseball comprehend baseball-related text better than they comprehend computer-related text (Gough et al., 1996). And McNamara and McDaniel (2004) demonstrated that college students with more general knowledge perform better on a popular reading comprehension test than students with less general knowledge.

A Brief Characterization of Reading Comprehension

Previously, I acknowledged several prerequisites for reading comprehension: decoding skill, general intelligence, general knowledge, and knowledge of the language, including vocabulary knowledge. In turn, those fundamental contributors influence three processes that occur during comprehension: parsing, disambiguation, and integration.

Parsing

Parsing is the process by which we determine the relationships among words in a sentence. Syntax, the rules governing the order of words within phrases, helps us parse those phrases. For example, the sentence, *The dog jumped toward the fence,* is comprehensible partly because the order of the words tells us what was jumping (the dog) and where it went (toward the fence). Without the rules of syntax, the sentence could have been written, *Fence jumped dog the the toward*, resulting in a more opaque meaning.

During reading, punctuation can simplify the parsing process, as shown in the following sentences (source unknown):

The woman, without her man, is nothing.

The woman, without her, man is nothing.

The only difference between the preceding sentences is the placement of one comma, which determines the way we parse each of them. Clearly, parsing influences interpretation.

Disambiguation

Another process, important to comprehension as well as this research, is disambiguation. Words can be ambiguous, as can entire phrases. An ambiguous word or phrase has multiple meanings.

Most ambiguous words are homographic homophones, like *ring*; their meanings are spelled and pronounced the same way. Least common are homographic heterophones, like *bow*; different meanings are spelled the same but pronounced differently, i.e., one meaning of *bow* rhymes with *low*, the other with *cow*. Some ambiguous words are heterographic homophones, like *rain* and *reign*, which sound the same but are spelled differently.

Of the different types of ambiguous words, heterographic homophones are the only ones that can be disambiguated when they are presented in isolation. That is, the meanings of words like *rain* and *reign* can be determined without context clues. To fully disambiguate a written homographic homophone, like *ring*, or a homographic heterophone, like *bow*, context is necessary. Context is not always sufficient, however. A complete sentence can still be ambiguous, e.g., *Iraqi head seeks arms* (source unknown). Disambiguating that sentence requires the application of both context and prior knowledge.

Integration

The third reading comprehension subprocess, integration, is especially important to the present research. Integration is the process by which we link what we are presently reading to what we read earlier in the text. Integration can occur locally within a phrase or a short sequence of phrases, e.g., *It was dark in the forest, so the cousins were scared*. As we read the phrase, *so the cousins were scared*, integration allows us to infer from the prior phrase, *it was dark in the forest*, that the darkness in the forest contributed to the cousins' fright.

Integration can also occur globally, across a large discourse. We can understand the plot of a novel because we integrate new information with information appearing earlier in the text. Both integration and disambiguation will be discussed throughout this paper.

The Many Faces of Suppression

Suppression has multiple meanings, three of which will be discussed in this section.

Suppression can mean the deliberate, conscious disregard of something that would otherwise command our attention. This type of suppression occurs during the

Stroop (1935/1992) task. The Stroop task consists of color names like *red* and *blue* printed in various colors. Typically, the color described by a word and the color in which the word appears are different, e.g., the word *blue* is printed in green. Most skilled readers have no trouble simply reading the words, but they do have trouble naming the words' colors, especially when they are inconsistent with the words themselves. Common sense tells us that this phenomenon occurs because skilled readers habitually attend to and read print, but they do not habitually attend to and name the colors they encounter. Ignoring a dominant stimulus, like a word, to attend to a subordinate stimulus, like a color, necessitates suppressing the automatic response of reading print.

Suppression can additionally mean the active ignoring of relatively subordinate information. More common and less extreme than the type of conscious suppression used in the Stroop task, it occurs when many of us solve arithmetic word problems, as suggested by a study of Italian fourth graders (Passolunghi, Cornoldi, & de Liberto, 1999). Good and poor problem solvers listened to twelve problems, including the one below.

> Four good friends go to a "pizzeria." Each of them eats a pizza which costs 8,500 Liras and orders a drink which costs 2,500 Liras. What does the bill come to? If one of them pays with a bill of 50,000 Liras, how much change will he receive?

The good problem solvers recalled more relevant information, e.g., the price of pizza, and less irrelevant information, e.g., the pizzeria, than the poor problem solvers recalled. Passolunghi et al. (1999) argued that the good problem solvers were better able to inhibit the irrelevant information.

The active ignoring of subordinate information additionally occurs outside the realm of problem solving and inside the realm of reading for enjoyment. For example, if the information in the above word problem appeared within a storyline, not a word problem, you might have remembered the “irrelevant” information but ignored the “relevant” information. In the word problem, the irrelevant information was the setting and characters, which are important to a story; the relevant information in the word problem was the cost of pizza and a drink, typically not an important plot point.

Of course, the notion of irrelevance is subjective. One reader might find a piece of information utterly useless, while another might think that it is not important, but it does enhance his or her understanding of the text. That is, information can be relevant or irrelevant. The relevant information varies in importance, which can range from unimportant to very important. Unimportant information might not be remembered as well as important information, but it is still somehow relevant. In contrast, we view irrelevant information as unrelated to what we wish to read about, so we inhibit it and focus on the relevant information.

An edition of Roget’s thesaurus (Lewis, 1964) defines *inhibit* and *suppress* synonymously. In the context of cognitive research, they both suggest the quashing of activation, i.e., stimulation, relevant to a cognitive representation. Much of the literature concerning what happens to various types of irrelevant information makes no effort to differentiate inhibition and suppression, sometimes interchangeably using the terms. An exception to the norm, Gorfein, Berger, and Bubka (2000) define inhibition as the consequence of suppression, meaning that the suppression mechanism swiftly quashes activation, and the inhibition mechanism prevents further activation for some time.

Describing the suppression and subsequent inhibition of irrelevant information seems simple in the context of stimuli that entail the conscious disregard of irrelevant information, e.g., the stimuli in the Stroop task. Nonetheless, it becomes complicated in the context of other stimuli, specifically ambiguous words.

A number of lexical access theories have suggested that when we read an ambiguous word, all of its meanings should initially be activated (Gorfein, 2001). Most words have multiple meanings, but we are rarely aware of that while we read. Therefore, we must partake in a rapid, disambiguation process.

A breadth of literature has attempted to explain what happens to the irrelevant meanings of the ambiguous words we encounter (for recent discussion, see Gorfein, 2001). Because debate continues on the topic, this space will be used to describe the suppression mechanism that inspired the present research. This particular suppression mechanism appears in Gernsbacher's structure building framework (1990/1996, 1991, 1997).

Gernsbacher (1990/1996, 1991, 1997) wrote that memory nodes are key to the structure building framework. When we comprehend words and phrases, multiple memory nodes are activated, and they connect to form the foundation of a *mental structure*. When nodes representing related information are activated, they connect to the existing structure. Mental structures, therefore, can contain a great deal of information.

Although all of the information in each structure is somehow related, the importance of each bit of information varies with context.

Gernsbacher claims that activated memory nodes send signals to other activated nodes, and the signals enhance or suppress the other nodes' activation depending on how pertinent they are to the context in which the mental structure was activated. The enhanced nodes aid in structure building while the suppressed nodes do not.

The aforementioned theory applies to ambiguous words with two, equiprobable meanings. To examine how we might activate and suppress meanings in real time, Gernsbacher and St. John (2002) simulated the meaning-activation and suppression mechanisms using sentences like those in Gernsbacher et al.'s study (1990, experiment 4). One of the sentences was *Pam was diagnosed by a quack*. *Quack*, of course, can refer to a dishonest doctor or a duck's dialect. Consistent with Gernsbacher's prior research, (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990) both the "doctor" and "duck" meanings of *quack* were initially activated in the simulation. The context of the sentence then allowed the doctor meaning to suppress the duck meaning.

Gernsbacher and St. John (2002) performed two other simulations in which each sentence-final word had one frequent meaning and one infrequent meaning, as opposed to two equally frequent meanings. Like *quack* in the previous simulation, both of the sentence-final words in these new simulations were homographic homophones. The first simulation utilized a word with a relatively small frequency difference between its two main meanings. The second utilized a word with a relatively large frequency difference between its two main meanings. In each simulation, the sentence context favored the less-frequent meaning.

In the first simulation, in which the frequency difference was relatively low, both meanings were activated, and the sentence context allowed the relevant meaning to suppress the irrelevant meaning. In the end, the relevant meaning had been activated more than the irrelevant meaning, though the irrelevant meaning was more frequent. Nonetheless, the relevant meaning was activated less than the (relevant) doctor meaning in the “quack” simulation, so the contextually appropriate meaning’s activation level seemed to depend on its frequency.

In the second simulation, in which the frequency difference was relatively high, the more frequent but irrelevant meaning was activated, but the relevant, though less frequent, meaning was not activated at all. Similarly, a word within the sentence that supported the relevant meaning was suppressed. Gernsbacher and St. John (2002) claimed that the less-frequent meaning was not activated because the sentence context was too weak to clarify that meaning’s relevance. They additionally claimed that the suppressed word had been suppressed because it was largely unconnected to the more frequent meaning of the ambiguous word.

Those simulations imply that if context supports the more frequent meaning of a sentence-final, homographic homophone, the less frequent meaning might be activated and then suppressed. If the less frequent meaning is extremely infrequent, it might not be activated at all. Hence, the speed and intensity of any meaning’s activation might depend on both its degree of contextual appropriateness and its frequency (see, e.g., Duffy, Kambe, & Rayner, 2001; Martin, Vu, Kellas, & Metcalf, 1999). This notion is further supported by McNamara and McDaniel's (2004) and Zwaan and Truitt's (2000) studies. Both studies demonstrated that prior knowledge of a topic influences the rate at which an ambiguous word with one meaning related to that topic is suppressed. Because an expert in a topic has probably encountered certain meanings more often than the average person, the studies also suggest that the amount of exposure to one meaning of an ambiguous word affects its processing in all contexts.

Although the structure building framework has been described here in terms of ambiguous words, Gernsbacher (1997) argued that structure building, including enhancement and suppression, is a general cognitive process, so memory nodes representing any type of stimulus can be activated, combined, enhanced, and suppressed. That implies that the relatively unimportant, irrelevant information that we more consciously suppress, like certain information in word problems, can be suppressed in the same way as the irrelevant meanings of ambiguous words. Later in this paper, that implication will fuel the notion that central executive processing and the suppression of irrelevant information might causally relate.

Comprehension Ability: Relative of Suppression Ability?

When we read a text, we encounter information necessary for understanding the remainder of the discourse, but we also encounter unnecessary information. To successfully comprehend, we must be able to distinguish relevant from irrelevant information, so we can base our interpretations on the relevant information. Research suggests that the ability to suppress irrelevant information differentiates good and poor comprehenders.

Morton Ann Gernsbacher and her colleagues have claimed that poor comprehenders have difficulty suppressing the inappropriate meanings of ambiguous words (Gernsbacher, 1993; Gernsbacher & Faust, 1991; Gernsbacher & Robertson, 1995; Gernsbacher, Varner, & Faust, 1990, experiment 4). A classic experiment examined whether skilled and less-skilled comprehenders can be differentiated according to the rate at which they suppress irrelevant word meanings (Gernsbacher et al., 1990, experiment 4).

The experiment was designed in several steps. First, 80 homographic homophones (heretofore called *homographs* for short) were chosen. The homographs' two most popular meanings were considered to be equally frequent by Gernsbacher et al. (1990). Next, 80 experimental sentences were designed, each with a homographic homophone (heretofore called *homograph* for short) in the final position, e.g., *She put on the ring*. Then, 80 control sentences were derived from the experimental sentences by changing the sentence-final homographs to unambiguous words. The unambiguous words were related to or synonymous with the appropriate meanings of the ambiguous words. Thus, the control version of *She put on the ring* was *She put on the necklace*. Afterward, each experimental-control sentence pair was assigned to a test word. For the experimental-control sentence pair described above, the test word was *bell*. Test words were always related to the inappropriate meaning of the homograph in the experimental sentence and unrelated to the meaning of the control sentence. In other words, the test word *bell* related to neither the experimental nor the control sentence, but it did relate to the irrelevant meaning of *ring*. During the experiment, the test words appeared after their corresponding sentences. For example, if a subject saw *She put on the ring*, it was followed by the test word *bell*. Because of counterbalancing, that same subject did not see the corresponding control sentence, *She put on the necklace*. Therefore, each subject saw 40 experimental sentences and 40 control sentences, each followed by a test word, on a computer screen. After seeing the test word, the subject had to indicate, via key press, whether the test word related to the entire sentence that he or she had just read. For both the experimental and control sentences, the correct response was always negative because none of the test words related to their corresponding sentences.

In addition to the ambiguity of the sentence-final word, Gernsbacher et al. (1990) manipulated the time interval between the presentation of each sentence and its corresponding test word. Specifically, each sentence disappeared from the screen before its corresponding test word appeared. Sometimes, the test word appeared “immediately,” or 100 ms, after the sentence disappeared. In the “delayed” condition, the test point occurred 850 ms after the sentence disappeared.

The stimuli were counterbalanced according to both variables—sentence-final word ambiguity and test point—so each subject saw 20 experimental sentences followed immediately by their test words, plus 20 experimental stimuli in the delayed condition, 20 control stimuli in the immediate condition, and 20 control stimuli in the delayed condition.

Because all of the correct responses to the experimental and control sentences were negative, 80 “filler” stimuli were devised to enable correct, affirmative responses. Approximately half of the filler sentences represented experimental sentences because they ended with homographs, and the rest represented control sentences because they ended with unambiguous words. Half of the filler sentences were followed immediately by their test words, and half were separated from their test words by the 850-ms delay. For example, one filler stimulus was *He wanted the award* followed immediately by the test word *trophy*. Each subject saw the same 80 filler stimuli, including approximately 20 “experimental” filler stimuli in the immediate condition, 20 experimental filler stimuli in the delayed condition, 20 “control” filler stimuli in the immediate condition, and 20 control filler stimuli in the delayed condition.

Gernsbacher et al. then diagnosed skilled and less-skilled comprehenders according to their performance on Gernsbacher and Varner’s (1988) Multi-Media Comprehension Battery. Those subjects participated Gernsbacher et al.’s task (1990, experiment 4). Overall, skilled and less-skilled comprehenders responded more slowly to the experimental stimuli than the control stimuli in the immediate condition. That means that all of the subjects found it harder to respond to stimuli containing ambiguous sentence-final words. However, skilled comprehenders, unlike less-skilled comprehenders, did not respond more slowly to the experimental stimuli in the delayed condition. That means that the less-skilled comprehenders still found it difficult to respond to stimuli containing ambiguous sentence-final words, but the skilled comprehenders did not have such trouble. Gernsbacher et al. interpreted those results as an indicator that good comprehenders suppress irrelevant information more efficiently than poor comprehenders. Gernsbacher has additionally claimed that difficulty suppressing irrelevant information causes comprehension problems (Gernsbacher, 1997).

Central Executive Processing: An Abbreviated Biography

Like the picture of Dorian Gray (Wilde, 1891/1982), the conceptual portrait of the central executive has changed since its first appearance. First described as one of three components of working memory, the central executive was considered a general process that coordinates the functions of the other two components, now known as the phonological loop and the visuospatial sketchpad, and acts as a supplementary storage unit (Baddeley, 2002; Baddeley & Hitch, 1974). Eventually, its role evolved into attentional control (Baddeley, 1986, 2002). Most recent research has focused on the central executive as one or many attentional mechanisms (for discussions, see Baddeley, 2002; Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000).

The central executive has been associated with a number of high-level mechanisms called *executive functions*. Many researchers have proposed that executive functions are associated with specialized components of the central executive (Baddeley, 2002; Baddeley, Emslie, Kolodny, & Duncan, 1998; Bull, Johnston, & Roy, 1999; Lehto, 1996; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Miyake et al., 2000; Towse, 1998). Others, such as Duncan, Emslie, Williams, Johnson, and Freer (1996), refer to executive functions more abstractly.

Executive Functions' Identity Crises

Exactly how many executive functions exist, and what they all do, remains a mystery. A number of researchers have suggested a number of possible executive functions. Though some of those executive functions seem to stand alone, e.g., the inhibition of prepotent responses (Friedman & Miyake, 2004; Salthouse, Atkinson, & Berish, 2003), evidence has suggested that they are interrelated (Friedman & Miyake, 2004; Miyake et al., 2000; Salthouse et al., 2003). For example, the executive function, *planning*, might be the result of simpler executive functions working together.

Adding to the uncertainty, some have referred to an executive function called "working memory" (e.g., Nigg et al., 2004). That name was bestowed despite the classical definition of working memory as a construct that encompasses the phonological loop, the visuospatial sketchpad, and, of course, the central executive.

Working memory (the executive function) is the namesake of assessments referred to as "working memory tasks", or "tests of working memory capacity". Friedman and Miyake (2004), Miyake et al. (2000), and Salthouse et al. (2003) statistically modeled relationships among several executive functioning tasks, including some attributed to working memory. Their results implied that the working memory tasks might measure more than one executive function. It is therefore reasonable to acknowledge that working memory tasks are general tests of central executive processing, as many researchers have already done (e.g., Engle, Tuholski, Laughlin, & Conway, 1999; Lieberman & Rosenthal, 2001; Morris & Jones, 1990; Whitney, Arnett, Driver, & Budd, 2001).

The Bond Between the Central Executive and Reading Comprehension Ability

The central executive is associated with both reading comprehension ability and problem solving ability. Specifically, performance on comprehension and problem solving tasks is linked to performance on executive function tasks, namely on verbal and mathematical variations of the span tasks (for a review, see Daneman & Merikle, 1996; also see De Beni & Palladino, 2000; De Beni, Palladino, Pazzaglia, & Cornoldi, 1998; Palladino, Cornoldi, De Beni, & Pazzaglia, 2001; Passolunghi et al., 1999; Passolunghi & Siegel, 2001; Waters & Caplan, 1996; Yuill, Oakhill, & Parkin, 1989, experiment 1).

Devised to measure working memory, the first span tasks were the reading and listening span tasks (Daneman & Carpenter, 1980). The earliest version of the reading span task consisted of groups of individually presented sentences. The number of sentences in the groups varied. Subjects had to read the sentences aloud while remembering all of the sentence-final words appearing in each group. At the end of each group, the subjects were instructed to recall all of the sentence-final words they could remember from that group. The listening span task was like the reading span task, but it required the subjects to listen to the stimuli rather than read them. Scores on those tasks were called *spans*, and reading spans correlated .80 with listening spans.

Years later, Turner and Engle (1989) proposed the operation span task, which was similar to the reading and listening span tasks but involved mathematical equations instead of sentences. Overall, verbal span tasks and mathematical span tasks have produced similar results (Daneman and Merikle, 1996).

Verbal and mathematical span tasks tend to correlate moderately with measures of general comprehension, some of which assess both comprehension and vocabulary knowledge, and more with measures that specifically assess integration (Daneman & Merikle, 1996). Integration is the process by which we link what we are presently reading to what appeared earlier in the text.

Sentence memory, a construct that has been associated with children's reading comprehension ability (for a review, see Scarborough, 1998), was found to be moderately correlated with adults' performance on the *n*-back task, another central executive measure (Roberts & Gibson, 2002; for more on the *n*-back task, see Jonides et al., 1997, and Smith & Jonides, 1997; also see the next chapter). However, Roberts and Gibson (2002) also found sentence memory to be unreliably correlated with reading span and, at most, moderately correlated with a mathematical span task. Overall, span was uncorrelated with *n*-back performance.

Central Executive Processing and Suppression: A Family Resemblance?

Two executive functions, inhibition and updating (Miyake et al., 2000; Salthouse et al. 2003), appear consistent with suppression.

Describing inhibition, like describing suppression, is difficult. If inhibition is the result of suppression (Gorfein et al., 2000), there might be several types of inhibition, just as there are several types of suppression. Although no one has empirically assessed Gernsbacher et al.'s (1990, experiment 4) or Gernsbacher and Faust's (1991) suppression measures in terms of executive functioning, Wilson and Kipp (1998) theorize that the suppression of irrelevant homograph meanings represents a particular type of cognitive inhibition—*unintentional inhibition*—by which we unconsciously suppress irrelevant information. Recent theory and research support the presence of multiple inhibition mechanisms but have not concurred on how to categorize all of them (Friedman & Miyake, 2004; Harnishfeger, 1995; Nigg, 2000; Wilson & Kipp, 1998).

This lack of agreement stems partly from the notion that tasks purportedly measuring one type of inhibition often measure multiple types of inhibition, and even more likely, constructs other than inhibition. The span tasks are such measures.

Although the span tasks have been described as general assessments of central executive functioning, they appear to require inhibition, and therefore, suppression. Specifically, the conscious ignoring of relatively unimportant information that occurs during the span tasks appears similar to what occurs when we read word problems. That is because we must ignore irrelevant information, e.g., the sentences in the reading span task and the settings of the word problems, to focus on more important information, e.g., the items to memorize in the span tasks and information related to the relevant numbers in the word problems. In fact, Whitney et al. (2000) found that “susceptibility to interference” factored into performance on the reading span test, and Friedman and Miyake’s (2004) statistical models suggested that the reading span task is related to a type of inhibition called *resistance to proactive interference (resistance to PI).* Kane and Engle (2000) additionally established a relationship between operation span and resistance to PI. In contrast, however, Miyake et al.’s (2000) models suggest that the operation span might not measure the inhibition of a dominant response, but it might instead represent one or two other executive functions. Miyake et al.'s (2000) most economical statistical model involved only one of those functions, updating.

