Published in final edited form as:

Read Res Q. 2016 ; 51(4): 391–402. doi:10.1002/rrq.148.

Word-Decoding Skill Interacts With Working Memory Capacity to Influence Inference Generation During Reading

Stephen Hamilton, Erin Freed, and Debra L. Long

University of California at Davis, USA

Abstract

The aim of this study was to examine predictions derived from a proposal about the relation between word-decoding skill and working memory capacity, called verbal efficiency theory. The theory states that poor word representations and slow decoding processes consume resources in working memory that would otherwise be used to execute high-level comprehension processes, such as the generation of inferences. Previous research has yielded inconsistent findings about the importance of word decoding in adult readers, and the hypothesis has never been tested experimentally. Verbal efficiency theory was tested in this experiment by manipulating the difficulty of grapheme–phoneme conversion and assessing the extent to which readers made bridging inferences. Participants read two-sentence passages and then responded to lexical decision targets. Some of the passages required a bridging inference to integrate the first and second sentences. Decoding difficulty was manipulated such that the second sentence in some passages was written using pseudohomophones. Participants also received tasks to assess their working memory capacity and decoding ability. Inference priming was found in both the Standard American English and pseudohomophone contexts but was stronger in the former than in the latter. The advantage in priming for the Standard English relative to the pseudohomophone condition was predicted by an interaction between decoding skill and working memory capacity. Poor decoders who scored high on the span tests were less impaired by the pseudohomophone manipulation than were poor decoders who scored low on the tests. The results suggest that working memory capacity compensates for poor decoding skills even among proficient adult readers.

Reading comprehension depends on our ability to transform complex symbols of a written orthography into coherent mental representations, or discourse models (Johnson-Laird, 1983; Kintsch, 1988). One operation that is central to the construction of a discourse model is the integration of ideas within and across sentences. As each incoming event or action in a text is encoded, readers monitor and update their mental representations to integrate new information with preceding information (Zwaan, Langston, & Graesser, 1995). When these events or actions are perceived as incoherent, domain knowledge is used to construct

DEBRA L. LONG (corresponding author) is a professor of psychology at the University of California, Davis, USA; dlong@ucdavis.edu.

STEPHEN HAMILTON was a postdoctoral researcher at the University of California, Davis, USA, when this research was conducted. He is now a senior user experience researcher at Sony Computer Entertainment America, San Mateo, California, USA; stephen.hamilton7@gmail.com.

ERIN FREED is a doctoral candidate and lecturer at the University of California, Davis, USA; emfreed@ucdavis.edu.

inferences that go beyond what is stated literally, establishing both local and global coherence (e.g., McNamara & Kintsch, 1996). Consider the following example: “John reached into his bag and pulled out the flask. The orange juice was delicious.” To establish coherence between these sentences, readers need to make at least three bridging inferences. They must infer that John opened the flask, that he drank from it, and that the flask contained orange juice.

Much of the early empirical work on inference generation during reading was concerned with identifying the types of inferences that are made routinely and the conditions under which they were generated. In most studies, participants read passages that encouraged the construction of a particular inference. The passages were often followed by single-word targets that were related to the inference that should be drawn (e.g., *drink* in the previous example). Responses to target words were usually compared in two conditions: when the target was semantically related to an inference that was supported by the preceding passage and when the target was unrelated to the passage. The underlying rationale of these studies is that responses to the target word will be faster when readers have generated the inference than when the inference is not made. Numerous studies have employed this logic and found facilitation for inference-related words in reading times (e.g., Keefe & McDaniel, 1993), recognition response times (e.g., McKoon & Ratcliff, 1986), lexicality judgments (e.g., Potts, Keenan, & Golding, 1988), naming speed (Gras, Tardieu, & Nicolas, 2012), and efficacy as retrieval cues (e.g., Anderson & Ortony, 1975).

Of interest in this study is research on the reader characteristics that predict inference generation during comprehension. Much of this research has focused on the role of background knowledge and high-level cognitive abilities, such as working memory (WM) capacity (Bohn-Gettler & Kendeou, 2014; Boudewyn, Long, & Swaab, 2012, 2013; Long, Prat, Johns, Morris, & Jonathan, 2008; Long, Wilson, Hurley, & Prat, 2006; Traxler, Caplan, Long, & Waters, 2011). For example, Long et al. (2008) examined the likelihood that inferred ideas were included in readers’ memory representations as a function of several individual-difference factors: word decoding, general reasoning, vocabulary, WM capacity, and background knowledge. All of the factors were significant predictors of inference generation with the exception of word decoding.