Updating is the process by which we replace irrelevant information with relevant information. More specifically, when we encounter a stream of information in which the relevant facts are constantly changing, we must actively keep track of the information as we hear it, determine what is relevant, and replace obsolete information with the most current, relevant information (Miyake et al., 2000; Morris & Jones, 1990). Imagine listening to a horse race on the radio while keeping track of the order of the three fastest horses. In a close race, the placement of the horses might change several times. Continually updating your representation of the horses' positions is more efficient than keeping track of the current and prior configurations, so each time you hear that the placement of the horses has changed, you eschew the obsolete information and update your knowledge of the race's outcome thusfar.

Perhaps eschewing irrelevant information is not very different from unintentionally suppressing it: regardless of how consciously we ignore irrelevant information, the means by which we suppress it might be the same. That idea is analogous to the theory of comprehension discussed by Gough and colleagues (Gough & Tunmer, 1986; Hoover & Gough, 1990; Gough et al., 1996), in which the comprehension mechanism is constant whether the initial input is visual or aural. In support of a universal suppression mechanism, Miyake et al. (2000) hypothesized that statistical relationships existed among executive tasks purportedly measuring updating and inhibition because they all involved the suppression of irrelevant information. It is therefore possible that suppressing irrelevant information is a general characteristic of central executive processing.

Family Ties: Suppression, Central Executive Processing, and Reading Comprehension Ability

Induction suggests that suppression and central executive processing are related. While individuals' spans have been related to performance on comprehension tests specifically measuring integration (Daneman & Merikle, 1996), Gernsbacher and colleagues' suppression tasks require subjects to determine the relationship between a test word and a sentence by integrating information from the sentence with the meaning of the test word. Additionally, the constructs measured by Gernsbacher's suppression tasks and the span tasks might be related because both tasks require the linking of present information to information appearing previously. And as aforementioned, span and performance on Gernsbacher's suppression tasks have been associated with reading comprehension ability.

But despite their commonalities, no conclusive evidence indicates whether suppression is dependent on central executive processing, or vice-versa, or whether they independently function within the realm of reading comprehension ability.

It is possible that the central executive processing might facilitate suppression, in accordance with the Central-Executive-As-Mother hypothesis. Engle and colleagues consistently argue that general central executive tasks, i.e., working memory tasks, measure controlled attention (Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2000; Engle, 2001; but see Friedman & Miyake, 2004). Kane and Engle (2000) found that high-span subjects, who perform better than low-span subjects in a number of inhibition-related tasks, were more adversely affected when they were required to divide their attention between an inhibitory task and another task. They argued that the low-span subjects were essentially unaffected by extraneous attentional loads because they already lacked the controlled attention to perform well on the inhibitory tasks. In contrast, the high-span subjects' inhibitory capabilities appeared to be worsened by an additional attentional load, implying that some of the attention that the high-span subjects normally allocated to resisting the interference of irrelevant information was diverted to the second task. Similarly, Miyake et al. (2000) mused that controlled attention might be the source of statistical relationships among multiple executive functions. In that vein, controlled attention is minimal in Gernsbacher and colleagues' suppression tasks but more abundant in central executive tasks, implying that the central executive should encompass the suppression mechanism.

Then again, suppression might help executive functions, like updating and inhibition, function, speculated Miyake et al. (2000). That speculation implies the Central-Executive-As-Daughter hypothesis. Prior research has suggested that the ability to inhibit irrelevant information influences central executive processing, as measured by span (Hasher & Zacks, 1988; May, Hasher, & Kane, 1999).

Alternatively, studies of central executive processing—as measured by span—and suppression—as measured by the Stroop task—favor the Sisters hypothesis, i.e., the independence of central executive processing and the suppression of irrelevant information. Specifically, span and Stroop performance have both been associated with scores on nonverbal- and performance-intelligence measures (for discussions, see Dempster & Corkill, 1999; Engle et al, 1999). Span and Stroop performance have additionally exhibited small, though significant, correlations with each other (see, e.g., Dempster & Corkill, 1999; Engle et al., 1999; Friedman & Miyake, 2004; Miyake et al, 2000). Dempster and Corkill (1999) and Engle and colleagues (1999) have suggested that such results might reflect individual differences in central executive processes. Nonetheless, no definitive evidence confirms that differences in general, fluid intelligence, as assessed by nonverbal and performance measures, are due to central executive processing differences. The relationship could be inverted. That is, similar elements of suppression and central executive processing might be byproducts of general intelligence, which, in this study of reading comprehension ability, is part of the Given. If that notion is correct, accounting for intelligence before examining the two mechanisms in terms of reading comprehension ability might yield results implying independent effects of each.

Because no evidence has falsified the existence of a causal association, researchers continue to speculate on the directionality of the relationship between suppression and central executive processing. Such speculation might perpetuate because that relationship is interactive rather than unidirectional. Or perhaps the two mechanisms share common variance that is related to reading comprehension ability but is a byproduct of a construct outside the Given. Inconclusive results implying these conjectures would favor the Conjoined Sisters hypothesis.

The issues discussed here are complex. Suppression is a complex mechanism, and central executive processing is perhaps more complex. Although suppression, central executive processing, and reading comprehension ability were rarely studied simultaneously (see Borella and de Ribaupierre, 2001; 2003), the available research has provided evidence in favor of distinct hypotheses concerning those three general constructs (see Figure 1). Accurately testing these hypotheses requires multiple measures.

Physical Exams: Rationales for Multiple Measures

The present research used multiple measures to assess central executive processing, suppression, and decoding ability for two reasons. First, convergent validity might have implied that an assessment purportedly measuring one construct might have been more valid than other assessments due to a higher correlation with reading comprehension. Multiple measures gave me the option to choose some measures over others. Second, multiple measures enhanced the statistical probability that my research questions would be answered—there were "fallback options" in case unexpected problems arose during the study and invalidated a measure.

Because reading comprehension ability is the centerpiece of this research, three reading comprehension measures were carefully chosen to thoroughly examine that construct.

Nelson-Denny Reading Comprehension Subtest

This popular test of general reading comprehension ability has been moderately correlated with the written subtest of Gernsbacher & Varner's (1988) Multi-Media

Comprehension Battery ($r = .46$, $p < .05$, Maki, Jonas, & Kallod, 1994), which was used to determine comprehension ability in Gernsbacher and colleagues' suppression studies. Normed for college students (The Riverside Publishing Company, 2004), the Nelson-Denny reading comprehension subtest consists of several passages, each followed by multiple-choice items. Many of the correct answer choices can be matched to phrases within the passages, meaning that little integration and inference is required.

The Nelson-Denny reading comprehension subtest (Brown, Fishco, & Hanna, 1991) is a timed test. It is also a purely *speeded* assessment, meaning that that an examinee with enough time to attempt an item will almost certainly mark the correct answer choice.

Advanced Degrees of Reading Power

This measure, normed for grades 6-12 (Touchstone Applied Science Associates, 2002, p. 51), was chosen because of its dissimilarity to the Nelson-Denny reading comprehension subtest. While the Nelson-Denny test requires little or no integration, publishers of the *Advanced Degrees of Reading Power* (Touchstone Applied Science Associates, 1995) claimed that the items were "designed to access the ability to integrate propositions over ever-increasing amounts of text" (Touchstone Applied Science Associates, 2002, p. 31). Because Daneman and Merikle (1996) found central executive processing to be more highly correlated with integration ability than general comprehension ability, the *Advanced Degrees of Reading Power* might answer the theoretical research questions differently than the Nelson-Denny.

Another difference between the *Advanced Degrees of Reading Power* and the Nelson-Denny stems from their design. While the Nelson-Denny is a timed, purely-speeded test, the *Advanced Degrees of Reading Power* is an untimed, purely *power* test. A score on a power test is purely dependent on the number of correct responses; a score on a speeded test depends on the amount of progress made by the examinee.

Like the Nelson-Denny reading comprehension subtest, the *Advanced Degrees of Reading Power* consists of several passages, each followed by multiple-choice items. But the Nelson-Denny requires examinees to respond to items about the passages, whereas the *Advanced Degrees of Reading Power* requires examinees to choose sentences to fit within the passages.

Semantic-decision task

Unlike the Nelson-Denny and *Advanced Degrees of Reading Power* tests, the semantic-decision task is a computerized task that follows a similar procedure to other computerized laboratory tasks that were also utilized in the present research. Gough and his colleagues found a correlation of approximately .60 between college students' performance on the semantic-decision task and the Nelson-Denny reading comprehension subtest (2002).

Research Questions: Locating Branches on the Family Tree

The present research seeks to answer four important, theoretical questions, which are illustrated in Figure 1. Through that process, it also asks three less-important, methodological questions.

Theoretical Questions

1. Does the suppression mechanism used during reading comprehension stem from central executive processing, according to the Central-Executive-As-Mother hypothesis?

2. Or do the central executive processes associated with successful reading comprehension stem from the suppression mechanism, according to the Central-Executive-As-Daughter hypothesis?

3. Or do general central executive processing and the suppression of irrelevant information independently predict reading comprehension ability, according to the Sisters hypothesis?

4. Or alternatively, are the two constructs inconclusively related, according to the Conjoined Sisters hypothesis?

The theoretical questions could be explored in a multi-step process, first by performing an exploratory principal components factor analysis with an oblique rotation on the non-comprehension measures. The oblique rotation was chosen because some of the factors, namely those representing central executive processing and suppression ability, should be related according to the evidence presented in this chapter. Then, another principal components factor analysis could be used to obtain a composite reading comprehension ability score. Afterward, two stepwise multiple regressions could be performed with reading comprehension ability as the dependent variable. To account for the influence of the Given and decoding ability, any factors representing either should be entered first into each regression. Then, in one regression, the central executive processing factor, followed by the suppression ability factor, would be entered. In the other regression, the suppression ability factor would be entered prior to the central executive processing factor. If the suppression ability factor accounted for an insignificant amount of the remaining reading comprehension ability variance in the first regression but significantly accounted for variance in the second regression, the results would support the Central-Executive-As-Mother hypothesis. If the central executive processing factor significantly accounted for reading comprehension ability variance in the first regression but not the second, the results would support the Central-Executive-As-Daughter hypothesis. If the order in which the factors were entered into each regression made no difference, i.e., each consistently accounted for the same amount of reading comprehension ability variance in each regression, the results would support the Sisters hypothesis. Lastly, if each factor accounted for a significant amount of reading comprehension ability variance in both regressions, but the amount depended on the order in which each factors was entered, the results would support the Conjoined Sisters hypothesis.

Methodological Questions

1. How well do multiple reading comprehension measures correlate?
2. How similar are certain measures of the central executive?
3. How similar are certain measures of suppression?

Chapter Two: Method

Participants

Upperclassmen (N = 101, 27 males, 74 females) from Philip Gough's Psychology of Reading class participated in the study as part of the requirements for the course. Twenty-nine subjects were removed from the analysis. One subject was eliminated for failing to follow directions on multiple tasks. Criteria for the removal of the other 28 subjects were the ability to read non-English text as well as or better than English text, spending fewer than five secondary school years in the United States[1], colorblindness, difficulty learning to read, a diagnosis of dyslexia or attention problems, attention medication use, self-reported below-average test-taking skills, and self-reported test anxiety. Participation was part of the course requirement, so all of the students were asked to participate in all of the tasks; in a traditional recruitment paradigm, subjects meeting any of the aforementioned criteria would not have participated.

Of the 72 subjects (19 males, 53 females) that remained in the analysis, three were paid $5 to retake the *Nelson-Denny Test of Reading Comprehension* (Brown et al., 1991). All subjects completed either form G or H of the Nelson-Denny, but data from two men were lost, and one woman reported prior exposure to the Nelson-Denny form she had completed. The woman and one man had initially completed Form H, so they retook the Nelson-Denny using form G. The other man initially completed form G, so he retook the Nelson-Denny using Form H. In addition, one woman did not take Raven's *Advanced Progressive Matrices* (Raven, 1947), a test of nonverbal intelligence, meaning that only

---

[1] The semantic-decision and which-came-first tasks contained items related to United States history and culture.

71 subjects completed that test. Finally, seven subjects' (2 males', 5 females') response patterns on one or more *n*-back subtasks indicated that they misunderstood the instructions. Those subjects retook the *n*-back.

Tasks and Their Administration

Subjects completed the personal-information survey in Appendix A, form G or H of the Nelson-Denny reading comprehension and vocabulary subtests (here abbreviated NDComp and NDVocab, respectively), form T-2 of the *Advanced Degrees of Reading Power* (ADRP), the semantic-decision task (sem), sections I and II of Raven's *Advanced Progressive Matrices* (Raven's APM), the which-came-first task (source unknown, n.d.), the motor speed task, the lexical-decision task, the pseudoword-decoding task, the Stroop task, the homograph-suppression task, the homophone-suppression task, the reading span task (RSpan), the operation span task (OSpan), and the *n*-back task.

The sem, which-came-first, motor-speed, pseudoword-decoding, lexical-decision, RSpan, and OSpan tasks were performed using personal computers. For the sem, lexical-decision, and which-came-first tasks, the *z* key on each keyboard was marked with a circular, red sticker, and the */* key on each keyboard was marked with a circular, green sticker. The experimenters referred to the *z* and */* keys as the *red button* and *green button*, respectively.

The Stroop, homograph-suppression, homophone-suppression, and *n*-back tasks were performed on Macintosh computers using the SuperLab program. The *x* key on each keyboard was marked with a circular, red sticker, and the . key on each keyboard was marked with a circular, green sticker. The experimenters referred to the *x* and . keys as the red button and green button, respectively.[2,3]

Design and Procedure

General Procedure for Computerized Tasks

For all computerized tasks except the RSpan, OSpan, and pseudoword-decoding tasks, subjects were instructed to press the red button with their left index finger and the green button with their right index finger. Subjects were instructed to keep their index fingers on the buttons throughout timed tasks. Stimuli appearing on the PC monitors were printed in white and centered on a black background. Non-Stroop stimuli appearing on the Macintosh monitors were printed in black and centered on a white background.

Nelson-Denny Test of Reading Comprehension (NDComp and NDVocab)

The NDComp, a reading comprehension measure, consisted of seven passages followed by multiple-choice items: the first passage spanned one page and was followed by eight items; each of the remaining six passages was 2-4 paragraphs long and followed by five items. The NDVocab, the vocabulary knowledge measure, consisted of 80 words, each followed by five, one-word answer choices. Subjects marked the answer choice that was most synonymous with the original word.

---

[2] Some of the PC tasks were run before the Macintosh tasks were programmed into SuperLab. During a test run of the homograph-suppression task, the author noticed that all of the */* key's responses, correct or incorrect, were marked as erroneous by SuperLab. To rectify the problem, the red and green buttons were moved inward by one key on the Macintosh keyboards.

[3] The Stroop, suppression, and *n*-back tasks were not performed on the PCs due to limitations in the PC program used to perform experiments in the lab. The other tasks were not performed using SuperLab because of a limited number of Macintoshes in the lab. The setup described allowed the experimenters to run all of the subjects in these and related experiments within a limited number of sessions and a limited amount of time.

The NDComp and NDVocab were administered during class, with make-up exams individually administered in the laboratory. Subjects were given 15 minutes to take the NDVocab, followed by 12 minutes[4] to take the NDComp. Responses were recorded on UT-Austin general-purpose answer sheets. Of the 72 subjects whose data were analyzed, 37 were given Nelson-Denny form G, while 35 were given form H.

ADRP

The ADRP, an untimed, reading comprehension measure, was composed of eight, 4-5 paragraph passages. Three paragraphs within each passage ended with a blank, which signified a missing sentence. Corresponding to each missing sentence were three multiple-choice items, which consisted only of five sentences. The subjects were instructed to choose the sentence that best fit in each blank. Each subject completed the ADRP in the classroom, the laboratory, or a convenient location. Responses were recorded on UT-Austin general-purpose answer sheets.

Sem

Following 10 practice sentences, subjects viewed 100 short sentences that varied in length and complexity (see Appendix B). The sentences were randomized for each subject and appeared individually on a computer screen.

Fifty of the sentences in this reading comprehension measure were devised to be obviously true, e.g., “A blind man cannot see.” The other fifty sentences were devised to be obviously false, e.g., “Most of us keep diamonds in the dishwasher.” The subjects were instructed to press the green button if the sentence was true or the red button if the

---

[4] Examinees are typically given 20 minutes to take the NDComp, but to maximize score variance, these upper-level undergraduates were given less time.

sentence was false. They were also told to respond as quickly as possible without sacrificing accuracy. To reduce rhythmic responding, a random interstimulus interval of 800-1,200 ms followed each response. Subjects had 10 seconds to respond to each sentence.

Raven's APM

Sets I and II of Raven's APM, a popular test of non-verbal intelligence, were administered. Set I consisted of 12 items, and set II consisted of 36. Each item included a 3 x 3 matrix of shapes and lines that together formed a pattern. The cell in the bottom, right-hand corner of each matrix was missing. Subjects had to locate each missing cell among eight incorrect answer choices. Responses were recorded on specialized answer sheets.

Subjects were administered both sets of Raven's APM during one of many 90-minute sessions. The test was untimed, so some students attended multiple sessions to complete it.

Which-Came-First

The which-came-first task was used to measure general knowledge. Subjects viewed the names of 50 individual pairs of historical people and events (see Appendix C). The members of each pair were adjacent to each other and separated by four spaces. Although the placement of each pair's members was not counterbalanced across subjects, the order of the pairs was randomized for each subject.

The subjects indicated which member of each pair came first historically. Subjects were instructed to press the red button if they believed that the leftmost item came first or the green button if they believed that the rightmost item came first. All but three subjects were instructed to press the spacebar if they were completely unsure of which came first. The subjects were advised to think about their answers and were given 20 seconds per response. A 1,500-ms interstimulus interval followed each response.

Motor Speed

Because many of the computerized measures required the subjects to respond quickly, their reaction times might have reflected motor speed as well as cognitive abilities. Measuring motor speed allowed me to statistically account for it while analyzing the experimental tasks.

Subjects were administered identical tasks for the right and left hands. The order in which the subjects performed these tasks depended on the order in which they arrived at the lab, with every other student participating in the right-hand condition first.

The tasks consisted of 30 strings of seven asterisks, the first five of which were practice trials. To better simulate a typical, speeded task, subjects placed their right index finger on the green button and their left index finger on the red button. In the right-hand condition, subjects were instructed to immediately press the green button each time the asterisks appeared. In the left-hand condition, subjects were instructed to immediately press the red button each time the asterisks appeared. To reduce rhythmic responding, a random interstimulus interval of 800-1,200 ms followed each response.

Lexical Decision

The lexical-decision task was used to measure decoding ability. Following 10 practice trials, subjects viewed 25 five-letter words and 25 five-letter pseudowords on a computer screen (see Appendix D). The stimuli were randomized for each subject. The subjects were instructed to respond as quickly and accurately as possible by pressing the green button if the stimulus was a word or the red button if the stimulus was not a word. To reduce rhythmic responding, a random interstimulus interval of 800-1,200 ms followed each response. Subjects were given 10 seconds to respond.

Pseudoword Decoding

A measure of decoding ability, this untimed task consisted of 25 pseudowords (Wren, 1995), some with more than one correct pronunciation. The order of the pseudowords was randomized for each subject, and the subjects were instructed to decode each pseudoword aloud. An experimenter recorded each subject’s responses on a form, like the one in Appendix E, and then hit a key to present the next pseudoword. The experimenters also recorded when a subject slowly attempted to sound out a pseudoword.

Stroop

The Stroop task was used to measure suppression ability. The stimuli appeared in four colors: red, blue, yellow, and green. Forty-eight stimuli were neutral, consisting of individual strings of five asterisks. Forty-eight stimuli were incongruent, meaning that the words did not describe the colors in which they were printed, e.g., *blue* printed in red. Eight stimuli were congruent, meaning that the words described the colors in which they were printed, e.g., *blue* printed in blue.

Subjects viewed ten practice items, followed by all 104 stimuli. Subjects were instructed to ignore the words on the screen and loudly name the text’s color into the microphone. They were additionally instructed to name those colors as quickly as possible without sacrificing accuracy.

The stimulus order was randomized once, and all subjects saw the stimuli in the same random order. Each response was followed by a 1,000-ms interstimulus interval, as the program in which the experiment was run could not incorporate random interstimulus intervals.

An experimenter remained with the subject to ensure an understanding of the task and record incorrect responses. Correct responses to the items were listed on a form, and the experimenter marked the items that did not match the subjects' responses. To help the experimenters keep track of which response corresponded with each stimulus, the stimuli were divided into blocks. The 10 practice items were their own block, and the 104 experimental stimuli were divided into 10 blocks of 10 items and one block containing the final four items. Between each block was an intermission in which the screen instructed the subject to ask the experimenter if he or she was ready to continue. After confirmation from the experimenter, the subject pressed a key. The instructions disappeared, and after 1,500 ms, the next block of stimuli began.

Homograph Suppression

This suppression task was based on Gernsbacher et al.'s (1990) fourth experiment, which was described in the first chapter. The present experiment utilized homographs whose meanings were considered fairly equiprobable by Gernsbacher and her colleagues (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990). Sixty experimental sentences, 60 corresponding control sentences, 60 test words corresponding to each experimental-control sentence pair, and 60 filler stimuli were taken or modified from Gernsbacher et al.'s (1990) experiment and a similar experiment by Gernsbacher and Faust (1991, experiment 4).

General design. Each experimental sentence had a homograph in the final position, e.g., *She put on the ring*. Each of the 60 control sentences was identical to its corresponding experimental sentence, but instead of ending with a homophone, it ended with an unambiguous word that was related to or synonymous with the appropriate meaning of the ambiguous word. Thus, the control version of *She put on the ring* was *She put on the necklace*. Each experimental-control sentence pair was assigned to a test word. For the sentence pair described above, the test word was *bell*. Test words were always related to the inappropriate meaning of the homograph in the experimental sentence and unrelated to the meaning of the control sentence. In other words, the test word *bell* related to neither the experimental nor the control sentence, but it did relate to the inappropriate meaning of *ring*. During the experiment, the test words appeared after their corresponding sentences. For example, if a subject saw *She put on the ring*, it was followed by the test word *bell*; however, that same subject did not see the corresponding control sentence, *She put on the necklace*. Therefore, each subject saw 30 experimental sentences and 30 control sentences, each followed by a different test word, on a computer screen. The subjects had to press a key to indicate whether they thought a test word was related to the sentence that preceded it. The subjects should have responded negatively to all of the experimental and control stimuli because no test word related to its corresponding sentence, regardless of the sentence-final word.

In addition to the ambiguity of the sentence-final word, a second variable was manipulated: the time interval between the presentation of the sentence and the presentation of the test word. Specifically, each sentence disappeared from the screen before its corresponding test word appeared. Sometimes, the test word appeared immediately, or 100 ms, after the sentence disappeared. In the delayed condition, the test point occurred 1,000 ms after the sentence disappeared.

The stimuli were counterbalanced according to both variables—sentence-final word ambiguity and test point—so each subject saw 15 experimental sentences followed immediately by their test words, plus 15 experimental stimuli in the delayed condition, 15 control stimuli in the immediate condition, and 15 control stimuli in the delayed condition. This counterbalancing resulted in four "forms," and the subjects were assigned to forms according to the order in which they entered the lab. A subject with one form saw, *She put on the ring*, followed immediately by *bell*. A subject with a different form saw, *She put on the necklace*, followed immediately by *bell*. Yet another subject saw, *She put on the ring*, followed by a 1,000-ms delay and then *bell*. A fourth subject saw, *She put on the necklace*, followed by a 1,000-ms delay and then *bell*. Since *bell* related to none of the sentences, the appropriate response to those stimuli was pressing the red button.

Because all of the correct responses to the experimental and control sentences were negative, 60 "filler" stimuli were used to enable correct, affirmative responses. Approximately half of the filler sentences represented experimental sentences because they ended with homographs, and the rest represented control sentences because they ended with unambiguous words. Half of the filler sentences were followed immediately by their test words, and half were separated from their test words by the 1,000-ms delay. For example, one filler stimulus was *He wanted the award* followed immediately by the test word *trophy*. Each subject saw the same 60 filler stimuli, including approximately 15 “experimental” filler stimuli in the immediate condition, 15 experimental filler stimuli in the delayed condition, 15 “control” filler stimuli in the immediate condition, and 15 control filler stimuli in the delayed condition.

A list of filler, experimental, and control sentences and their corresponding test words appears in Appendix F.

Stimulus design and experimental procedure. The stimuli were presented to minimize the effects of reading speed. Specifically, each word in each sentence was presented individually. Although the first word in each sentence was capitalized, the sentences did not end with punctuation marks.

Before the presentation of each sentence, subjects saw a plus sign in the center of the screen for 500 ms. Following a 150-ms interval, the words of each sentence individually appeared in the same place as the plus sign, with a 150-ms interval between the words. Each word’s presentation duration was based on the following function, rounded to the nearest millisecond:

$$\text{duration (ms)} = 300 + 16.7 \times \text{number of characters}.$$

After the conclusion of each sentence, its corresponding test word appeared immediately or after a 1,000-ms delay. The test word was capitalized and flanked on both sides by a group of two asterisks separated from the word by a space. For example, the subjects saw the test word *bell* as "** BELL **". The subjects were given 10 seconds to respond as quickly as possible without sacrificing accuracy. Pressing the green button indicated a positive response; pressing the red button indicated a negative response. Each response caused the test word to disappear, and visual feedback on the response's correctness subsequently appeared for 750 ms. Another stimulus followed the feedback.

Ten practice stimuli preceded the 120 stimuli described above. The presentation of those experimental, control, and filler stimuli was randomized for each subject.

Homophone Suppression

A modification of Gernsbacher and Faust's experiment (1991, experiment 1), the homophone-suppression task was identical in purpose, design, and procedure to the homograph-suppression task mentioned above. Each of 60 test words corresponded to an experimental-control sentence pair. The experimental sentence ended with a homophone, and the control sentence ended with a word related to or synonymous with the homophone. Approximately half of the sixty filler sentences also ended with homophones. A complete list of filler, experimental, and control sentences and their corresponding test words appears in Appendix G.

All subjects participated in this task after completing the homograph-suppression task. That was because the sentences in the present task were only disambiguated by their sentence-final words. That means that, hypothetically, the subjects could have performed well on the task by focusing only on the sentence-final words. In contrast, to understand the meaning of a sentence in the *homograph*-suppression task, reading all of the sentence's words was necessary. Therefore, any subjects who participated in the present task prior to the homograph-suppression task were in danger of assuming that sentence-final words could disambiguate the sentences in the homograph-suppression task. Such expectations could have produced task-order effects. Therefore, all subjects participated in the homograph-suppression task before the homophone-suppression task.

All subjects exposed to form 1 of the homograph-suppression task were exposed to form 1 of the homophone-suppression task, with the exception of one female subject, who was exposed to form 1 of the homograph-suppression task and form 2 of the homophone-suppression task. Although the subjects could only disambiguate the homophones according to their spellings, they were sufficiently common, ensuring that spelling ability would have little or no effect on performance.

RSpan and OSpan

Both span tasks used in the present study were modified from existing tests of the central executive (Engle, 2003a, 2003b). A RSpan stimulus was a multi-clause sentence followed by a noun, and an OSpan stimulus was a one-part addition or subtraction equation followed by a noun. For example, a RSpan stimulus was as follows,

When it is cold, my mother always makes me wear a cap on my head. ? BOX

and an OSpan stimulus was

DOES 4 + 2 = 6 ? MAP

Half of the RSpan sentences made sense, but half were nonsensical, like the sentence in the stimulus below.

All parents hope their list will grow up to be intelligent. ? COW

Half of the OSpan equations were correct, but half were incorrect, like the equation in the stimulus below.

DOES 6 + 1 = 15 ? PET

Each span task consisted of 42 stimuli, which were collated into 12 groups. Specifically, there were three groups of two stimuli, three groups of three stimuli, three groups of four stimuli, and three groups of five stimuli. They were preceded by three practice groups, each consisting of two stimuli, that were not scored.

The two tasks were nearly identical (see Appendix H for stimulus lists). First, both tasks' stimuli were presented in a fixed, quasi-random order. In addition, the order of the groups in the RSpan task corresponded to the order of the groups in the OSpan task, i.e., both tasks began with a three-stimulus group, followed by a five-stimulus group. Within those groups, the placement of true and false sentences and equations corresponded. The nouns following each sentence or equation were yoked in frequency, as measured by U statistics appearing in the CD-ROM edition of *The Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1996). The concreteness of the nouns was estimated to ensure that the nouns in one task were approximately as concrete as those in the other task.

To prepare the subjects for the span tasks, they first participated in a brief, word-memory task modeled on the span tasks but lacking sentences or equations. The task employed ten nouns divided into three groups, one containing two nouns, the others five and three. The nouns in each group were presented individually, in capital letters, on a computer screen. Subjects were instructed to read each noun aloud, as soon as it appeared onscreen, and remember it. An experimenter in the room with each subject pressed the spacebar as soon as each noun was named. Subjects were told that the end of each group would be signaled by a string of three question marks, and upon seeing it, they should record the group's nouns in order. They were also informed that there was no guessing penalty. Responses were recorded on forms like the one in Appendix I. A form contained a row for each group and six blanks in each row, each blank representing a noun.[5]

Both span tasks were exactly like the word-memory task, except the nouns were separated by sentences or equations to which the subjects had to respond. For example, as shown earlier in this section, the RSpan task required subjects to view individual sensical and nonsensical sentences, each followed by a question mark and a capitalized noun. The subjects were instructed to read each sentence aloud as soon as it appeared on the screen, then say, "yes," if the sentence made sense or "no" if it was nonsensical, and immediately read the noun aloud. They were also instructed on the importance of making correct affirmative and negative responses and remembering the noun that followed each sentence. The subjects were additionally told that they could take as long as they needed to decide whether each sentence made sense, but as soon as they said, "yes" or "no," they had to immediately name the noun following the sentence. For example, given,

We were fifty lawns out at sea before we lost sight of land. ? CHART

---

[5] Because evidence (Daneman & Carpenter, 1980; De Beni & Palladino, 2000; De Beni et al., 1998; Palladino et al., 2001) indicates that Short-Term Memory capacity, or the number of stimuli a person can remember, is less related to comprehension ability than performance in span tasks, the task was not scored.

the correct response was, “We were fifty lawns out at sea before we lost sight of land…no chart.”

As a subject responded to a stimulus, the experimenter quickly noted the subject’s yes-or-no response and pressed the space bar to present the next stimulus. After each group, the subject saw a string of three question marks in the center of the screen, signaling him or her to record the group’s nouns in order on a form like the one in Appendix J. The form contained one row per group and five blanks per row, each blank representing a noun.

The OSpan task was very similar to the RSpan task. Given,

DOES 3 + 1 = 2? CLOUD

the correct response was, “Does three plus one equal two?…no cloud.” Subjects recorded the nouns they memorized on a form like the one used in the RSpan task.

Although all the subjects completed the word-memory task first, approximately half of the subjects participated in the RSpan task before the OSpan task. The task order was determined according to the order in which the subjects entered the lab.

N-back

Reaction time and accuracy were both measured in this central executive task, which was divided into three subtasks, the 2-back, 3-back, and 4-back. The foundation of the *N*-back was the stimuli, made up of thirteen 3- and 4-letter, high-frequency, high-imagery nouns taken from Paivio, Yuille, and Madigan (1968). The stimulus words were *arm*, *book*, *boy*, *camp*, *cat*, *door*, *girl*, *gold*, *home*, *meat*, *ship*, *sky*, and *tree*.

In each subtask, the stimuli appeared individually on a computer screen. Subjects had to determine whether each stimulus appeared *n* stimuli prior. Subjects were instructed to press the green button if the stimulus appeared *n* stimuli ago and the red button if it did not.

For example, in the 2-back subtask, the stimuli might have appeared in this order:

girl arm girl tree.

The subject should have responded negatively to the first stimulus, *girl*, because it did not appear two stimuli ago. The subject should have responded negatively to the second stimulus for the same reason. The subject should have responded affirmatively to the third stimulus because it matched the first stimulus, which appeared two stimuli before. Finally, the subject should have responded negatively to the fourth stimulus, *tree*, because it did not match the second stimulus, *arm*.

Below are examples of stimuli that might have appeared in the 3-back and 4-back subtests. The stimuli in brackets merit affirmative responses.

3-back: *boy camp tree* [*boy*] *ship* [*tree*]

4-back: *home meat sky cat* [*home*] *book*

Each subtask consisted of five practice stimuli followed immediately by 30 experimental stimuli. The stimuli appeared in the same order for each subject, and each subject performed the subtasks in the same order, beginning with the 2-back subtask and ending with the 4-back subtask. In each task, fifty percent of the preferred responses were affirmative.

Subjects were given 2,500 ms to respond to each stimulus and 500 ms between stimuli, as they were in Roberts and Gibson's (2002) study. Although they were instructed to respond as quickly as possible without sacrificing accuracy, they were warned that each stimulus would only disappear after it had been onscreen for 2,500 ms. That was to ensure that they had enough time to look at and remember each stimulus.

Chapter Three: Results

Scoring

Scores were obtained for all 101 subjects, but only scores from 72 subjects were analyzed.

Untimed Tasks

NDComp, NDVocab, ADRP, Raven's APM. The number of correct responses was determined for each test. Correct responses from both sets of Raven's APM were added to form a composite Raven's APM score.

Which-came-first. Most of the subjects' scores consisted of the number of correct responses. Scores for the three subjects who were not told to press the spacebar when they would otherwise guess were determined by subtracting the number of incorrect responses from the number of correct responses. These "corrected" scores were similar in value to other subjects' scores.

Pseudoword decoding. Each pseudoword pronounced correctly on the first attempt was worth 1 point. Each correctly pronounced pseudoword that was sounded out slowly, as well as each pseudoword pronounced correctly on the second attempt, was worth .5 point. A pseudoword that was skipped, incorrectly pronounced, or pronounced correctly on the third try was worth 0 points.

RSpan and OSpan. The design of the RSpan and OSpan forms (see Appendix J for a sample) allowed the RSpan and OSpan tasks to be scored in two ways, thus obtaining for each subject an absolute and total span for each task (see La Pointe & Engle, 1990). On the form, each row represented a group, and each blank represented a word. If all the words in a group were recorded in their appropriate blanks, that group was perfectly recalled. Each of a subject's absolute spans was computed by summing the number of words in only the perfectly recalled groups. In contrast, each of a subject's total spans was computed by summing all of the words appearing in their correct blanks, regardless of whether any of the words were part of a perfectly recalled group.

Timed Tasks

General procedure for scoring a timed task. Reaction times representing incorrect responses were dropped from the analysis. Then, each subject's score was determined by computing the median of the remaining data. Tasks with additional scoring specifications appear below.

Motor speed. Scores for the right and left hands were determined separately.

Lexical decision. Only responses to words were used in the analysis.

Stroop. First, stimuli-specific scores were determined for the neutral and incongruent stimuli. Stroop interference scores were computed by subtracting each subject's neutral-stimuli score from his or her incongruent-stimuli score. The Stroop-interference scores were entered into the final analysis.

N-back. The *n*-back task was scored twice, once for speed and once for accuracy.

Subtask accuracy scores were determined by subtracting the number of incorrect responses from the number of correct responses and then dividing the quotient by the number of possible responses, thus producing a percent-correct score corrected for guessing. Composite *n*-back accuracy scores were computed with the following equation (see Roberts & Gibson, 2002):

N-back Accuracy = 1 + (2-back, % correct) + (3-back, % correct) + (4-back, % correct)

Subtask speed scores were computed by simply obtaining each subject's median reaction time for correct responses on each subtask. A composite speed score was computed by obtaining the median reaction time of all of the correct responses.

Homograph and Homophone Suppression

A main expectation of these tasks was an indication that poorer comprehenders would have more difficulty suppressing irrelevant word meanings than better comprehenders. Each task was analyzed once in conjunction with each of the three comprehension measures. Subjects performing in the top and bottom third (n = 24 per group) of each comprehension measure were used in each analysis. Low performers on the NDComp had scores ranging between 13 and 20, and high performers' scores ranged from 25 to 37. Low performers on the ADRP had scores ranging between 8 and 17, and high performers' scores ranged from 20 to 24. Low performers, i.e., slow responders, on the sem had median reaction times between 1,866.5 and 2,808.0 ms, and high performers' median reaction times ranged from 1,198.5 to 1,576.5 ms. Because the comprehension measures were moderately or unreliably correlated ($r_{\text{NDComp, sem}} = .42$, $p < .001$; $r_{\text{NDComp, ADRP}} = .17$, $p = .149$; $r_{\text{ADRP, sem}} = .19$, $p = .119$), each analysis compared different groups of subjects.

Tables 1-6 show, for high and low performers on each comprehension measure, mean median reaction times for test words presented after experimental and control sentences—which ended with ambiguous and unambiguous words, respectively—at the immediate and delayed test points. Within the immediate and delayed conditions of each

Table 1

Low and High NDComp Performers' Mean Median Homograph Suppression Times (ms)

	Low NDComp		High NDComp	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	834 (156)	775 (158)	740 (151)	711 (137)
Delayed	784 (182)	761 (150)	707 (132)	683 (117)

Note. Numbers in parentheses are standard deviations. NDComp = Nelson-Denny Reading Comprehension Subtest; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

Table 2

Low and High ADRP Performers' Mean Median Homograph Suppression Times (ms)

	Low ADRP		High ADRP	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	745 (150)	736 (168)	795 (182)	746 (168)
Delayed	711 (132)	700 (119)	722 (157)	716 (152)

Note. Numbers in parentheses are standard deviations. ADRP = Advanced Degrees of Reading Power; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

Table 3

Low and High Sem Performers' Mean Median Homograph Suppression Times (ms)

	Low Sem		High Sem	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	905 (136)	877 (122)	687 (126)	648 (86)
Delayed	862 (151)	820 (113)	649 (114)	620 (71)

Note. Numbers in parentheses are standard deviations. Sem = Semantic Decision; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

Table 4

Low and High NDComp Performers' Mean Median Homophone Suppression Times (ms)

	Low NDComp		High NDComp	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	1,012 (265)	931 (199)	870 (207)	798 (144)
Delayed	926 (214)	903 (228)	783 (140)	783 (166)

Note. Numbers in parentheses are standard deviations. NDComp = Nelson-Denny Reading Comprehension Subtest; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

Table 5

Low and High ADRP Performers' Mean Median Homophone Suppression Times (ms)

	Low ADRP		High ADRP	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	893 (218)	837 (147)	947 (258)	874 (214)
Delayed	822 (172)	808 (182)	860 (206)	821 (192)

Note. Numbers in parentheses are standard deviations. ADRP = Advanced Degrees of Reading Power; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

Table 6

Low and High Sem Performers' Mean Median Homophone Suppression Times (ms)

	Low Sem		High Sem	
Test Point	Ambig. SFW	Unambig. SFW	Ambig. SFW	Unambig. SFW
Immediate	1,136 (220)	1,010 (183)	770 (131)	743 (105)
Delayed	1,004 (181)	970 (174)	726 (126)	709 (120)

Note. Numbers in parentheses are standard deviations. Sem = Semantic Decision; Ambig. = Ambiguous; SFW = sentence-final word; Unambig. = Unambiguous.

task, interference was computed by subtracting each subject's median reaction time for "control" test words from each subject's median reaction time for "experimental" test words. Figure 2 represents expected results, based on those of Gernsbacher and Faust (1991, experiment 1) and Gernsbacher et al. (1990, experiment 4). A reliable amount of interference was expected in the immediate condition for both low and high performers. In the delayed condition, low performers, but not high performers, were expected to show a reliable amount of interference. Figures 3-8 compare mean homograph and homophone interference levels for each comprehension measure.

Here, and in similar studies by Gernsbacher and colleagues (Gernsbacher et al., 1990; Gernsbacher & Faust, 1991), the statistical significance of interference levels was examined via one-way, repeated-measures ANOVAs. For example, an ANOVA was used to determine whether high performers on the ADRP experienced a reliable amount of interference in the immediate condition. Specifically, that ANOVA compared reaction times for test words that followed experimental and control sentences. A different one-way ANOVA was used to determine whether high performers on the ADRP experienced a reliable amount of interference in the delayed condition.

Furthermore, a 2 x 2 x 2 (performance level x sentence-final word x test point) factorial ANOVA was expected to result in a 3-way interaction indicating that only high performers reacted faster at the delayed test point than they did at the immediate test point.

This is not the first time in which Gernsbacher and colleagues' studies could not be replicated (Watts & Gough, 1995), and of the six analyses described below, only one, which examined the relationship between NDComp performance and homophone suppression, produced near-expected results (for expected results, see Figure 2; for actual, near-expected results, see Figure 6). No reliable three-way interactions were obtained from any the analyses. However, in accordance with Gernsbacher and colleagues (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990), the subjects, overall, reacted slower in the immediate condition than the delayed condition, and slower to the experimental sentences than the control sentences (Tables 1-6). Nonetheless, those differences were not always outside the realm of the standard errors.

Homograph Suppression

Homograph suppression and NDComp performance. Figure 3 illustrates that low performers experienced a reliable amount of interference in the immediate condition, *F*(1, 23) = 10.32, *MSE* = 4,102.24, *p* = .004, but not the delayed condition, *F*(1, 23) = 1.46, *MSE* = 4,251.00, *p* = .239. High performers showed no reliable interference in either the immediate or delayed conditions, *F*(1, 23) = 3.99, *MSE* = 2,439.52, *p* = .058, F(1, 23) = 3.52, *MSE* = 1,998.28, *p* = .073, respectively.

A 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F*(1, 46) = 1.19, *MSE* = 2,644.20, *p* = .281, and no sentence-final word x test point interaction, *F*(1, 46) = 1.90, *MSE* = 2,644.20, *p* = .175. Performance level did not interact with sentence-final word or test point, *Fs* < 1. Overall, there were main effects of sentence-final word,

F(1, 46) = 14.54, *MSE* = 3,751.32, *p* < .001, and test point *F*(1, 46) = 8.31, *MSE* = 5,617.78, *p* = .006.