WM capacity was also an important predictor of inference generation in a study by Bohn-Gettler and Kendeou (2014). Interestingly, they found that its role depended on participants’ reading goals and the structure of the text. Low-capacity readers made more inferences when they comprehended compare/contrast and problem/response texts than when they comprehended chronological and descriptive texts. The effect of text structure, however, occurred only when low-capacity readers comprehended the texts with the goal of entertainment. No influence of text structure was found when low-capacity readers comprehended the texts with the goal of preparing for an exam.

Our goal in this study is to examine individual variation in inference generation with a particular focus on how WM capacity and word decoding influence the extent to which readers make bridging inferences. In beginning readers, comprehension is primarily affected by decoding skill. As young readers acquire a more complete knowledge of the alphabetic

writing system, letter patterns that are repeated across words and pronunciations become consolidated, and word reading becomes an automated process based more on recognition and analogy than on a detailed analysis of phonology (see Ehri, 1995).

The influence of word decoding on comprehension in proficient adult readers is more controversial. Comparisons across studies can be difficult because they vary significantly with respect to the measures that are used to assess word decoding, whether or not measures of other reader characteristics are included and the analytic methods are used. Nonetheless, most studies have found substantial correlations between various measures of word decoding and reading comprehension in proficient readers (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Cunningham, Stanovich, & Wilson, 1990; Dixon, LeFevre, & Twilley, 1988; Georgiou & Das, 2014; Hannon, 2012, 2014; Holmes, 2009; Landi, 2010; Macaruso & Shankweiler, 2010). Many of these studies have also found that the relation between decoding and comprehension disappears or is significantly diminished when other reader characteristics are included in the analyses (Baddeley et al., 1985; Cromley et al., 2010; Cunningham et al., 1990; Dixon et al., 1988; Landi, 2010; Macaruso & Shankweiler, 2010).

For example, Baddeley et al. (1985) found that lexical decision performance accounted for 26% of the variance in comprehension when it was entered first into a regression analysis, but dropped to 15% when it was entered after WM span. Similarly, Cunningham et al. (1990) and Braze, Tabor, Shankweiler, and Mencl (2007) found that nonword naming accounted for substantial variance in comprehension when considered alone or entered first in a regression analysis (17% and 24%, respectively), but it accounted for much less variance when other reader characteristics were included (7% and 6%, respectively). Bell and Perfetti (1994) reported that the relation between nonword naming and comprehension diminished when shared variance with other reader characteristics was considered. In addition, they found that the relation varied significantly with the type of text; nonword naming accounted for significant unique variance in comprehension of science texts (11%) but not history or fiction texts (0.3% and 3.4%, respectively).

In Dixon et al.'s (1988) study, lexical decision speed accounted for a significant 15% of variance in performance on the Nelson–Denny Comprehension Test but was not a significant predictor of comprehension when other reader characteristics were included. However, not all studies have found that the relation between word decoding and comprehension diminishes with the inclusion of other reader characteristics. Hannon (2012, 2014) reported a substantial influence of word decoding on comprehension in university students even when high-level comprehension measures (e.g., knowledge integration, WM capacity) were included.

Most of the research described previously was conducted to examine the unique role of word decoding in comprehension among adult readers, primarily using multiple-regression analyses. Little research has focused on how word decoding and other reader characteristics may affect comprehension in combination. An important exception is a study by Hannon (2012); she examined mediation as one possibility for the joint influence of word decoding and high-level skills on comprehension. She used structural equation modeling to examine

direct and indirect effects of word decoding on comprehension. She compared a model in which word decoding was allowed to affect comprehension directly and another in which WM capacity mediated the relation between decoding and comprehension (i.e., word decoding had a path to WM capacity, and WM capacity had a path to comprehension). She found no evidence for mediation. The best fitting model was the one in which decoding had a direct effect.