Homograph suppression and ADRP performance. Figure 4 illustrates that high

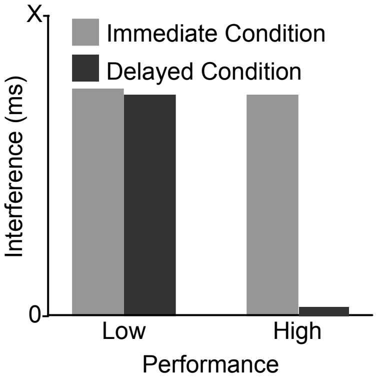

Figure 2. Expected interference of irrelevant word meanings in the immediate and delayed conditions for low and high performers on any comprehension measure. High performers were not expected to show interference in the delayed condition.

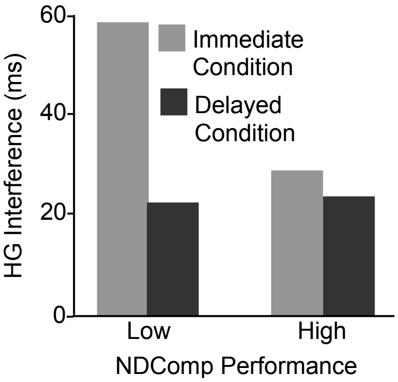

Figure 3. Interference, in the immediate and delayed conditions, of irrelevant homograph meanings (HG) exhibited by 24 low and 24 high performers on the Nelson-Denny reading comprehension subtest (NDComp).

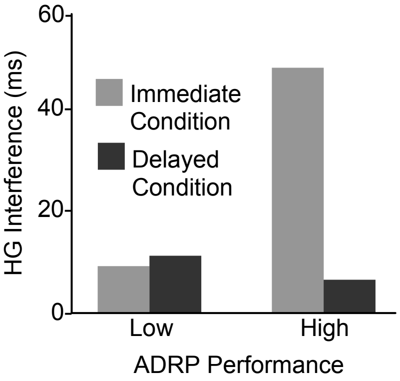

Figure 4. Interference, in the immediate and delayed conditions, of irrelevant homograph meanings (HG) exhibited by 24 low and 24 high performers on the Advanced Degrees of Reading Power (ADRP).

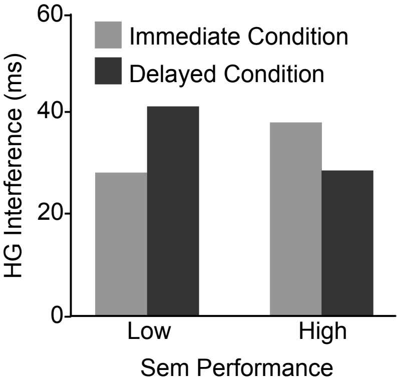

Figure 5. Interference, in the immediate and delayed conditions, of irrelevant homograph meanings (HG) exhibited by 24 low and 24 high performers on the semantic-decision task (Sem).

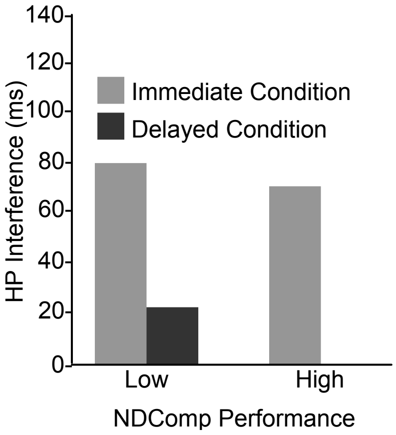

Figure 6. Interference, in the immediate and delayed conditions, of irrelevant homophone meanings (HP) exhibited by 24 low and 24 high performers on the Nelson-Denny reading comprehension subtest (NDComp).

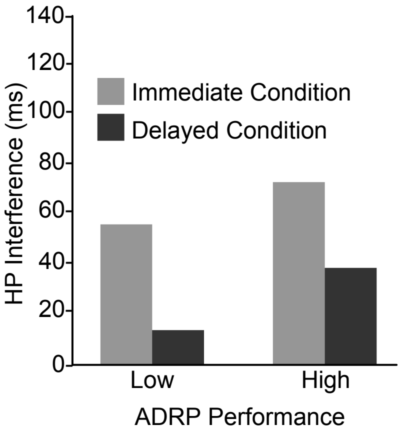

Figure 7. Interference, in the immediate and delayed conditions, of irrelevant homophone meanings (HP) exhibited by 24 low and 24 high performers on the Advanced Degrees of Reading Power (ADRP).

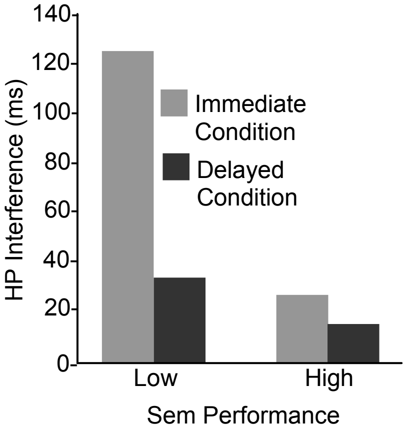

Figure 8. Interference, in the immediate and delayed conditions, of irrelevant homophone meanings (HP) exhibited by 24 low and 24 high performers on the semantic-decision task (Sem).

performers experienced a reliable amount of interference in the immediate condition point, *F*(1, 23) = 10.48, *MSE* = 2,677.53, p = .004, but not the delayed condition, *F* < 1. Low performers showed no reliable interference in either condition, *F* < 1.

A 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F*(1, 46) = 2.28, *MSE* = 2,525.79, *p* = .138, and no sentence-final word x test point interaction, *F*(1, 46) = 1.95, *MSE* = 2,525.79, *p* = .169. Performance level did not interact with sentence-final word, *F*(1, 46) = 2.10, *MSE* = 1583.97, *p* = .154, or test point, *F* < 1. Overall, there were main effects of sentence-final word, *F*(1, 46) = 10.87, *MSE* = 1,583.97, *p* < .002, and test point, *F*(1, 46) = 21.84, *MSE* = 4,113.95, *p* < .001.

Homograph suppression and sem performance. Figure 5 illustrates that high performers experienced a reliable amount of interference in the immediate condition, *F*(1, 23) = 9.57, *MSE* = 1,879.15, *p* = .005, but not the delayed condition, *F*(1, 23) = 4.00, *MSE* = 2,438.63, *p* = .058. In contrast, low performers did not show a reliable amount of interference in the immediate condition, *F*(1, 23) = 2.00, *MSE* = 4,645.59, *p* = .170, but reliable interference was evident in the delayed condition, *F*(1, 23) = 5.81, *MSE* = 3,593.76, *p* = .024. For each performance level, interference did not reliably differ from one test point to another, *Fs* < 1.

A 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F* < 1, or an interaction between sentence-final word and test point, *F* < 1. Performance level did not interact with sentence-final word or test point, *Fs* < 1. Overall, there were main effects of sentence-final word, *F*(1, 46) = 14.15, *MSE* = 3,967.53, p < .001, and test point, *F*(1, 46) = 15.32, *MSE* = 5,429.59, p < .001.

Homophone Suppression

Homophone suppression and NDComp performance. Figure 6 illustrates that high performers experienced a reliable amount of interference in the immediate condition, $F(1, 23) = 9.09$, $MSE = 6{,}797.74$, $p = .006$, but not the delayed condition, $F < 1$. Low performers also experienced a reliable amount of interference in the immediate condition, $F(1, 23) = 9.56$, $MSE = 8{,}118.05$, $p = .005$, but not the delayed condition, $F < 1$. Although not a significant difference, $F < 1$, Table 4 means and Figure 6 illustrate that the low performers experienced more interference in the delayed condition than did the high performers.

Of all of the suppression-task results, these (see Figure 6) most closely resembled the expected results in Figure 2. Nonetheless, even these similar results are different from those of the original homophone-suppression task, designed by Gernsbacher and Faust (1991). In the delayed condition, Gernsbacher and Faust’s poor performers experienced a noticeable amount of interference (effect size f = .32) whereas the poor NDComp performers in this study experienced a negligible amount of interference, with an effect size f of = .15. A large effect size f would have been .40. Furthermore, the power of the results presented here was .18, while Gernsbacher and Faust’s (1991) results had a power of approximately .71. Given the effect size of .15, the present experiment would have required 214 poor NDComp performers to obtain results as powerful as Gernsbacher and Faust’s (1991).

In addition, a 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F* < 1, but it did indicate a sentence-final word x test point interaction, *F*(1, 46) = 6.89, *MSE* = 7,293.30, *p* = .012. Performance level did not interact with sentence-final word or test point, *F* < 1. Overall, there were main effects for sentence-final word, *F*(1, 46) = 12.17, *MSE* = 7,542.10, *p* = .001, and test point, *F*(1, 46) = 15.253, *MSE* = 9,162.40, *p* < .001.

Homophone suppression and ADRP performance. Figure 7 illustrates that the high performers experienced a reliable amount of interference in the immediate condition, *F*(1, 23) = 5.88, *MSE* = 10,987.66, *p* = .024, but not the delayed condition, *F*(1, 23) = 3.72, *MSE* = 4,762.66, *p* = .066. Similarly, the low performers experienced a reliable amount of interference in the immediate condition, *F*(1, 23) = 7.22, *MSE* = 5,055.61, *p* = .013, but not the delayed condition, *F* < 1. Though Table 5 and Figure 7 illustrate that the high performers experienced more interference than the low performers in both conditions, the differences were not statistically significant, *Fs* < 1.

A 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F* < 1, or sentence-final word x test point interaction, *F*(1, 46) = 2.61, *MSE* = 6,752.05, *p* = .113. Performance level did not interact with sentence-final word or test point, *F* < 1. Overall, there were main effects for sentence-final word, *F*(1, 46) = 12.70, *MSE* = 7,689.43, *p* = .001, and test point, *F*(1, 46) = 27.61, *MSE* = 6,261.54, *p* < .001.

Homophone suppression and sem performance. Figure 8 illustrates that the high performers experienced no reliable interference in the immediate condition, *F*(1, 23) = 2.68, *MSE* = 3,524.23, *p* = .115, or the delayed condition, *F*(1, 23) = 1.36, *MSE* = 2,705.80, *p* = .256. The low performers experienced a reliable amount of interference in the immediate condition, *F*(1, 23) = 18.47, *MSE* = 10,344.79, $p < .001$, but not the delayed condition, *F*(1, 23) = 1.57, *MSE* = 8,589.58, $p = .223$.

A 2 x 2 x 2 factorial ANOVA indicated no three-way interaction, *F*(1, 46) = 2.92, *MSE* = 7,109.81, $p = .094$, or sentence-final word x test point interaction, $F < 1$. Although performance level interacted with sentence-final word, *F*(1, 46) = 7.46, *MSE* = 5,337.39, $p = .009$, it did not interact with test point, *F*(1, 46) = 3.13, *MSE* = 8,577.75, $p = .083$. Overall, there were main effects for sentence-final word, *F*(1, 46) = 23.43, *MSE* = 5,337.39, $p < .001$, and test point, *F*(1, 46) = 21.57, *MSE* = 8,577.75, $p < .001$.

Representing irrelevant-word-meaning suppression in future analyses. The homograph- and homophone-suppression tasks each produced two interference scores, one for the immediate condition and one for the delayed condition. The expected results (see Figure 2) included a reliable amount of interference in the immediate condition for the high and low performers. In addition, the expected results suggested that the low performers would experience more interference in the delayed condition than the high performers. Results similar to those expected were only found in the homophone-suppression analysis comparing high and low performers on the NDComp (see Figure 6). Because the expected results only indicated a difference between high and low performers in the delayed condition, performance in the delayed condition of the homophone-suppression task, heretofore called HPDelay, was chosen to represent one’s ability to suppress irrelevant word meanings.

Detecting Stroop Interference

A Wilcoxon Signed Rank test compared the subjects' incongruent-stimulus reaction times to their neutral-stimulus reaction times. Subjects took significantly less time to respond to the neutral stimuli, $Z = -7.38$, $p < .001$. Additionally, transformed Stroop interference scores (see next section for discussion of the transformation) were significantly different from zero, $t(71) = -43.38$, $p < .001$.

Data Characteristics

Table 7 shows means, standard deviations, and skewnesses and kurtoses with standard errors for all three measures of reading comprehension, Raven's APM, the which-came-first task, the lexical-decision task, the pseudoword-decoding task, and the motor-speed task for each hand. Also included in Table 7 are Stroop interference, HPDelay, absolute and total scores on the RSpan and OSpan tasks, accuracy scores on each of the three *n*-back subtasks, composite *n*-back accuracy scores, median reaction times on each of the three *n*-back subtasks, and composite median *n*-back reaction times.

Transformations. According to Table 7, some of the measures' distributions were skewed. To minimize the influence of the distributions' shapes on the analysis, distributions with skewnesses greater than the standard error of the skew (0.283 for 72 subjects, 0.285 for the 71 subjects who completed Raven's APM) were transformed. Such skewed distributions were transformed in two ways, and the transformation that produced the least skewness for a distribution was used for the remainder of the analysis. For example, distributions with negative skews were transformed by squaring each score. They were also transformed by adding a constant of 5 to each squared score before

Table 7

Untransformed Descriptive Statistics for Measures to be Correlated

Measure[a]	*M* (*SD*)	Skewness[b]	Kurtosis
NDComp	22.986 (5.212)	0.441	0.075
ADRP	18.125 (3.460)	-0.404	0.057
Semantic decision (s)	1.775 (0.361)	1.064	0.930
NDVocab	69.972 (6.840)	-1.193	1.790
Raven's APM[c]	37.662 (5.767)	-0.649[d]	0.109
Which came first	25.556 (6.614)	-0.041	-0.483
Motor speed: Left hand (s)	0.216 (0.351)	1.816	3.989
Motor speed: Right hand (s)	0.211 (0.287)	0.556	0.642
Lexical decision (s)	0.555 (0.750)	1.419	4.027
Pseudoword decoding	22.917 (2.136)	-2.410	9.063
Stroop interference (s)	0.131 (0.501)	0.652	0.642
HPDelay (s)	0.203 (0.108)	-0.011	2.099

Note. Continued on next page. NDComp = Nelson-Denny reading comprehension subtest; ADRP = Advanced Degrees of Reading Power; NDVocab = Nelson-Denny Vocabulary subtest; APM = Advanced Progressive Matrices; HPDelay = Homophone-suppression interference, delayed condition.

[a] n = 72, unless noted. [b] standard error = 0.283, unless noted. [c] n = 71. [d] standard error = 0.285.

Table 7, cont.

Untransformed Descriptive Statistics for Measures to be Correlated

Measure[a]	*M* (*SD*)	Skewness[b]	Kurtosis
Reading span (Absolute score)	6.500 (4.913)	1.162	1.529
Reading span (Total score)	18.849 (5.586)	0.294	-0.108
Operation span (Absolute score)	12.569 (4.122)	0.109	-0.118
Operation span (Total score)	22.597 (4.770)	-0.169	0.360
2-back accuracy	22.889 (4.889)	-1.763	6.528
2-back reaction time (s)	9.084 (0.241)	0.669	0.142
3-back accuracy	18.292 (4.957)	-0.451	0.123
3-back reaction time (s)	1.020 (0.233)	0.611	-0.127
4-back accuracy	14.194 (6.439)	-0.014	-0.551
4-back reaction time (s)	1.002 (0.267)	0.298	-0.103
N-back accuracy	2.846 (0.385)	-0.211	-0.658
N-back reaction time (s)	0.964 (0.206)	0.567	-0.341

obtaining the reciprocal. Distributions with positive skews were transformed by obtaining the square root of each score. They were also transformed by adding a constant of 5 to each square-root score before obtaining the reciprocal. The chosen transformation and resulting skewness and kurtosis for each task appear in Table 8.[6]

To make the data easier to interpret, transformed scores on three measures (NDComp, 4-back reaction time, and Stroop interference) were multiplied by -1, so better performance was associated with higher scores.

Reliabilities. Internal consistency on most measures was computed with the Spearman-Brown split-half formula. In split-half reliability analyses, one half of the items are compared to the other half. Typically, odd-numbered items are relegated to one half, while even-numbered items are relegated to the other half, so each half contains a similar amount of early-appearing and late-appearing items. With multiple item types, such as the true and false sentences in the sem, the items can first be sorted by type and then by order of appearance. From that list, odd-numbered items are relegated to one half, and even-numbered items are relegated to another. Because most of the measures in this study contained multiple item types, each half was determined as aforementioned. Afterward, the number of correct items in each half, or for some measures, the median reaction time for each half, was entered into each Spearman-Brown formula. Internal consistency reliability coefficients appear in Table 9.

---

[6] Transformed interference scores for HPDelay were obtained by subtracting each subject's transformed mean median reaction time to test words that followed experimental sentences from each subject's mean median reaction time to test words that followed control sentences. The skew of the resulting distribution was not greater than the standard error of the skew.

Table 8

Score Transformations and Resulting Descriptive Statistics for Measures to be Correlated

Measure[a]	Transformation	*M* (*SD*)	Skewness	Kurtosis
NDComp	$-1/\sqrt{5+\text{score}}$	-0.103 (0.006)	-0.141	-0.236
ADRP	score^2	340.319 (121.768)	0.103	-0.513
Semantic decision	$1/\sqrt{5+\text{score}}$	0.021 (0.002)	-0.359	-0.005
NDVocab	score^2	4,803.301 (891.968)	-0.857	0.637
Raven's APM[b]	score^2	1,451.211 (415.541)	-0.278	-0.591
Which came first	not applicable	25.556 (6.614)	-0.041	-0.483
Motor speed: Left hand	$1/\sqrt{5+\text{score}}$	0.051 (0.003)	-1.118	1.923
Motor speed: Right hand	$1/\sqrt{5+\text{score}}$	0.051 (0.003)	0.003	0.383
Lexical decision	$1/\sqrt{5+\text{score}}$	0.035 (0.002)	-0.639	1.088
Pseudoword decoding	score^2	529.674 (87.782)	-1.712	4.601
Stroop interference	$-\sqrt{\text{score}}$	-11.254 (2.201)	-0.029	0.240
HPDelay	$1/\sqrt{5+\text{score}}$	0.000 (0.001)	-0.027	0.289

Note. Continued on next page. NDComp = Nelson-Denny reading comprehension subtest; ADRP = Advanced Degrees of Reading Power; NDVocab = Nelson-Denny Vocabulary subtest; APM = Advanced Progressive Matrices; HPDelay = Homophone-suppression interference, delayed condition.

[a] n = 72, unless noted. [b] n = 71.

Table 8, cont.

Score Transformations and Resulting Descriptive Statistics for Measures to be Correlated

Measure$_a$	Transformation	*M* (*SD*)	Skewness	Kurtosis
Reading span (Absolute score)	$\sqrt{score}$	2.309 (1.088)	-0.366	-0.379
Reading span (Total score)	$\sqrt{score}$	4.293 (0.654)	-0.139	-0.025
Operation span (Absolute score)	not applicable	12.569 (4.172)	0.109	-0.119
Operation span (Total score)	not applicable	22.597 (4.770)	-0.169	0.360
2-back accuracy	$score^2$	547.472 (190.426)	-0.357	-0.405
2-back reaction time	$1/\sqrt{5+score}$	0.029 (0.003)	0.119	-0.523
3-back accuracy	$score^2$	358.819 (173.758)	0.351	0.385
3-back reaction time	$1/\sqrt{5+score}$	0.027 (0.003)	0.091	-0.212
4-back accuracy	not applicable	14.194 (6.439)	-0.014	-0.551
4-back reaction time	$-\sqrt{score}$	-31.370 (4.264)	-0.114	0.205
N-back accuracy	not applicable	2.846 (0.385)	-0.211	-0.658
N-back reaction time	$1/\sqrt{5+score}$	0.028 (0.003)	-0.023	-0.683

Table 9

Measure[a]	Internal Consistency Measure	Coefficient
ADRP	Spearman-Brown split half	.71
Semantic decision	Spearman-Brown split half	.97
Raven’s APM	Spearman-Brown split half	.83
Which came first	Spearman-Brown split half	.81
Motor speed: Left hand	Spearman-Brown split half	.96
Motor speed: Right hand	Spearman-Brown split half	.95
Lexical decision	Spearman-Brown split half	.93
Pseudoword decoding	coefficient alpha	.71
Stroop interference	Spearman-Brown split half	.97
Homograph suppression: Form 1[b]	Spearman-Brown split half	.99
Homograph suppression: Form 2[c]	Spearman-Brown split half	.99
Homograph suppression: Form 3[d]	Spearman-Brown split half	.97
Homograph suppression: Form 4[c]	Spearman-Brown split half	.97
Homophone suppression: Form 1[c]	Spearman-Brown split half	.98
Homophone suppression: Form 2[b]	Spearman-Brown split half	.96
Homophone suppression: Form 3[d]	Spearman-Brown split half	.91
Homophone suppression: Form 4[c]	Spearman-Brown split half	.97

Note. Continued on next page. ADRP = Advanced Degrees of Reading Power. APM = Advanced Progressive Matrices.
[a] n = 72, unless stated. [b] n = 19. [c] n = 18. [d] n = 17.

Table 9, cont.

Internal Consistencies of Measures and Their Forms

Measure$_a$	Internal Consistency Measure	Coefficient
Reading span	coefficient alpha	.78
Operation span	coefficient alpha	.74
2-back accuracy	coefficient alpha	.47
2-back reaction time	Spearman-Brown split half	.91
3-back accuracy	coefficient alpha	.29
3-back reaction time	Spearman-Brown split half	.85
4-back accuracy	coefficient alpha	.51
4-back reaction time	Spearman-Brown split half	.89

Another reliability measure, the coefficient alpha, also called Cronbach's alpha, was used to analyze the consistency of the span, *n*-back accuracy, and pseudoword-decoding scores because they contained polytomous items, i.e., items that are not scored as simply right or wrong.

Internal consistency was not computed for the Nelson-Denny subtests because they were speeded, meaning that the number correct was determined by the number of items completed. Most believe that speeded measures produce inflated internal-consistency coefficients, though one report argues that some aspects of a speeded test can lead to a deflated coefficient (Educational Testing Service, 2004). In contrast, though the sem task was also speeded, each subject saw each item, meaning that I could measure the internal consistency using the reaction times for each item.

According to Koch (2003), reliability coefficients between .70-.80 are typical for scores on experimental tasks, so only the accuracy scores on the *n*-back subtests were distinctly unreliable, with coefficients ranging from .29-.51 (Table 9). Koch also noted that standardized-test-score reliability coefficients are typically above .90, but coefficients between .60 and .70 can occur on classroom tests and might indicate homogeneity of subjects.

Span reliability was determined by first separating the sentence-clusters into three groups, as was done by Engle, Cantor, and Carullo (1992) and Engle et al. (1999). The first group contained the first two-sentence cluster that appeared, the second three-sentence cluster that appeared, the third four-sentence cluster, and the first five-sentence cluster. The second group contained the second two-sentence cluster, the third three-sentence cluster, the first four-sentence cluster, and the second five-sentence cluster. The third group contained the third two-sentence cluster, the first three-sentence cluster, the second four-sentence cluster, and the third five-sentence cluster. Those scores were used to compute the coefficient alpha.

The consistency of the *n*-back accuracy scores was determined in a similar fashion. Items on each subtest were divided into three groups. Fifty percent of the items in each group had appeared *n* items prior during the task, and fifty percent of the items had not. Accuracy scores were computed for each group as they were previously computed for the full subtask, and those scores were used to compute the coefficient alpha. A coefficient alpha was additionally obtained for the pseudoword-decoding items.

Some measures, e.g., the NDComp, NDVocab, and homograph- and homophone-suppression tasks, involved multiple forms. Both forms of the NDComp and NDVocab were compared for equality of variances, $F < 1$, and score differences. NDComp scores on form G ($M = 22.08$, $SD = 5.11$) were slightly lower than on form H ($M = 23.94$, $SD = 5.22$), but an independent-samples t-test showed no reliable difference, $t(70) = 1.53$, $p = .131$. Similar results were obtained for the NDVocab scores. Scores on form G ($M = 68.95$, $SD = 6.87$) were slightly lower than scores on form H ($M = 46.89$, $SD = 6.85$), but they were not reliably different, $t(70) = .79$, $p = .432$.

The homograph and homophone-suppression tasks each consisted of four forms. Of the four homograph-suppression forms, form 1's scores were the most variable and the highest on average ($M = 773.16$, $S^2 = 27{,}596.06$), and form 3's scores were the least variable and the lowest on average ($M = 702.68$, $S^2 = 12{,}246.22$). An F-test for variances, $F(18, 16) = 2.25$, $p = .054$, and an independent-samples t-test, $t(34) = 1.48$, $p = .148$, showed no reliable difference between either the means or the variances of these two forms. Similarly, of the four homophone-suppression forms, form 1's scores again were the most variable and the highest on average ($M = 892.03$, $S^2 = 42{,}768.46$), and form 3's were again the least variable and the lowest on average ($M = 812.65$, $S^2 = 19{,}879.18$). Again, an F-test for variances, $F(17, 16) = 2.15$, $p = .066$, and an independent-samples t-test, $t(34) = 1.30$, $p = .204$, showed no reliable differences between the forms.

Correlations Between Tasks

Table 10 shows correlations between scores on each measure. Of note, the speeded comprehension measures correlated moderately with each other, $r_{\text{NDComp, sem}} = .42$, $p < .001$, but neither reliably correlated with the ADRP, an untimed test. Scores on the NDVocab correlated with all of the comprehension measures ($r_{\text{NDComp, NDVocab}} = .48$, $p < .001$; $r_{\text{ADRP, NDVocab}} = .50$, $p < .001$; $r_{\text{sem, NDVocab}} = .46$, $p < .001$). Only scores on the ADRP reliably correlated with scores on Raven's APM**,** $r = .49$, $p < .001$; neither test was timed.

The span and *n*-back tasks, which were assumed to measure central executive processing, were mostly uncorrelated, and according to Table 10, neither group of tasks was overall a better predictor of comprehension-measure performance. The highest correlation between a span measure and an *n*-back measure—between total operation span and 3-back subtask accuracy—was .25, $p = .034$. Absolute operation span did not reliably correlate with 3-back accuracy, $r = .11$, $p = .363$, suggesting that scoring method

Table 10

Pearson Correlations for Comprehension Measures and Measures Considered for the Principal Components Factor Analysis

	Measure$_a$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	NDComp	–																							
2.	ADRP	17	–																						
3.	Sem	.42**	-.08	–																					
4.	NDVocab	.47**	.49**	.46**	–																				
5.	Raven's APM$_b$	12	.46**	-.08	16	–																			
6.	WCF	.20	14	.04	.25*	15	–																		
7	MSLeft	12	-.09	.28*	.09	-.04	.02	–																	
8.	MSRight	.21	.01	.34**	.20	.03	.15	79**	–																
9.	LD	.25*	15	.62**	.32**	-.05	-.08	.35**	.41**	–															
10.	Pseudo	.25*	.08	.43**	.37**	14	-.01	11	15	10	–														
11.	Stroop	-.18	.02	.21	.05	.07	12	.29*	.21	10	.19	–													
12.	HPDelay	-.20	-.08	.03	.06	.07	.09	.06	.05	-.12	-.12	-.18	–												
13.	RSpanAbs	.24*	-.04	.26*	.09	.04	-.03	11	15	14	.28*	.06	-.02	–											
14.	RSpanTot	17	.09	19	.13	19	.01	17	.20	13	.25*	.18	-.15	78**	–										
15.	OSpanAbs	.26*	.09	.15	.05	.15	-.13	17	19	-.03	.24*	.00	.06	.48**	.40**	–									
16.	OSpanTot	.29*	14	.11	.02	.35**	-.06	.06	.07	-.04	.22	.01	.00	.34**	.43**	.63**	–								
17	2BackAcc	.27*	19	-.02	12	.32**	-.04	-.10	-.07	.07	.06	.08	-.10	.08	19	13	18	–							
18.	2BackRT	12	-.09	.40**	.09	-.01	-.17	18	.37**	.35**	.27*	10	-.13	.06	.03	-.01	-.03	.25*	–						
19.	3BackAcc	.05	11	-.17	-.12	.21	.08	-.11	-.16	-.08	-.05	-.12	-.23	-.02	16	.11	.25*	.06	-.08	–					
20.	3BackRT	11	-.02	.34**	12	-.23	-.02	.25*	.40**	.22	.24*	-.01	-.05	.05	15	-.08	-.05	.06	.60**	-.04	–				
21	4BackAcc	.03	.25*	-.12	-.02	17	.10	-.16	-.13	-.03	-.07	.03	-.19	-.01	10	.02	12	.19	-.02	.47**	.06	–			
22.	4BackRT	15	-.06	.20	.12	-.11	-.10	.16	15	.09	.18	-.10	.05	-.10	-.04	-.06	-.05	.15	.37**	-.07	.67**	-.20	–		
23.	NBackAcc	14	.28*	-.13	-.01	.32**	.05	-.20	-.20	-.01	-.04	.02	-.24*	.00	19	.10	.24*	.54**	.02	.02	.02	.83**	-.08	–	
24.	NBackRT	13	-.06	.36**	.12	-.12	-.12	.24*	.36**	.28*	.26*	.01	-.09	-.05	.03	-.08	-.08	.18	.78**	-.03	.88**	-.03	.80**	.03	–

Note. NDComp = Nelson-Denny reading comprehension subtest; ADRP = Advanced Degrees of Reading Power; Sem = Semantic decision; NDVocab = Nelson-Denny vocabulary subtest; APM = Advanced Progressive Matrices; WCF = Which came first; MSLeft = Motor speed: Left hand; MSRight = Motor speed: Right hand; LD = Lexical decision; Pseudo = Pseudoword decoding; Stroop = Stroop interference; HPDelay = homophone-suppression interference, delayed condition; RSpanAbs = Reading span (Absolute score); RSpanTot = Reading span (Total score); OSpanAbs = Operation span (Absolute score); OSpanTot = Operation span (Total score); 2BackAcc = 2-back accuracy; 2BackRT = 2-back reaction time; 3BackAcc = 3-back accuracy; 3BackRT = 3-back reaction time; 4BackAcc = 4-back accuracy; 4BackRT = 4-back reaction time; NBackAcc = *N*-back Accuracy; NBackRT = *N*-back reaction time.

[a] n = 72, unless noted. [b] n = 71.

influenced the magnitude of the correlations. Correlations between absolute and total reading span, $r = .78$, $p < .001$, and absolute and total operation span, $r = .63$, $p < .001$, indicate that the two scoring methods produced similar, but not entirely consistent results. Within the *n*-back subtasks, speed scores were essentially uncorrelated with accuracy scores (Table 10), indicating that more accurate performers did not necessarily respond faster.

Stroop interference and HPDelay, the two measures that were assumed to measure the suppression of irrelevant information, were unreliably correlated, $r = -.18$, $p = .124$. Nonetheless, they loaded together in the factor analysis described in the next section.

Analyzing Reading Comprehension

Reading comprehension was decomposed and analyzed in three steps. First, a factor analysis was performed on the non-comprehension measures via a principal components extraction, and one subject was automatically removed from the analysis for failure to take Raven's APM. Then, another principal components factor analysis was performed on the three comprehension measures to obtain a composite comprehension factor. For reasons to be explained later, two of the measures, NDComp and sem, but not ADRP, loaded onto a single factor to form a composite reading comprehension ability score. Afterward, the non-comprehension factors were entered as independent variables into two stepwise multiple linear regressions. The dependent variable in one regression was the composite NDComp-sem factor (heretofore called *NDCompSem*), and the dependent variable in the other was ADRP.

Factor Analysis for Non-Comprehension Measures

Variables. The following 15 variables were entered into the factor analysis: NDVocab, Raven's APM score, which-came-first score, right-hand motor speed, left-hand motor speed, lexical-decision time, pseudoword decoding accuracy, Stroop interference, HPDelay, 3-back accuracy, 4-back accuracy, 3-back reaction time, 4-back reaction time, RSpan (absolute score), and OSpan (absolute score).

Some variables were left out of the analysis, including accuracy on the 2-back subtask. While accuracy on the 3- and 4-back subtasks was reliably correlated, $r = .47$, $p < .001$, 2-back accuracy was minimally and unreliably correlated with 3-back or 4-back accuracy (Table 10). Furthermore, 2-back accuracy, unlike 3- and 4-back accuracy, reliably correlated with scores on Raven's APM, $r = .32$, $p = .007$. Those data implied that some subjects had difficulty understanding the task and needed more practice than others, meaning that 2-back accuracy probably measured learning speed. Because 2-back accuracy was thus inappropriate for the analysis, examining reaction times on the 2-back subtask was also deemed inappropriate.

To prevent redundancy and enhance the power of the analysis, *n*-back accuracy and median *n*-back reaction time were not entered. For similar reasons, total RSpan and OSpan scores were not entered.

Absolute spans, instead of total spans, were entered for three reasons. First, performance on similar versions of the span tasks that appear in the present research are typically assessed with the absolute span (Engle, 2003c). Second, total OSpan was reliably correlated with performance on Raven's APM, $r = .345$, $p = .003$, but all other spans were not. In a preliminary factor analysis using the total, but not the absolute, spans, Raven's APM performance loaded onto the same factor as the spans. Nonetheless, Table 11 indicates that Raven's APM performance did not load with RSpan or OSpan when the absolute spans, but not the total spans, were entered. Thus, using the absolute spans more clearly separated the Given (measured via NDVocab, Raven's APM, and which-came-first performance) from the variables being examined, enabling a simpler analysis. Finally, the absolute spans reliably correlated with NDComp performance (see Table 10). Only total OSpan, and not total RSpan, correlated with performance on the NDComp, and there was no reliable relationship between any of the total spans and ADRP or sem performance. Because an empirical purpose of this study was to determine the importance of the relationship between span and performance on comprehension measures, an established scoring method that would allow a relationship to exist is appropriate.

Computation. One subject was automatically eliminated from the analysis for failure to complete Raven's APM. As many of the variables related to each other, the factor analysis was first performed with a direct oblimin rotation. Six factors were extracted according to the eigenvalue-greater-than-one rule, but the factors were, at best, minimally and unreliably correlated. Hence, the factor analysis was performed again with a varimax rotation. Fortunately, the absence of multicollinearity among the factors obtained with the varimax rotation made the subsequent multiple regressions easier to compute and interpret—multicollinearity among independent variables can make

Table 11

Principal Components Exploratory Factor Analysis for Non-Comprehension Measures

	Factor					
Measure	1	2	3	4	5	6
NDVocab	.07	.22	.10	-.17	**.76**	.14
Raven’s APM	-.02	-.24	.19	.32	**.53**	-.17
Which came first	.16	-.17	-.30	.16	**.64**	-.15
Motor speed: Left hand	**.90**	.07	.11	-.09	-.03	-.02
Motor speed: Right hand	**.90**	.16	.12	-.06	.12	-.05
Lexical decision	**.50**	.19	.03	-.07	-.09	.32
Pseudoword decoding	-.05	**.32**	**.50**	-.15	**.44**	**.39**
Stroop interference	.37	-.29	-.03	-.11	.17	**.56**
HPDelay	.06	-.03	-.01	-.29	.18	**-.78**
Reading span (Absolute score)	.11	-.05	**.77**	-.03	.03	.09
Operation span (Absolute score)	.13	-.09	**.84**	.11	-.04	-.17
3-back accuracy	-.11	.02	.06	**.85**	-.02	.00
3-back reaction time	.30	**.85**	-.04	.07	-.02	.07
4-back accuracy	-.06	-.04	-.05	**.80**	.07	.14
4-back reaction time	.07	**.86**	-.05	-.10	.01	-.11

Note. Varimax rotation. Boldface loadings indicate factor assignments. Factors 1-6 accounted for 14.53%, 12.45%, 11.47%, 11.15%, 10.28%, and 8.82% of the total variance, respectively. 1 = Motor Speed. 2 = Correct Decision Speed. 3 = Working Memory and Processing Ability. 4 = Central Executive Functioning. 5 = Knowledge and Intelligence. 6 = Suppression Ability. NDVocab = Nelson-Denny vocabulary subtest. APM = Advanced Progressive Matrices. HPDelay = Homophone-suppression interference, delayed condition.

multiple regressions difficult or even impossible to calculate, and it can also result in independent-variable coefficients with high standard errors.

From the chosen factor analysis, six factors, accounting for a total of 68.71% of the total variance, were extracted according to the eigenvalue-greater-than-one rule (for percent of variance attributed to each factor, see notes to Table 11). All but one of the variables, pseudoword-decoding accuracy, loaded onto one factor substantially more than the others. Pseudoword-decoding accuracy loaded similarly onto Factors 2, 3, 5, and 6. Possible constructs measured by the task will be discussed in the upcoming paragraphs.

Indicated in Table 11, Factor 1 clearly represents Motor Speed. Right- and left-hand motor speed loaded heavily onto the factor, and the lexical-decision task, a rather intellectually simple speeded task, also loaded.

Loading mainly onto Factor 2 were 3- and 4-back reaction time, but because the reaction times were for correct responses only, the factor was named Correct-Decision Speed. Pseudoword decoding accuracy, which loaded .32 onto that factor, also estimated Correct-Decision Speed. Subjects who slowly sounded out a pseudoword lost .5 point, and subjects who suggested one or more incorrect pronunciations before a correct one lost .5 point per pronunciation, for up to two incorrect pronunciations.

Loading mainly onto Factor 3 were RSpan and OSpan performance, and to a lesser degree, pseudoword-decoding accuracy. Two issues might have contributed to the loading of pseudoword-decoding accuracy. First, the pseudoword-decoding task measured a known contributor to span-task performance, processing speed. Subjects who could quickly process the pseudowords would not have lost points for slowly sounding them out. Second, the span tasks and the pseudoword-decoding task both required the subjects to pronounce unusual words or phrases aloud to an experimenter. Performance anxiety might have distracted some subjects from the stimuli, causing processing difficulties that made it difficult to accurately decode aloud. That, in turn, might have led to lower scores. In the pseudoword-decoding task, decoding errors resulted in a loss of points. In the span tasks, decoding errors often caused the subjects to pause or correct themselves. Recognizing the errors tended to further distract the subjects and slow their performance, increasing the likelihood of the target words' decay and impeding the full recall of the noun groups. In essence, generally poorer processors and distracted subjects might have been at a disadvantage in the span and pseudoword-decoding tasks.

Factor 3 was named Working Memory and Processing Ability. The name derived from the fact that the span tasks traditionally measured a construct called *working memory capacity*, and performance on all three tasks likely depended on how well the stimuli were processed.

Loading mainly onto Factor 4 were accuracy on the 3- and 4-back tasks. Because *n*-back subtasks were widely considered tests of the central executive, Factor 4 was named Central Executive Functioning.

Loading mainly onto Factor 5 were NDVocab score, Raven's APM score, which-came-first score, and to a lesser extent, pseudoword-decoding accuracy. The NDVocab and which-came-first tasks measured knowledge of words and history, and a high score on the pseudoword-decoding task indicated knowledge of the English orthographic cipher. Raven's APM has traditionally measured nonverbal intelligence. If not for the presence of pseudoword-decoding accuracy, this factor would have been called Given, but because the Given does not include decoding ability, Factor 5 was called Knowledge and Intelligence.

Loading mainly onto Factor 6 were HPDelay, Stroop interference, and to a lesser degree, pseudoword-decoding accuracy. Both the homophone-suppression and Stroop tasks have been said to involve the suppression of irrelevant information, though unexpectedly, HPDelay loaded negatively, while Stroop interference and pseudoword-decoding accuracy loaded positively.

Introspection tells us that when we attempt to decode pseudowords, we occasionally think of possible meanings or associate them with existing words. These associations and speculations can distract us from the pseudowords themselves, so accurate pseudoword decoding should require us to repress those distracters. Observation of the subjects in this study indicated that sometimes, the distracters were not repressed. Upon seeing the pseudoword *scorth*, some subjects pronounced it as a similar-looking word, *scorch*. And upon seeing the pseudoword *whulse*, some subjects pronounced it "wussle"; trigram frequencies for *sle* are slightly higher than they are for *lse* (*Distribution of trigrams (three-letter combinations) in a 1 mil. word corpus of English*, n.d.). Because the tasks loading onto Factor 6 involved some form of suppression, the factor was named Suppression Ability.

Factor Analysis for Comprehension Measures

To obtain a composite reading comprehension ability score, NDComp, ADRP, and sem were entered into an exploratory principal components factor analysis, which produced one factor according to the eigenvalue-greater-than-one rule. The factor produced from the analysis accounted for 51.22% of the total variance among the three measures. NDComp, ADRP, and sem loaded a respective .79, .53, and .80 onto the factor. However, performing a Cronbach's alpha on the three measures produced a coefficient of .00. Removing ADRP scores from the Cronbach's alpha analysis raised the coefficient to .39, meaning that the ADRP scores were irrelevant to the factor. A second principal components factor analysis of just NDComp and sem produced a single factor, NDCompSem, accounting for 70.88% of the total variance. Both measures loaded .84. The resulting factor scores for NDCompSem had a skewness of 0.04, and a kurtosis of -0.14.

Regression Analyses

Two stepwise, multiple regression analyses were conducted. In the first analysis, the dependent variable was NDCompSem, and in the second, it was ADRP. The 6 non-comprehension factors acted as the independent variables.

Assumptions. Examination of scatterplots indicated homoscedasticity among all six independent variables and both dependent variables. The distributions of the dependent variables—NDCompSem and ADRP—were normal (see Table 8 for descriptive ADRP statistics). The distributions of two independent variables, Motor Speed and Knowledge and Intelligence, were reliably negatively skewed, though the most negative skew was -0.80. The distribution of Knowledge and Intelligence was also reliably leptokurtic (Table 12). The distributions of Motor Speed, Knowledge and Intelligence, and Working Memory and Processing Ability each had one

Table 12

Descriptive Statistics for Independent Variables in Regression Analyses

Independent Variable	Skewness$_a$	Kurtosis$_b$
Motor Speed	- 0.80	0.86
Correct-Decision Speed	-0.26	-0.46
Working Memory and Processing Ability	-0.32	0.81
Central Executive Functioning	0.15	-0.30
Knowledge and Intelligence	-0.74	1.19
Suppression Ability	-0.12	0.06

[a]standard error = 0.29. [b]standard error = 0.56.

outlier at least three standard deviations below the mean, but removing them had minimal effects on the regression analyses.