A second possibility for how word decoding might combine with other reader characteristics to influence comprehension is interaction; that is, the relation between word decoding and comprehension might be moderated by other reader characteristics. This possibility is distinct from the mediated relation that was examined by Hannon (2012). In mediation, one variable has its influence through its relation with another. For example, Hannon examined the possibility that word decoding might influence comprehension because it was resource consuming. If so, word decoding would influence available WM resources, which would, in turn, influence comprehension. The influence of a moderating variable, in contrast, can be seen when the influence of one variable is dependent on the level of another. The effect of word decoding on comprehension might be modulated by another reader characteristic. For example, word decoding might have an effect on comprehension only among a subset of readers, such as those who are also low in WM capacity.

WM capacity as a moderating variable was investigated recently by Hamilton, Freed, and Long (2013). They conducted a study to examine how word decoding interacts with WM capacity to influence reading in proficient adult readers. They used multilevel modeling to estimate reading times as a function of text properties and cognitive abilities. Five real-world texts were used, and each sentence was coded for the number of new noun concepts that it contained. The researchers analyzed reading times for a target sentence (sentence N) and the following sentence (sentence N + 1). Of particular interest were reading times for sentences N + 1. Variance in reading times for these sentences was predicted by the number of new noun concepts in the previous sentence. This spillover in reading times increased as a function of decoding latencies (measured with a nonword naming task) but was modulated by WM capacity. When sentences were challenging, that is, when they had many new noun concepts to be integrated, poor decoders were slower to integrate them into their sentence-level representations, which was reflected in slow reading times for the following sentence. WM capacity moderated the costs associated with new noun integration. Importantly, the integration effect was significantly correlated with accuracy on comprehension questions that assessed readers' ability to make inferences.

Hamilton et al. (2013) interpreted their results in light of Perfetti's lexical quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2002) and verbal efficiency theory (Perfetti, 1985). The lexical quality hypothesis emphasizes the role of lexical knowledge and processing in comprehension. According to the hypothesis, skilled readers have large vocabularies, and their representations of individual words are high in quality, consisting of detailed orthographic-to-phonological mappings and rich semantic meanings. A high-quality lexical representation has orthographic, phonological, and semantic specificity, allowing word meanings to be accessed rapidly and accurately in context. Verbal efficiency theory is related to the broader lexical quality hypothesis by describing one way in which poor lexical

abilities can impair comprehension. According to verbal efficiency theory, readers who are slow at mapping orthographic representations to phonological ones use WM resources for word identification that would normally be dedicated to comprehension processes, such as inference generation. In Hamilton et al.'s study, poor decoders had more difficulty than good decoders in integrating new concepts into their discourse representations, presumably because word decoding consumed resources that would otherwise be used for integration. Importantly, this difficulty was diminished in high-capacity readers presumably because they had more resources to devote to decoding and integration.

The results provided by Hamilton et al. (2013) are consistent with the claim that word decoding and WM capacity interact to influence high-level comprehension processes, particularly when sentence-processing demands are high, but the study had no explicit manipulation of sentence difficulty. The aim of the present study was to examine the interaction of word decoding and WM capacity on inference generation by experimentally manipulating the difficulty of grapheme–phoneme conversion processes. We examined this issue in proficient adult readers for two reasons. First, readers with poor decoding skills in this population are likely to decode words well enough to understand individual sentences in a difficult grapheme–phoneme condition even though they may be impaired in their ability to generate inferences based on the content. Second, the correlation between measures of WM capacity and reading comprehension in adults is well documented (for reviews, see Daneman & Merikle, 1996; Van Dyke, Johns, & Kukona, 2014). Indeed, theories about the role of WM capacity in reading comprehension were developed in this population.

Participants read brief passages, some of which required a bridging inference to integrate ideas across sentences. The final sentence of each passage was written in Standard American English (SAE) or in pseudowords. Lexical decision targets followed the passages and were inference related, unrelated, or nonwords. Participants also performed tasks to assess their WM capacity and word-decoding ability.

We predicted that decoding skill and WM capacity would interact to influence inference generation, primarily in the pseudohomophone condition. Specifically, priming for inference targets should be smaller in the pseudohomophone than in the SAE condition. Crucially, the difference between the two contexts should be related to WM capacity and decoding skill: Poor decoders should show a larger impairment of inference generation as a function of decoding difficulty than good decoders do. This reduction, however, should be modulated by WM capacity. That is, WM capacity should help compensate for poor decoding skill, allowing poor decoders to make the bridging inference even though their decoding skills are challenged in the pseudohomophone condition.