Computation. First, a stepwise multiple regression was performed using NDCompSem as the dependent variable (see Table 13). Four independent variables—Knowledge and Intelligence, Correct-Decision Speed, Working Memory and Processing Ability, and Motor Speed—significantly regressed onto the dependent variable. Knowledge and Intelligence accounted for 17.6% of the total NDCompSem variance. Correct-Decision Speed independently accounted for an additional 9.5% of the variance. Working Memory and Processing Ability accounted for a further 9.5% of the variance. And Motor Speed independently accounted for 7.5% of the total variance.

Second, a stepwise multiple regression was performed using ADRP as the dependent variable (see Table 14). Two independent variables, Knowledge and Intelligence and Central Executive Functioning, significantly regressed onto the dependent variable. Knowledge and Intelligence accounted for 23.5% of the ADRP variance, and Central Executive Functioning independently accounted for 4.9% of the variance.

Table 13

Regression Coefficients for NDCompSem

Independent Variable	*R*	R^2	*p of change in variance*
Knowledge and Intelligence	.42	.18	< .001
Correct-Decision Speed	.52	.27	< .001
Working Memory and Processing Ability	.61	.37	< .001
Motor Speed	.66	.44	< .001

Note. Stepwise multiple regression. Central Executive Functioning and Suppression Ability did not regress. NDCompSem = Factor derived from Nelson-Denny Reading Comprehension subtest and semantic-decision scores.

Table 14

Regression Coefficients for Advanced Degrees of Reading Power

Independent Variable	*R*	R^2	*p of change in variance*
Knowledge and Intelligence	.49	.24	< .001
Central Executive Functioning	.53	.28	.035

Note. Stepwise multiple regression. Motor Speed, Correct-Decision Speed, Working Memory and Processing Ability, and Suppression Ability did not regress.

Chapter Four: Discussion

The present research sought to determine whether two mechanisms associated with reading comprehension ability—central executive processing and the suppression of irrelevant information—were immediately related. It was hypothesized that one mechanism might cause the other (the Central-Executive-As-Mother and Central-Executive-As-Daughter hypotheses), or the two mechanisms might function independently within the realm of reading comprehension ability (the Sisters hypothesis), or they might be inconclusively related within the realm of reading comprehension ability (the Conjoined Sisters hypothesis).

The data from this study did not strongly support any of those hypotheses (see Figure 1). Specifically, the factors representing suppression and the various executive processes were statistically independent, and the suppression factor never regressed onto reading comprehension ability, violating two major assumptions of this research. Because all of the factors were statistically independent, there was no evidence that central executive processing, of any sort, is a prerequisite for suppression, in accordance with the Central-Executive-as-Mother hypothesis, nor was there evidence that suppression is necessary for central executive processing, in accordance with the Central-Executive-as-Daughter hypothesis. Also due to statistical independence, the mechanisms were not inconclusively related, in accordance with the Conjoined Sisters hypothesis. The statistical independence could have permitted results appearing to favor the Sisters hypothesis, which involves independent central executive and suppression mechanisms within the realm of reading comprehension ability. Nonetheless, the Sisters hypothesis states that while the mechanisms are not related within the realm of reading comprehension ability, they should otherwise be related. That means that even if suppression ability regressed onto reading comprehension ability, the Sisters hypothesis would not be supported.

And of course, the measures assumed to assess central executive processing were not all represented by one factor. Thus, this study, which was intended to examine the relationship between two constructs within the realm of reading comprehension ability, examined the relationship among three constructs as related to comprehension ability and each other.

In light of the results, the operational definitions of the central executive processing and suppression mechanisms are inappropriate in this context. Because the null hypothesis of unrelated mechanisms could not be rejected, the results should be examined in light of what can otherwise be inferred.

Validity of the Measures

Because the results were fairly unexpected, one might wonder whether the data were sufficiently valid to draw conclusions from them. Convergent validity was evident among several sets of scores. The comprehension scores were valid because they were reliably correlated with the vocabulary scores. Similarly, the vocabulary scores were valid because they were reliably correlated with scores on the three comprehension measures. Scores on the which-came-first task, which focused mainly on prior historical knowledge, were significantly related to NDVocab scores. Likewise, Stanovich and Cunningham (1992) found that NDVocab scores were significantly related to a test of historical and literary knowledge. Furthermore, data from the Stroop task showed a clear Stroop effect, converging with decades of prior research. Finally, the decoding scores produced from the lexical-decision and pseudoword-decoding tasks were valid because they were significantly more positively correlated with comprehension measures containing easily comprehensible text. Specifically, the correlation between NDCompSem and lexical decision was .52 ($p < .001$), and the correlation between ADRP and lexical decision was .15 ($p = .21$), and the two correlations were significantly different, $t(69) = 6.19$, $p < .001$; similarly, the correlations between NDCompSem and pseudoword decoding and ADRP and pseudoword decoding were .40 ($p < .001$) and .09 ($p = .47$), respectively, and the correlations were significantly different, $t(69) = 3.65$, $p < .001$. Those significant differences illustrate how measures of reading comprehension assess both decoding and comprehension. Measures containing easily comprehensible text account for more decoding variance than measures containing relatively difficult text. That is because most college students are capable of understanding simple concepts, so a reading comprehension measure using easily comprehensible text must assess decoding ability—if the students can decode the text, they should have no trouble comprehending it. College students performing poorly on such an assessment probably have difficulty decoding.

While the above measures were characterized by common relationships, other measures were characterized by common estrangements, indicating divergent validity. The motor-speed scores were valid because they did not correlate with the untimed, computerized measures. Although the HPDelay data bore little resemblance to

Gernsbacher and Faust's (1991) data and data from similar studies by Gernsbacher and colleagues, they might represent the abilities of the population appearing in the present research. That is, the subjects were likely of higher ability, and more homogeneous, than subjects in similar, previous studies, explaining the diverging results. Dissimilarities indicated divergent validity among other data, as well. As in Roberts and Gibson's (2002) study, span and *n*-back accuracy scores were uncorrelated. In addition, prior research has shown that span and nonverbal intelligence scores are minimally related, (Cunningham et al., 1990; Carr, Brown, & Vavrus, 1985), and the present study, which showed an unreliable relationship between span and Raven's APM scores, provided further converging evidence.

A Cast of Complex Characters

Although the results did not conform to the hypotheses, this study has produced further information about reading comprehension ability. It appears, for example, that knowledge and intelligence can predict reading comprehension ability fairly well (see Tables 13 & 14). And given the type of research performed and complexity of what influences reading comprehension ability, the factors appearing in the regression analyses accounted for a substantial proportion of variance in reading comprehension ability; the remaining variance might be explained by some of the other elements, e.g., motivation, discussed later in this chapter.

But the results also imply that the notion of reading comprehension ability is too general. In fact, we already know that comprehension is multifaceted, requiring parsing, integration, and disambiguation, plus a host of other abilities and strategies. Perhaps that explains the differences among the reading comprehension measures. The most noticeable difference was between the speeded measures—NDComp and sem—and the power measure, ADRP. Although it is quite possible that the relationship between the speeded measures and their minimal relationships with the ADRP were exclusively due to the speededness of each measure, it is also likely that the tasks measured different types of reading comprehension ability.

That hypothesis conforms to the data, which showed that while Working Memory and Processing Ability, which included performance on the span tasks, regressed onto NDCompSem, Central Executive Functioning, consisting of performance on the 3- and 4-back tasks, regressed onto ADRP performance. Although both sets of central executive tasks might have measured central executive processing, they might have done so in very different ways. In this section, I will first discuss how the tasks were different. Then, I will explain how that difference was relevant to reading comprehension abilities.

Miyake et al. (2000) wrote that *n*-back tasks most likely measure updating ability. Like the notion of reading comprehension ability, the notion of updating ability might be too general because some types of updating might require more inhibition than others. In the first chapter, I explained updating according to the processes that we could use to continually note the identity and positions of three leading horses in a close race. That updating process, as described, should have required more inhibition than the *n*-back subtasks.

I hypothesize that for each of the *n*-back subtasks, the subjects stored representations of the stimuli in an updatable queue that was *n* stimuli long. The presentation of each new stimulus caused the stimulus representations already in the queue to move one place closer to the back of the queue. Once a stimulus appeared more than *n* places prior to the stimulus being presented, its representation was eliminated from the queue. However, all of the stimulus representations were relevant until they reached the back of the *n*-stimulus queue. Furthermore, they always remained in the same order relative to each other. Because the stimulus representations always remained in the same order throughout their time in the queue, the most salient updating elements in the *n*-back task involved the addition, not the subtraction, of information. In contrast, subtracting information might have been more important than adding information in the horserace example in the first chapter. Specifically, the horses might not have remained in a constant order throughout the race, and other horses might have entered the top three places as the initially faster horses fell behind. Forgetting the horses' prior placement might have been advantageous in that situation.

Similar to the horse race example, the span tasks probably required a good deal of subtractive updating. Each relevant element in those tasks, i.e., the nouns to be memorized, was separated from the others by a fairly irrelevant sentence or equation. To prevent distraction as they updated their list of nouns to recall, the subjects had to subtract the irrelevant sentences or equations from their working memories. Failure to do so resulted in incorrect responses. For example, in the RSpan task, a number of subjects mistakenly recalled words that made the sentences nonsensical. Therefore, success on the RSpan task probably required the subjects to ignore those salient, but irrelevant, words. The importance of ignoring those words, sentences, and equations is consistent with the research of Miyake and his colleagues, who linked span not only to updating ability (Miyake et al., 2000) but to other cognitive processes, including inhibition (Friedman & Miyake, 2004; Miyake et al., 2000). Similarly, Baddeley (2002) has claimed uncertainty over exactly how working memory span fits with his theory of working memory, though Engle and colleagues (e.g., Engle, 2001; Engle et al., 1999) have suggested that the span tasks assess attention in the face of distraction.

Notions of attention and additive updating can be applied to the assessment of reading comprehension. For example, the ADRP required mostly additive updating. Specifically, it measured the global integration of information over varying amounts of text. To correctly respond to the items, the subjects had to combine information from the text into a coherent entity. Most of the text in the passages was directly relevant to the items or the passages as a whole, so the conscious subtraction of information was relatively unnecessary. In contrast, the NDComp and sem tasks required more attention to detail than did the ADRP. Many correct NDComp answer choices very closely resembled material in their respective passages. Attentive subjects were probably able to locate the answers in the passages more quickly than less attentive subjects, and because the NDComp was a speeded test, attentiveness was advantageous. Similarly, subjects who more aptly focused on the semantic decisions in the sem task probably responded more quickly overall. In short, the nature of the comprehension involved in the NDComp, sem, and ADRP was reflected by the factors associated with them.

Nonetheless, the data did not conform to some of the major assumptions of this research. Although little-to-no prior research had compared *n*-back performance to comprehension ability—meaning that a relationship between the two was not guaranteed—span was expected to be moderately correlated with reading comprehension ability in accordance with numerous other studies (for a review, see Daneman & Merikle, 1996). In this study, however, the highest correlation between an absolute span and a reading comprehension measure was .26. Such results might have occurred for several reasons.

First, the relationship between reading comprehension ability and span might not be robust enough to transcend different span-task formats. Many researchers have taken varying efforts to ensure that the subjects pay attention to the "distracting" elements of the span stimuli, such as the complete sentences in the RSpan task and the full equations in the OSpan task. During reading and listening-span tasks, for example, experimenters have asked the subjects questions about the sentences they read. Daneman and Carpenter (1980, experiment 2) presented their subjects with semantically and factually true sentences, plus sentences like, *The Supreme Court of the United States has eleven justices*, and asked them to indicate whether they were true or false. However, the use of purely nonsensical sentences, such as, *All parents hope their list will grow up to be intelligent*, which was used in the present research, was rarely used in the verbal span tasks. Into their listening span tasks, Baddeley, Logie, Nimmo-Smith, and Brereton (1985, experiments 1 and 2) incorporated short, single-clause sentences, half sensical and half purely nonsensical, like, *The girl sang the water*. Subjects indicated, via key press, whether each sentence made sense. Performance on the tasks correlated between .46 and .49 with an earlier edition of NDComp, and .33 and .10, respectively, on a measure of vocabulary knowledge. In their review, Daneman and Merikle (1996) obtained the averages of the correlations between span and NDComp and span and vocabulary knowledge for each experiment and arrived at .40 and .31, respectively. Both correlations were lower than the .53 obtained by Daneman and Carpenter (1980, experiment 2), who correlated listening span with performance on the verbal SAT, a measure including vocabulary knowledge. It appears that a working memory task involving purely nonsensical sentences might require additional strategies that are not pertinent to reading comprehension, thus producing a lower correlation between span and reading comprehension ability.

Nonsensical sentences notwithstanding, the format of the span tasks used in this study was quite similar to that used by Engle et al. (1999), whose RSpan tasks were noticeably different from traditional RSpan tasks. In traditional RSpan tasks, subjects were asked to memorize a word appearing within each sentence in each group. In this and Engle et al.’s (1999) studies, the subjects were asked to memorize an unrelated noun appearing after each sentence in each stimulus group to minimize the influence of reading comprehension ability on span. Without that influence, it was not surprising that RSpan was not very highly correlated with reading comprehension ability in Engle et al.’s (1999) research and the present research.

Even more importantly, in Engle et al.’s (1999) study, RSpan and OSpan correlated .51, and they correlated a similar .48 in the present research. That implies that Engle and colleagues’ tasks, and the tasks used here, measured similar cognitive mechanisms. Nonetheless, Engle et al. (1999) found significant correlations between span and nonverbal intelligence, though the present research did not, as well as higher correlations between span and reading comprehension ability than were found in the present research. But the reading comprehension measure was the verbal SAT test, which measures a number of skills in addition to reading comprehension ability, and the nonverbal intelligence measure was Raven's Standard Progressive Matrices (Raven, Court, & Raven, 1977), as opposed to Raven's APM.   Furthermore, while the OSpans in Engle et al.'s (1999) and the present research were approximately equally reliable, Engle et al.'s RSpans were much less reliable than those in the present research—Engle et al.'s had a coefficient alpha of .53, as compared to a .78 in the present research. Had Engle et al.'s scores been more reliable, and had they compared their spans to different measures of reading comprehension ability and nonverbal intelligence, their results might have looked more like those in the present research.

In addition to the structure of the span tasks, the characteristics of the subjects participating in them might have influenced the present results. That is, Baddeley et al. (1985), Daneman and Carpenter (1980), and Engle et al. (1999) might have obtained higher correlations between span reading comprehension ability than the present research because they used subjects with a wider range of abilities. Baddeley et al.'s (1985) subjects were community members, while Daneman and Carpenter (1980) and Engle et al.'s (1999) subjects were college students of all levels. In contrast, this research utilized upper-level undergraduates. Similarly, Light and Anderson (1985) used upper-level students and alumni in their studies and obtained non-significant RSpan-comprehension correlations of .26 and .07.

The characteristics of the subjects might have also contributed to the unexpected results on the homograph- and homophone-suppression measures (see Figures 2-8). Specifically, high performers on the comprehension measures were expected to suppress irrelevant word meanings, while low performers were not. Some of the present results, however, indicated that both the high and low performers suppressed the irrelevant word meanings. Similarly, Watts and Gough (1995) used upper-level undergraduates in their experiments and found unexpected results. In contrast, the college students and military recruits participating in Gernsbacher and colleagues' original research (Gernsbacher & Faust, 1991, experiment 1; Gernsbacher & Faust, experiment 4) were more heterogeneous.

Many aspects of the meaning-suppression protocol might have resulted in the inconsistencies. The most noticeable differences between the present research and Gernsbacher and colleagues' original suppression studies pertained to the assessment of comprehension ability. Watts and Gough (1995), who measured reading comprehension ability via the NDComp, attempted several times to produce the expected results but were unable to do so. Thusfar, the only published research obtaining the expected results utilizing the experiments discussed here was by Gernsbacher and her colleagues. It is therefore possible that they can only be obtained using one comprehension measure, the Multi-Media Comprehension Battery (Gernsbacher & Varner, 1988), which Gernsbacher and her colleagues have used in all of their published suppression research. The relatively low correlations among the three reading comprehension measures used in this research support that possibility.

Despite the inconsistencies with Gernsbacher and colleagues' data (Gernsbacher and Faust, 1991; Gernsbacher et al., 1990; see Figure 2), a distinct pattern of results did emerge among the low and high performers (see Tables 1-6). On both meaning-suppression tasks, low performers on the NDComp and sem responded more slowly, overall, than the high performers, whereas the low performers on the ADRP tended to respond more quickly on the meaning-suppression tasks than the high performers. The suppression tasks, like the NDComp and sem tasks, required the subjects to respond as quickly as possible, with the sem task quite methodologically similar to the meaning-suppression tasks. In contrast, many high performers on the untimed ADRP might have spent extra time rereading the passages or checking over their answers, but applying similar caution on the meaning-suppression tasks would have led to relatively poor performance. That might explain why the high ADRP performers did relatively worse on the meaning-suppression tasks than the low ADRP performers.

Unexpected results involving the comprehension measures themselves, as well as other measures, might have been associated with a number of issues. Because most of the experiments did not begin until the second half of the semester, laboratory sessions were sometimes longer than an hour in order to ensure that all of the tasks were completed. Hence, fatigue from those long sessions might have adversely affected the data. Additionally, the subjects knew that they would receive credit for simply participating in all of the laboratory activities, which took a total of approximately six hours and sometimes more. The sheer number of tasks, in combination with the small rewards, might have resulted in low levels of motivation. Though the subjects' scores on most of the measures were internally consistent, the aforementioned issues might have lead to inconsistent performance from one laboratory session to the next. And while typical American psycholinguistic research has involved native English speakers, few have involved native-English-speaking, college upperclassmen without a history of reading difficulty, meaning that the subjects in this study were far more homogenous than most. Studying individual differences is difficult when the subjects are relatively similar, and external variables might be confounding the data.

It's All Relative

This study provided evidence supporting the idea that knowledge and intelligence are essential, but the other skills needed for reading comprehension, e.g., additive or subtractive updating, vary with the nature of the text to be comprehended. We have additionally learned that central executive processing and the suppression of irrelevant information are completely unrelated to each other, both within and outside the realm of reading comprehension ability. Plus, we have confirmed that the notion of central executive processing is extremely general, meaning that more work must be done to separate the individual executive functions from the different strategies people use during executive functioning tasks. That might entail the use of executive functioning tasks in which no strategy, or the same strategy, would be used by all of the participants. Lastly, we have learned that our ability to comprehend text might not be influenced by our capacity to deliberately or automatically suppress irrelevant information.

The results of the present study allowed for several interpretations because they might have been influenced by a number of factors. Although some of those factors were subject-related, others stemmed from a deficient knowledge of what some of the measures, including the comprehension measures and the span tasks, actually measured. Thus, future research on the independence or dependence of various cognitive mechanisms related to reading comprehension ability must entail the breaking down of the general constructs, such as reading comprehension and central executive processing, into more specific constructs. Some of the previously cited executive functioning research has already embarked on that task (cf. Friedman & Miyake, 2004; Miyake et al., 2000; Salthouse et al., 2003; for discussion, see Harnishfeger, 1995; Nigg, 2000; Wilson & Kipp, 1998), and research cited by Daneman and Merikle (1996) acknowledged the existence of specific types of reading comprehension.

Once we pinpoint a specific central executive or suppression mechanism, we can examine whether it relates to our ability to comprehend certain elements of discourse. For example, making causal inferences during reading might involve additive updating—we have to update our schema from its initial state to its resulting state—so better additive updaters might process text pertaining to the resulting state faster than their poorer counterparts. That is, if a group of subjects were given text representing an initial state, such as, *The milk carton had a small hole in it*, the better additive updaters might be able to process, *We had to clean the refrigerator,* faster than poorer additive updaters. The *n*-back task might already measure additive updating, and multiple versions, all somewhat different from the version used here, have been administered (see, e.g., Jonides et al. 1997; Lieberman & Rosenthal, 2001; Roberts & Gibson, 2002). Any of them might reliably measure additive updating.

If a form of central executive processing, such as additive updating, and a form of suppression were both associated with a specific element of reading comprehension, e.g., making causal inferences, we could then examine whether they shared a mother-daughter or sisterly relationship, or no relationship at all, within the realm of that aspect of comprehension. We could then continue the research with the investigation of other specific cognitive mechanisms and elements of reading comprehension. The result would likely be a number of family trees, each sharing at least one relative with another.

Appendix A

Personal-Information Survey

PSY 338K Questionnaire—Spring, 2004

Lab Number: ________ Initials: ____________ Age:_________ Year in College: ___________ Today's Date: ______________

Gender: __________ Date of Birth:__/___/___ Which hand do you write with? _____________

1. Have you ever been considered dyslexic? Y/N (circle one)
2. Did you have difficulty learning to read? Y/N (circle one)

3a. The first language you learned to speak fluently:________________________________

If the answer to question #3a was English, skip to question 4. Otherwise, please complete the following before moving onto question 4:

3b. The approximate age at which you started to use English frequently in your daily life: ________

3c. List the languages that you read **better** than English: ________________________________

3d. List the languages that you read **equally well as** English: ________________________________

4. Place of birth:________________

If you were born outside the United States, how old were you when you moved here?

5. Did you spend any school years outside the country? If so, please list below (include countries where you lived) or indicate Not Applicable.

6a. Before taking it in this class, had you ever taken the Nelson-Denny Test? Circle one. (Yes / No / I don't know)

6b. Before taking it in this class, had you ever taken a Raven Progressive Matrices Test, which measures nonverbal intelligence?
Circle one: (Yes / No / I don't know)

6c. Before taking it in this class, had you ever taken the Advanced Degrees of Reading Power Test? Circle one. (Yes / No / I don't know)

7a. Do you BELIEVE that you have attention problems? Circle one: (Yes / No)

7b. Do you take medicine to improve your attention? (Yes / No) If so, how often? ___________________

7c. Have you been diagnosed with any attention problems? (Yes / No) If so, when? ___________________

8a. Have you been diagnosed with Test Anxiety, or do you suspect that you have Test Anxiety? Circle one: (Yes / No)

8b. Rate your perceived multiple-choice test-taking skills (circle one.)

1	2	3	4	5
very poor				very good

Appendix B

Sem Stimuli

Air is lighter than lead.
A carrot is a vegetable.
Automobiles run on gasoline.
Wednesday is the day after Tuesday.
Jesus was the son of Mary.
Thomas Edison invented the electric light.
Adolf Hitler was a Nazi.
Birds fly.
Lenin and Stalin were communists.
Men seldom wear dresses.
London is located in England.
There are 12 months in the year.
It's women, not men, who give birth.
The current president of the United States is George W. Bush.
Democrats are generally more liberal than Republicans.
An astronaut has walked on the moon.
A coach is usually older than his players.
A blind man cannot see.
You keep milk in the refrigerator.
Rock bands tend to be louder than string quartets.
A surgeon uses a scalpel more often than a lawyer does.
Gasoline is more expensive than water.
The leaves turn color in autumn.
Horror movies are more frightening than comedies.
When he was a baby, Bill Gates was smaller than he is now.
You can get a hamburger at McDonald's.
Men tend to be taller than women.
Frogs can jump higher than turtles.
Women wear earrings more often than men.
Instant coffee can be prepared faster than regular.
Most people would rather be rich than poor.
Someone who wore a tire around his neck might be considered odd.
Nuns pray.
Poetry often rhymes.
There are 50 states in the U.S.
Crocodiles have more teeth than birds.
Water is wet.
Houston is in Texas.
January comes before February.
Elvis Presley died many years ago.
Employers want employees who work hard.
Dinosaurs are extinct.
Most Americans wear shoes.
It's probably hard to sleep on a bed of nails.
Letters come in envelopes.
Multiplication is harder than addition.
The Empire State Building is taller than a basketball player.
Romeo loved Juliet.
Los Angeles is in California.
Many cars are made in Japan.

Babe Ruth was a famous hockey player.
Columbus discovered Antartica in 1492.
The sun comes out at night.
Berlin is the capital of France.
Lincoln was the first president of the United States.
One is the square root of eight.
Marilyn Monroe was born in the 17th century.
Salads usually consist of wool.
San Antonio is north of Austin.
Coca Cola is yellow.
Automobiles are usually repaired by secretaries.
Tom Hanks, the movie actor, has a uterus.
Most bananas are imported from Canada.
A hamburger is a piece of meat between two pianos.
Dogs often chase planets.
You go to the dentist to get a haircut.
Basketballs are square.
The Daily Texan is published once a month.
Almost everyone would rather be dead than alive.
Breakfast is always served in the afternoon.
Members of sororities are usually men.
You could put a skyscraper in a wheelbarrow.
When a minister delivers a sermon, he yodels.
There are seven singers in a quartet.
The Alamo is in Florida.
Men's shirts have two sleeves but women's have three.
Drums boil.
George Washington fought in the Vietnam War.
Pencils cost about fifty dollars apiece.
Every book contains a cucumber.
At least a million people have climbed to the top of Mount Everest.
We eat supper with our hands in our pockets.
Lemons are sweeter than marshmallows.
William Shakespeare wrote the Bible.
The Amazon River flows into the Dead Sea.
Plastic consists of bread and water.
Napoleon married Muhammad.
A mule can give birth to a hummingbird.
Most of us keep diamonds in the dishwasher.
Albert Einstein was a professional sportscaster.
Rain makes the streets dry.
Godzilla was a mouse.
Our eyes are in the back of our heads.
There's only one bridge over the Mississippi River.
David Letterman can lift an elephant.
A triangle has four sides.
Walt Disney could probably swallow a cactus.
Telegraphy is a disease.
Every human being has four legs.
Hardly anyone in America can speak English.

Appendix C

Which-Came-First Stimuli

Alexander the Great   Frederick the Great
American Revolution   French Revolution
Thomas Acquinas   Aristotle
John Locke   John Stuart Mill
Picasso   Velaszquez
Franklin D. Roosevelt   Theodore Roosevelt
Harvard University   Oxford University
classical Greece   classical Rome
Ferdinand and Isabella   Victoria and Albert
Crusades   Punic Wars
Geoffrey Chaucer   William Shakespeare
calculus   geometry
Christianity   Judaism
Gettysburg   Waterloo
Christopher Wren   Frank Lloyd Wright
John Dewey   Alexis de Tocqueville
Charles Darwin   Isaac Newton
Winston Churchill   Benjamin Disraeli
Charlemagne   Louis XIV
Michelangelo   Rembrandt
cathedrals   pyramids
Lenin   Stalin
Columbus   Confucius
Simon Bolivar   Hernando Cortes
automobiles   railroads
Saint Augustine   Socrates
Beethoven   Vivaldi
Mozart   Tchaikovsky
Machiavelli   Mussolini
Martin Luther   John Wesley
Andrew Carnegie   Andrew Jackson
Jesus Christ   Mohammed
Herbert Hoover   Woodrow Wilson
Galileo Gallilei   Isaac Newton
Thomas Edison   Benjamin Franklin
Julius Caesar   Leonardo da Vinci
Byzantine Empire   Magna Carta
Mahatma Gandhi   Jean Jacques Rousseau
Marie Antoinette   Florence Nightingale
Immanuel Kant   Friedrich Nietzsche
Enrico Caruso   Mario Lanza
French & Indian Wars   Spanish-American War
Italy   Yugoslavia
Albert Einstein   President Truman
President Eisenhower   President Truman
U.S. Civil War   Russian Revolution
telephone   television
Vietnam War   World War II
radio   telegraphy
Great Depression   Prohibition

Appendix D

Lexical-Decision Stimuli

Words	**Pseudowords**
drain	droze
dense	doolb
trick	traff
pound	pleeg
grand	gloke
crime	crade
climb	clupe
rough	reasp
broad	bress
track	troud
serve	srutt
month	milth
noise	noote
stage	spond
fresh	frusk
fight	furch
chief	chent
brown	brump
stone	snept
bring	brull
clear	clore
space	sterk
south	sloom
front	fleek
group	goomb

Appendix E

Pseudoword Decoding Task Form

Lab Number:______ Subject must initial here:_______ Date:______

Word	1) Make a check mark below if you know the pronunciation is correct; otherwise record it in the space provided. 2) If applicable, indicate the number of incorrect pronunciations made before and after the correct one.	"X" here if sounded out **slowly** or leave blank if not.
ADJOIST		
BELTH		
CROOB		
DRICK		
FROTT		
GASSAGE		
GHUSKLY		
GRIMPLE		
HARIBEL		
LERNACE		
LOTTLE		
PHANK		
PRUCKLE		
RAPTION		
SCORTH		
SHENNY		
SHINTER		
SPARCH		
SPISMA		
STOAP		
TISPOR		
WHULSE		
WIMFUL		
YUTTON		
ZEEVISH		

Appendix F

Homograph-Suppression Task Stimuli

Test words (T) and their corresponding sentences: experimental (E) and control (C) or filler (F).

T: HEAD
E: He went to the temple
C: He went to the sanctuary

T: BELL
E: She put on the ring
C: She put on the necklace

T: CHURCH
E: He calculated the mass
C: He calculated the length

T: JOKE
E: She started to gag
C: She started to cough

T: HERB
E: He talked to the sage
C: He talked to the prophet

T: TENNIS
E: She lit the match
C: She lit the lamp

T: WAITER
E: She jabbed with the tip
C: She jabbed with the end

T: HAIR
E: He built the shed
C: He built the shack

T: CELERY
E: She tried to stalk
C: She tried to creep

T: PAN
E: He smoked the pot
C: He smoked the cigar

T: CATCHER
E: She drank from the pitcher
C: She drank from the thermos

T: COW
E: He wanted to steer
C: He wanted to guide

T: MUSIC
E: She picked up the rock
C: She picked up the pebble

T: COLUMN
E: He liked to row
C: He liked to swim

T: WATER
E: He played some bridge
C: He played some chess

T: QUALITY
E: She paid the fine
C: She paid the ticket

T: DRINK
E: He tried to punch
C: He tried to throw

T: FRUIT
E: She mentioned her date
C: She mentioned her appointment

T: LOOK
E: She liked the watch
C: She liked the clock

T: ARMY
E: She drank the draft
C: She drank the beer

T: SLIDE
E: She wore the slip
C: She wore the nightgown

T: JET
E: He killed a fly
C: He killed a moth

Appendix F, cont.

T: OUT
E: He paid the check
C: He paid the bill

T: ANIMAL
E: He helped to seal
C: He helped to shut

T: CARDS
E: He walked on the deck
C: He walked on the tile

T: DISHES
E: He went to China
C: He went to France

T: FOOL
E: She tried to jerk
C: She tried to yank

T: DUCK
E: She went to a quack
C: She went to a dentist

T: JELLO
E: He grew some mold
C: He grew some fungus

T: NEWS
E: He started to press
C: He started to push

T: FIGHT
E: She lifted the box
C: She lifted the crate

T: FIRST
E: She paused a second
C: She paused a while

T: KITCHEN
E: She did not want to sink
C: She did not want to drown

T: BRACELET
E: He wanted to charm
C: He wanted to smile

T: CEREAL
E: She used to bowl
C: She used to ski

T: MACHINE
E: He fed the crane
C: He fed the bird

T: KING
E: She drew using the ruler
C: She drew using the stencil

T: OVEN
E: He lived on the range
C: He lived on the farm

T: COMMAND
E: He bought a full order
C: He bought a full meal

T: MATTRESS
E: She looked forward to spring
C: She looked forward to summer

T: TOILET
E: He wanted to stall
C: He wanted to wait

T: TRUTH
E: She said the knife was blunt
C: She said the knife was sharp

T: BROKEN
E: He was part of the cast
C: He was part of the troupe

T: TOCK
E: She stepped onto the tick
C: She stepped onto the insect

T: DIFFICULT
E: He said the wood was hard
C: He said the wood was solid

T: TRASH
E: She gave away the litter
C: She gave away the puppy

T: TRANSPLANT
E: He played the organ
C: He played the piano

T: CEILING
E: She was a jazz fan
C: She was a jazz lover

Appendix F, cont.

T: FUNERAL
E: He did not want to wake
C: He did not want to sleep

T: FACTORY
E: She watered the plant
C: She watered the shrub

T: JUDGE
E: He nailed up the panel
C: He nailed up the siding

T: ALCOHOL
E: He played some gin
C: He played some solitaire

T: COMPLAINT
E: She ate the beef
C: She ate the lamb

T: PRAYER
E: She acted with grace
C: She acted with poise

T: STEREO
E: She introduced the speaker
C: She introduced the expert

T: INCH
E: He stepped on her foot
C: He stepped on her toe

T: CASH
E: He tasted the mint
C: He tasted the candy

T: CLOTHES
E: She bent the iron
C: She bent the metal

T: CRIME
E: She used the mug
C: She used the goblet

T: PORK
E: He began to chop
C: He began to saw

T: CONCEAL
F: She had to hide

T: ASSAULT
F: She had to attack

T: COLLEGE
F: She attended the school

T: SNACK
F: She tried the dip

T: ABILITY
F: She had the power

T: SERMON
F: He began the speech

T: QUIT
F: She wanted to stop

T: TEXT
F: He left a book

T: FEELING
F: He resisted the urge

T: PUZZLE
F: He solved the problem

T: PURCHASE
F: She went to shop

T: PET
F: He had a dog

T: ON
F: She flipped the switch

T: CHARITY
F: She helped the poor

T: ASSIGNMENT
F: He finished the homework

T: GEM
F: He had the jewel

T: PAPER
F: He prepared the report

T: TRIAL
F: She went to court

Appendix F, cont.

T: HIKE
F: He walked on the trail

T: TROPHY
F: He wanted the award

T: LIMB
F: She climbed the branch

T: GARBAGE
F: She went to the dump

T: SHOOT
F: He played some pool

T: HOME
F: He lived in the house

T: BAKE
F: She made a cake

T: PEPPER
F: He forgot the spice

T: TOWN
F: He went to the village

T: WHIRL
F: She began to spin

T: DEPART
F: He tried to leave

T: BROOK
F: He crossed the stream

T: LATE
F: She waited for him

T: GREASE
F: He fried the bacon

T: PRICE
F: He paid for it

T: GONE
F: He really missed her

T: MISTAKE
F: She dropped the platter

T: TOURIST
F: He was a traveler

T: CAB
F: She expected the taxi

T: SPORT
F: He joined the team

T: PAL
F: He had a friend

T: REPLY
F: She had to answer

T: LAUGH
F: He wanted to chuckle

T: START
F: She asked to begin

T: HOT
F: She cooked on the stove

T: NYLON
F: She washed the stockings

T: SOB
F: She felt so sad

T: MIX
F: She turned on the blender

T: TAKE
F: She tried to grab

T: ACTOR
F: She hated the movie

T: THINK
F: She used to believe

T: IDEA
F: He played with the thought

T: FETCH
F: He helped to bring

T: BAD
F: He did it poorly

Appendix F, cont.

T: FEMALE
F: He saw the woman

T: MEAL
F: She went to lunch

T: TOWN
F: He went to the village

T: POND
F: She liked the lake

T: PICTURE
F: He shot with a camera

T: ENTER
F: She walked through the door

T: SING
F: He heard the opera

T: COUCH
F: She sat on the sofa

Appendix G

Homophone-Suppression Task Stimuli

Test words (T) and their corresponding sentences: experimental (E) and control (C) or filler (F).

T: COLUMNS
E: She arranged the rose
C: She arranged the flowers

T: PLANK
E: He examined the timbre
C: He examined the vibration

T: SERGEANT
E: He just stared at the kernel
C: He just stared at the seed

T: LEARN
E: She had never been taut
C: She had never been rigid

T: POETRY
E: He wrote the pros
C: He wrote the officials

T: TREE
E: She looked at the fur
C: She looked at the scarf

T: SPIRIT
E: He mended his sole
C: He mended his shoe

T: GROCERY
E: He put away the sax
C: He put away the trumpet

T: BLOOD
E: She spoke of the vain
C: She spoke of the arrogant

T: LARGE
E: She was unsure of the sighs
C: She was unsure of the moans

T: CANDY
E: She just loved suites
C: She just loved lodges

T: CAMP
E: He changed the tense
C: He changed the verb

T: KETTLE
E: He prepared the tee
C: He prepared the racket

T: YOUTH
E: He was a miner
C: He was a logger

T: FRUIT
E: She liked the pair
C: She liked the group

T: STREAM
E: He walked toward the creak
C: He walked toward the noise

T: WEIRD
E: He liked the bazaar
C: He liked the market

T: DROPS
E: She felt the rein
C: She felt the leash

T: PIZZA
E: He handled the doe
C: He handled the calf

T: SAND
E: She took a picture of the beech
C: She took a picture of the elm

T: BALD
E: He lost his hare
C: He lost his collie

T: FOOD
E: She almost ruined the meet
C: She almost ruined the contest

Appendix G, cont.

T: THIRD
E: She walked forth
C: She walked forward

T: UNCLE
E: She didn’t like the ant
C: She didn’t like the moth

T: CALM
E: He had lots of patients
C: He had lots of students

T: SCRATCH
E: She was hurt by the long clause
C: She was hurt by the long contract

T: NICE
E: She was a real deer
C: She was a real goat

T: COTTON
E: He put up the bail
C: He put up the ransom

T: CARNIVAL
E: She really liked the fare
C: She really liked the price

T: PROPERTY
E: He fixed his gait
C: He fixed his stride

T: DRUG
E: He grabbed the heroine
C: He grabbed the hero

T: ENVELOPE
E: She was attracted to the male
C: She was attracted to the man

T: ACHE
E: He couldn’t get over the pane
C: He couldn’t get over the window

T: BASEMENT
E: He went down to see the seller
C: He went down to see the buyer

T: SMELL
E: She couldn’t identify the cent
C: She couldn’t identify the dime

T: GALLONS
E: She measured the quartz
C: She measured the topaz

T: CHURCH
E: She began to prey
C: She began to hunt

T: DOCK
E: She stood by her peer
C: She stood by her friend

T: MUSIC
E: He held the last cord
C: He held the wire

T: DARK
E: He aimlessly walked into the knight
C: He aimlessly walked into the king

T: DAY
E: She worked hard for all the weak
C: She worked hard for all the poor

T: VODKA
E: She was overcome by all the boos
C: She was overcome by all the applause

T: PANTS
E: She wished she had a new pair of genes
C: She wished she had a new pair of parents

T: HANDS
E: He couldn’t help noticing her feat
C: He couldn’t help noticing her accomplishment

T: OPPORTUNITY
E: She looked forward to the chants
C: She looked forward to the songs

T: BREATHE
E: She was disturbed by the heir
C: She was disturbed by the grandson

T: TELL
E: He could never imagine such an unusual tail
C: He could never imagine such an unusual neck

T: STEPS
E: She tried to get over all the stares
C: She tried to get over all the jeers

Appendix G, cont.

T: LOOK
E: He couldn’t believe it so he went to sea
C: He couldn’t believe it so he went to Europe

T: BEER
E: She was completely unaffected by the whine
C: She was completely unaffected by the complaint

T: WATER
E: He put away the hoes
C: He put away the shovels

T: THREAD
E: She was proud of what she had sown
C: She was proud of what she had planted

T: BARGAIN
E: She heard about the huge sail
C: She heard about the huge banner

T: POUNDS
E: He was hampered by the abnormal wait
C: He was hampered by the abnormal delay

T: PERMITTED
E: She realized her screams weren’t aloud
C: She realized her screams weren’t real

T: MOST
E: She said the last car was leased
C: She said the last car was rented

T: AUDITORIUM
E: They met in the big haul
C: They met in the big raid

T: SLICE
E: He begged at the table for another peace
C: He begged at the table for another truce

T: OPENING
E: She saw some of the whole
C: She saw some of the sections

T: GIFTS
E: She was intrigued by his presence
C: She was intrigued by his poise

T: SPICY
F: She likes the food chili

T: DISALLOWED
F: She was with the banned

T: MONEY
F: He spoke of the profit

T: RIVER
F: He cursed the dam

T: FINLAND
F: She said the fish was discovered by a Finn

T: SQUATTED
F: He kneeled to secure the ducked

T: CHILD
F: He was blinked by the son

T: GLUE
F: She talked about how they paste

T: WRITTEN
F: He was opposed to the verses

T: ABSENT
F: She said the light rain was missed

T: LEAVE
F: She wanted to stop the flea

T: PEDAL
F: He saw the brake

T: SUMMIT
F: He could see the peak

T: ROD
F: She picked up the pole

T: EMPEROR
F: He was aware of the throne

T: ADD
F: He thought of sum

T: CHOIR
F: She said her favorite was choral

T: SPECULATED
F: She sought after the guessed

Appendix G, cont.

T: LIVESTOCK
F: She heard the sound of the herd

T: BOAT
F: He knew the article was the oar

T: FIGHT
F: He was engaged in a duel

T: UNINTERESTING
F: She knew of a bore

T: OCEAN
F: He got tangled up in the wave

T: ROB
F: She was unaware of the steal

T: PAYMENT
F: He wanted to get the raise

T: BUILDING
F: He dined near the capitol

T: GOVERNMENT
F: She required the city council

T: THROAT
F: She talked to the hoarse

T: TIE
F: She did knot

T: ROYALTY
F: He framed the prince

T: SINGLE
F: She lived by herself

T: PRESIDENT
F: He overthrew the leader

T: SHARP
F: He posted the note with a tack

T: MOUTH
F: She examined the dog’s tongue

T: TRIVIA
F: He knew the information

T: LINK
F: She took apart the chain

T: CANVAS
F: He painted the portrait

T: CLOTHING
F: She hated doing the laundry

T: DESTINY
F: He realized his fate

T: PUNISHMENT
F: He said it was the penalty

T: EMPLOYED
F: She wanted to get hired

T: DIRT
F: She stood in the mud

T: AGREEMENT
F: He wanted to write the treaty

T: DISTURBANCE
F: She made quite a riot

T: ATTORNEY
F: He spoke to the lawyer

T: SUITCASE
F: She packed the luggage

T: FISH
F: He shot the salmon

T: BRIDE
F: She rehearsed the wedding

T: HOME
F: He moved into his new house

T: DINNER
F: She roasted the chicken

T: UP
F: He competed in the high jump

T: BRAVERY
F: She liked his courage

Appendix G, cont.

T: FAR
F: He knew about the distance

T: EARRING
F: She liked the diamond

T: BALLET
F: She began to dance

T: CHIN
F: He had a dimple

T: HOSPITAL
F: She performed the surgery

T: MANY
F: He had a bunch

T: ERASER
F: She looked at the pencil

T: RISE
F: He went up the escalator

Appendix H

RSpan and OSpan Stimuli

Reading Span	**Operation Span**
No matter how much we talk to him, he is never going to change. ? SKIN	IS 10 - 5 = 5 ? WIFE
The prosecutor's dish was lost because it was not based on fact. ? FLOOR	IS 10 - 1 = 11 ? CLASS
Every now and then I catch myself swimming blankly at the wall. ? CHAIN	IS 1 + 2 = 1 ? PAINT
We were fifty lawns out at sea before we lost sight of land. ? CHART	IS 3 + 1 = 2 ? CLOUD
He is afraid of heights and refuses to fly on a plane. ? SOAP	IS 2 - 1 = 1 ? PIPE
Throughout the ordeal, the hostages never seemed to lose hope. ? TEA	IS 1 + 3 = 4 ? EAR
The young pencil kept her eyes closed until she was told to look. ? BRICK	IS 9 + 2 = 17 ? FLAME
People tend to go on feathers when they want to lose weight. ? NEST	IS 9 - 7 = 4 ? BELT
I cheered loudly, knowing I would have a tall voice the next day. ? BUMP	IS 4 - 2 = 6 ? BEAN
She was asked to stop at the new mall to pick up several items. ? TOOTH	IS 9 - 6 = 3 ? SHELF
When she shops she always looks for the lowest flood. ? CROW	IS 4 + 1 = 4 ? FORK
When I get up in the morning, the first thing I do is feed my dog. ? PATH	IS 1 + 1 = 2 ? HOLE
When it is cold, my mother always makes me wear a cap on my head. ? BOX	IS 4 + 2 = 6 ? MAP
All parents hope their list will grow up to be intelligent. ? COW	IS 6 + 1 = 15 ? PET
When the couple moved to Japan, their wish had a huge garage sale. ? BIRD	IS 3 + 2 = 9 ? SAND
At school yesterday morning, his daughter heard a terrible plum. ? SACK	IS 6 - 3 = 2 ? JAIL
In the fall, my gift and I love to work together in the yard. ? JAR	IS 8 - 2 = 2 ? TIN
Unaware of the hunter, the deer wandered into his shotgun range. ? LAKE	IS 8 - 1 = 7 ? MILK
Since it was the last game, it was hard to cope with the loss. ? FROG	IS 1 + 6 = 7 ? CAVE
Because she slices early, she usually gets a good parking spot. ? HEAD	IS 6 + 3 = 3 ? HAND
The only furniture he had in his first bowl was his waterbed. ? SEAT	IS 5 - 1 = 5 ? NECK
Last year, he was given detention for running in the hall. ? RUST	IS 10 + 2 = 12 ? FERN
The huge clouds covered the silk tie, and the rain began to fall. ? BELL	IS 2 + 2 = 3 ? COAT
After one date I knew that her friend simply was not my type. ? CORN	IS 7 + 6 = 13 ? HALL
He broke his arm when he fell from the tree and onto the ground. ? DRILL	IS 1 + 2 = 3 ? BEAST
Most people agree that Monday is the worst stick of the week. ? CALF	IS 10 - 1 = 8 ? YARN
On warm and sunny afternoons, I like to walk in the neighborhood. ? KING	IS 4 + 6 = 10 ? FISH
With discipline and determination, he knew he could win the race. ? GOOSE	IS 6 - 1 = 5 ? CHEEK
My mother has always told me that it is not polite to shine. ? COAL	IS 3 + 2 = 4 ? STAR
A person should never be disliked based on his or her race. ? RASH	IS 4 + 5 = 9 ? GERM
The school marching band decided to play two out of three songs. ? GOAT	IS 1 + 7 = 8 ? DOCK
Raising children requires the ability to be firm and a lot of sip. ? FARM	IS 6 - 3 = 5 ? WALL
The gathering crowd turned to look when they heard the gun shot. ? ROPE	IS 9 - 2 = 7 ? FUEL
As soon as I get done taking this envy I am going to go home. ? WAX	IS 10 - 2 = 6 ? HEN
She opened her purse and realized that she did not have any money. ? CUP	IS 8 - 7 = 1 ? CAP
She wanted a garden in her yard, but the soil was mostly clay. ? MOUTH	IS 10 + 3 = 13 ? STORE
He said that they would get a surprise if they listened briskly. ? DRESS	IS 1 + 6 = 2 ? BEACH
She stopped dating the light when she found out he had a wife. ? CAKE	IS 3 + 2 = 7 ? LAMP
He was so tired of studying, he could not read another page. ? NET	IS 6 + 3 = 9 ? FOX
Although he is sarcastic at times, he can also be very sweet. ? COMB	IS 3 - 1 = 2 ? CONE
She will ask her agent how much the flight to Mexico will cost. ? SWEAT	IS 8 - 4 = 4 ? GRAPH
The sugar could not believe he was offered such a great deal. ? BRANCH	IS 9 + 1 = 11 ? BRIDGE

Appendix I

Word-Memory Task Form

Word Memory

Lab#______ S’s Initials________ Date______

1) __________ __________ __________ __________ __________ __________

2) __________ __________ __________ __________ __________ __________

3) __________ __________ __________ __________ __________ __________

Appendix J

Sample Span Task Form

Reading Span

Lab#_____ S’s Initials_______ Date_____

****
Practice

a) ____________ ____________

b) ____________ ____________

c) ____________ ____________

****

1) ____________ ____________ ____________ ____________ ____________

2) ____________ ____________ ____________ ____________ ____________

3) ____________ ____________ ____________ ____________ ____________

4) ____________ ____________ ____________ ____________ ____________

5) ____________ ____________ ____________ ____________ ____________

6) ____________ ____________ ____________ ____________ ____________

7) ____________ ____________ ____________ ____________ ____________

8) ____________ ____________ ____________ ____________ ____________

9) ____________ ____________ ____________ ____________ ____________

10) ____________ ____________ ____________ ____________ ____________

11) ____________ ____________ ____________ ____________ ____________

12) ____________ ____________ ____________ ____________ ____________

References

Baddeley, A., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory [Electronic version]. *The Quarterly Journal of Experimental Psychology, 51A*, 819-852.

Baddeley, A. D. (1986). *Working Memory*. New York: Oxford University Press.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7,* 85-97. Retrieved September 18, 2004, from PsycINFO database.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in Research and Theory* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluid reading [Electronic version]. *Journal of Memory and Language, 24*, 119-131.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74,* 506-521.

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. *Journal of Educational Psychology, 86*, 244-255.

Borella, E., & de Ribaupierre, A. (2001). Age differences in text comprehension: A question of memory? Poster retrieved March 2, 2004, from professor A. de Ribaupierre's Web site, University of Geneva Psychology department: http://www.unige.ch/fapse/PSY/persons/deribaupierre/downloads/conf/posters/Borella_Edimburgh_2001.pdf.

Borella, E., & de Ribaupierre, A. (2003, September). Individual differences in reading comprehension in children: The role of inhibition, working memory and decodings [sic] skills. Poster presented at the 13th conference of the European Society for Cognitive Psychology, Granada, Spain. Poster retrieved March 2, 2004, from professor A. de Ribaupierre's Web site, University of Geneva Psychology department: http://www.unige.ch/fapse/PSY/persons/deribaupierre//downloads/conf/posters/2003_ESCOP_Borella_poster.pdf.

Brown, J. L., Fishco, V. V., & Hanna, G. (1991). *The Nelson-Denny Reading Test.* Chicago: Riverside.

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology [Electronic version]. *Developmental Neuropsychology, 15*, 421-442.

Burt, J. S., & Fury, M. B. (2000). Spelling and adults: The role of reading skills and experience [Electronic version]. *Reading and Writing: An Interdisciplinary Journal, 13*, 1-30.

Calvert, J. B. (2001). *Alphabets, Ciphers, and Codes: A look at some interesting and useful lore.* Retrieved September 3, 2004, from professor J. B. Calvert's Web site: http://www.du.edu/~jcalvert/tel/codes.htm.

Cantwell, Z. M. (1966). Relationships between scores on the standard Progressive Matrices (1938) and on the D. 48 Test of non-verbal intelligence and three measures of academic achievement [Electronic version]. *Journal of Experimental Education, 34*, 28-31.

Carr, T. H., Brown, T. L., & Vavrus, L. G. (1985). Using component skills analysis to integrate findings on reading development. In T. H. Carr (Ed.), *The development of reading skills* (pp. 95-108). San Francisco, CA: Jossey-Bass.

Chen, R. S., & Vellutino, F. R. (1997). Predication of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research, 29*, 1-24.

Cunningham, A. E., Stanovich, K. E., & Wilson, M. R. (1990). Cognitive variations in adult college students differing in reading ability. In T. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 129-159). San Diego, CA: Academic Press.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72*, 656-669.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related to reading comprehension ability? [Electronic version]. *Learning and Individual Differences, 12*, 131-143.

De Beni, R., Palladino, P., Pazzaglia, F., & Cornoldi, C. (1998). Increases in intrusion errors and working memory deficit of poor comprehenders [Electronic version]. *The Quarterly Journal of Experimental Psychology, 51A*, 305-320.

Dempster, F. N., & Corkill, A. J. (1999). Individual differences in susceptibility to interference and general cognitive ability [Electronic version]. *Acta Psychologica, 101*, 395-416.

Distribution of trigrams (three-letter combinations) in a 1 mil. word corpus of English. (n.d.). Retrieved July 9, 2004, from Free University of Berlin, Institute for Computer Science Web site: http://www.inf.fu-berlin.de/lehre/SS96/DS/trigrams.tab.

Dixon, P., LeFevre, J. A., & Twilley, L. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*, 465-472.

Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S. Gorfein (Ed.), *On the Consequences of Meaning Selection* (pp.27-43). Washington, DC: American Psychological Association.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. Intelligence and the frontal lobe: The organization of goal-directed behavior [Electronic version]. *Cognitive Psychology, 30*, 257-303.

Educational Testing Service (2004, May). *Reliability of speeded number-right multiple-choice tests* (Report Series ETS-RR-04-15). Princeton, NJ: Author.

Engle, R. (2003a). [OSPAN.ebs]. Retrieved March 2, 2004, from the Georgia Institute of Technology, Attention and Working Memory Lab Web site: http://www.psychology.gatech.edu/renglelab/.

Engle, R. (2003b). [RSPAN.ebs]. Retrieved March 2, 2004, from the Georgia Institute of Technology, Attention and Working Memory Lab Web site: http://www.psychology.gatech.edu/renglelab/.

Engle, R. (2003c). [RSPAN readme.txt]. Retrieved March 2, 2004, from the Georgia Institute of Technology, Attention and Working Memory Lab Web site: http://www.psychology.gatech.edu/renglelab/.

Engle, R. W. (2001). What is working-memory capacity? In H. L. Roediger III & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297-314). Washington, DC: American Psychological Association. Retrieved October 25, 2004, from the Georgia Institute of Technology, Attention and Working Memory Lab Web site: http://www.psychology.gatech.edu/renglelab/.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses [Electronic version]. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972-992.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach [Electronic version]. *Journal of Experimental Psychology: General, 128*, 309-331.

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General, 133,* 101-135.

Gernsbacher, M. A. (1991). Cognitive processes and mechanisms in language comprehension: The structure building framework. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 27, pp. 217-263). San Diego, CA: Academic Press.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms [Electronic version]. *Psychological Science, 4*, 294-298.

Gernsbacher, M. A. (1996). The structure building framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of text understanding* (pp. 289-311). Hillsdale, NJ: Lawrence Erlbaum Associates. (Reprinted from *Language Comprehension as Structure Building*, chapter 6, by M. A. Gernsbacher, 1990, Hillsdale, NJ: Lawrence Erlbaum Associates)

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23*, 265-304.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 245-262.

Gernsbacher, M. A., & Robertson, R. R. W. (1995). Reading skill and suppression revisited [Electronic version]. *Psychological Science, 6*, 165-169.

Gernsbacher, M. A., & St. John, M. F. (2001). Modeling suppression in lexical access. In D. S. Gorfein (Ed.), *On the Consequences of Meaning Selection* (pp.47-65). Washington, DC: American Psychological Association.

Gernsbacher, M. A., & Varner, K. R. (1988). *The multi-media comprehension battery.* (Tech. Rep. No. 88-07). Eugene, OR: University of Oregon, Institute of Cognitive and Decision Sciences.

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 430-445.

Gorfein, D. S. (Ed.). (2001). *On the Consequences of Meaning Selection*. Washington, DC: American Psychological Association.

Gorfein, D. S., Berger, S. A., & Bubka, A. (2000). The selection of homograph meaning: Word associations when context changes. *Memory & Cognition, 28,* 766-773.

Gough, P. B. (2002). *Literacy*. Lecture presented to students in the Psychology of Reading class, University of Texas, Austin, TX.

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 30,* 171-176.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading Comprehension Difficulties: Processes and Intervention* (pp. 1-13). Mahwah, NJ: Lawrence Erlbaum Associates.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education, 7,* 6-10.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and Inhibition in Cognition* (pp. 175-204). San Diego, CA: Academic Press.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in Research and Theory* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Hess, A. M. (1982). An analysis of the cognitive processes underlying problems in reading comprehension. *Journal of Reading Behavior, 14,* 313-333.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading &Writing, 2,* 127-160.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., et al. (1997). Verbal working memory load affects regional brain activation as measured by PET [Electronic version]. *Journal of Cognitive Neuroscience, 9,* 462-475.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity [Electronic version]. *Journal of Experimental Psychology: General, 130,* 169-183.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval [Electronic version]. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 336-358.

Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference [Electronic version]. *Journal of Experimental Psychology: General, 132*, 47-70.

Koch, W.R. (2003). *Reliability*. Lecture presented to students in the Psychometrics: Theory and Methods graduate-level class, University of Texas, Austin, TX.

Kyllonen, P. C., & Christal, R. E. (1990.) Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389-453.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity [Electronic version]. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1118-1133.

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? [Electronic version]. *The Quarterly Journal of Experimental Psychology, 49A,* 29-50.

Levy, B. A., & Hinchley, J. (1990). Individual and developmental differences in the acquisition of reading skills. In T. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 81-128). San Diego, CA: Academic Press.

Lewis, N. (Ed.) (1964). Inhibition. In *The New Roget's Thesaurus in Dictionary Form* (p. 263). New York: G. P. Putnam's Sons.

Lieberman, M. D., & Rosenthal, R. (2001). Why introverts can't always tell who likes them: Multitasking and nonverbal decoding [Electronic version]. *Journal of Personality and Social Psychology, 80,* 294-310.

Light, L. L., & Anderson, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology, 40*, 737-747.

Maki, R. H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review, 1,* 126-129.

Martin, C., Vu, H., Kellas, G., & Metcalf, K. (1999). Strength of discourse contexts as a determinant of the subordinate bias effect [Electronic version]. *Quarterly Journal of Experimental Psychology, 52A,* 813-839.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27*, 759-767.

McCormick, C. & Samuels, S. J. (1979). Word recognition by second graders: The unit of perception and interrelationships among accuracy, latency, and comprehension. *Journal of Reading Behavior, 11*, 107-118).

McKeown, M. G., Beck, I. L., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior, 25*, 3-18.

McNamara, D. S., & McDaniel, M. A. (2004). Suppressing irrelevant information: Knowledge activation or inhibition? [Electronic version]. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 465-482.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis [Electronic version]. *Cognitive Psychology, 41*, 49-100.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*, 621-640. Retrieved October 25, 2004, from PsycINFO database.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive [Electronic version]. *British Journal of Psychology, 81*, 111-121.

Naglieri, J. A., & Ronning, M. (2000). The relationships between general ability using the NNAT and SAT reading achievement. *Journal of Psychoeducational Assessment, 18*, 230-239.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220-246. Retrieved October 25, 2004, from PsycINFO database.

Nigg, J. T., Glass, J. M., Wong, M. M., Poon, E., Jester, J. M., Fitzgerald, H. E. (2004). Neuropsychological executive functioning in children at elevated risk for alcoholism: Findings in early adolescence. *Journal of Abnormal Psychology, 113,* 302-314. Retrieved September 18, 2004, from PsycINFO database.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills [Electronic version]. *Language and Cognitive Processes, 18,* 443-468.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension [Electronic version]. *Memory & Cognition, 29,* 344-354.

Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language, 24,* 59-88.

Passolunghi, M. C., Cornoldi, C., & De Liberto, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition, 27,* 779-790.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving [Electronic version]. *Journal of Experimental Child Psychology, 80,* 44-57.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement, 76,* 1-25.

Perfetti, C. A. (1985). *Reading Ability*. New York: Oxford University Press.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67,* 461-469.

Raven, J. C. (1947). *Advanced Progressive Matrices Sets I & II*. Oxford, England: Oxford Psychologists Press.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. In R. W.Engle, S. W. Tuholski, J. E. Laughlin, & A. R. A. Conway. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach [Electronic version]. *Journal of Experimental Psychology: General, 128*, 309-331.

The Riverside Publishing Company (2004). *Nelson-Denny Reading Test—Overview*. Retrieved September 2, 2004, from http://www.riverpub.com/products/ /group/ndrt/overview.html.

Roberts, R., & Gibson, E. (2002). Individual differences in sentence memory [Electronic version]. *Journal of Psycholinguistic Research, 31*, 573-598.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults [Electronic version]. *Journal of Experimental Psychology: General, 132*, 566-594.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.

Singer, M. H. & Crouse, J. (1981). The relationship of context-use skills to reading: A case for alternative experimental logic [Electronic version]. *Child Development, 52*, 1326-1329.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging [Electronic version]. *Cognitive Psychology, 33*, 5-42.

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Reading Behavior, 25*, 33-50.

Stanovich, K. E. (2000). *Progress in Understanding Reading: Scientific Foundations and New Frontiers*. New York: The Guilford Press.

Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20*, 51-68.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19*, 278-303.

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. In *Journal of Experimental Psychology: General, 121*, 15-23. (Reprinted from J. R. Stroop [1935]. *Journal of Experimental Psychology, 18*, 643-662)

Touchstone Applied Science Associates (2002). *Advanced DRP Handbook: T & U Test Forms* (pp. 31, 51). Brewster, NY: Author.

Touchstone Applied Science Associates. (1995). *Degrees of Reading Power Advanced Form T-2*. Brewster, NY: Author.

Towse, J. N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology, 89,* 77-101. Retrieved October 25, 2004, from PsycINFO database.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? [Electronic version]. *Journal of Memory and Language, 28,* 127-154.

Vellutino, F. R., Scanlon, D. M, & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers. [Electronic version]. *Journal of Learning Disabilities, 33*, 223-238.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension [Electronic version]. *The Quarterly Journal of Experimental Psychology, 49A*, 51-79.

Watts, J. L., & Gough, P. B. (1995). The suppression of irrelevant information in comprehension: A failure to replicate. Paper presented at the annual meeting, Southwest Psychological Association, San Antonio, TX.

Whitney, P., Arnett, P. A., Driver, A., & Budd, D. Measuring central executive functioning: What's in a reading span? [Electronic version]. *Brain and Cognition, 45,* 1-14.

Wilde, O. (1982). The picture of Dorian Gray. In R. Ellmann (Ed.), *The Picture of Dorian Gray and Other Writings* (pp. 1-193). New York: Bantam Books. (Original work published 1891)

Wilson, S. P., & Kipp, K. (1998). The development of efficient inhibition: Evidence from directed-forgetting tasks [Electronic version]. *Developmental Review, 18,* 86-123.

Wren, S. A. (1995). [Psdofreq.stm]. Unpublished stimuli.

Yuill, N., Oakhill, J., & Parkin, A. J. (1989). Working memory, comprehension ability and the resolution of text anomaly [Electronic version]. *British Journal of Psychology, 80,* 351-361.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1996). *The Educator's Word Frequency Guide [CD-ROM, DOS version]*. Brewster, NY: Touchstone Applied Science Associates.

Zwaan, R. A., & Truitt, T. P. (2000). Inhibition of smoking-related information in smokers and nonsmokers [Electronic version]. *Experimental and Clinical Psychopharmacology, 8,* 192-197.

Vita

Robyn Michelle Honig, daughter of Ellen Honig, was born in Houston, Texas, where she worked as a freelance writer and graduated magna cum laude from Cypress Falls High School in 1998. Shortly after graduation, she left Houston for the University of Texas at Austin and in May, 2000, graduated Phi Beta Kappa, earning a Bachelor of Arts with Highest Honors, Honors in Psychology, and a minor in Journalism. She then continued her studies with her mentor, Philip Gough, and in December, 2002, graduated yet again from the University with a Master of Arts. The following summer, she served as an Assistant Instructor in the Psychology department.

Permanent Address: 15934 Pinyon Creek Drive, Houston, Texas 77095

This dissertation was typed by the author.