Method

Participants

Fifty-two monolingual English speakers were recruited from the University of California, Davis participant pool and the State of California's Yolo County community. Native language was determined by asking participants what language(s) they learned first, what language(s) they spoke before the age of 5, and what language(s) was taught in their first-

grade class. All participants answered “English” to these questions. Of the 52 participants, 20 were male, and 32 were female. Twenty-three identified as white, three as black/African American, six as Latino, 11 as South Asian, and nine as mixed race. Ages ranged from 18 to 40 years (mean [M] = 21.18, standard deviation [SD] = 3.51). Participants’ level of education varied: 14 (26.9%) had a high school diploma only, 26 (50.0%) were undergraduate college students, three (5.8%) had attained an associate’s degree, eight (15.4%) had attained a bachelor’s degree, and one (1.9%) had obtained a master’s degree. No participants had a diagnosed reading or learning disability. Participants were paid \$10 per hour for their time.

Materials

The materials were adapted from Virtue, van den Broek, and Linderholm’s (2006) study (experiment 1). They consisted of 48 four-sentence passages. Each passage was followed by a lexical decision target. Details about how the passages and the targets were constructed can be found in Virtue et al. In one condition, the passages were written such that the target was related to an inference that was needed to integrate the final sentence with the previous one. In the unrelated condition, the passages were written such that the lexical decision target was unrelated to the content of the passage. Thus, the target words were the same in the two conditions; their status as inference related or unrelated depended on the content of the passages. Target words were one- or two-syllable action verbs and were similar in number of letters and frequency across texts (see Virtue et al., 2006). See Table 1 for example passages and targets.

We added an additional condition to Virtue et al.’s (2006) materials to manipulate decoding difficulty. This resulted in a total of three conditions. In the SAE condition, the final sentence of each passage was written in SAE, and passages were followed by inference targets. In the pseudohomophone condition, content words in the final sentences were written in pseudohomophones, and the passages were followed by inference targets. In the unrelated condition, the final sentences were written in SAE, and the targets were unrelated to the passages. This resulted in 144 experimental passages. Filler passages were also written to be similar to the experimental passages. All filler passages were followed by nonword targets.

Procedure

Participants were seated in front of a computer screen and told that they would read stories, some of which would contain made-up words. These were not real words but would sound like real words if they were pronounced aloud. An example sentence and translation was provided. Participants received three practice trials, each consisting of a four-sentence passage. They pressed the space bar to proceed through the passage one sentence at a time. Yes and no responses were mapped to the *D* and *K* keys, respectively, with small, plastic *Y* and *N* letters. All practice passages ended with a sentence written in pseudohomophones. After participants had read each practice passage, a fixation cross was presented for 750 ms and was followed by a target item in the center of the screen. They were asked to indicate whether the target was a real word or a nonword as quickly and accurately as possible. After this, a comprehension question about the passage was presented, to which participants gave

a yes or no answer. The questions required participants to endorse or reject literal propositions that may or may not have been expressed in the passage (e.g., “Did Kathy want to move to Vermont?”).

Once they had completed the practice section, participants were asked whether they had any questions before the experiment began. Each passage in the experiment was preceded by a screen that asked participants to press the space bar to begin the passage. They proceeded through the passage one sentence at a time by pressing the space bar. After they finished reading the final sentence, a fixation cross appeared in the center of the screen for 750 ms and was replaced with a lexical decision target. As before, participants were asked to indicate whether the target was a real word or a nonword as quickly as possible. To ensure that participants attended to the passages, 27 passages were followed by yes/no comprehension questions.

WM and Decoding Tasks

Participants received two tests of WM capacity: the alphabet span and operation span tasks. The stimuli for both tasks were presented visually. In the alphabet span task, participants received word lists to remember and were asked to recall them in alphabetical order. Each list was preceded by a fixation cross for 1,000 ms, and each item was presented for 1,000 ms. Once the list had been presented, a response screen was provided for participants to type the items in alphabetical order, pressing Enter to submit their answers. Participants received three practice lists, each of which consisted of three items. The test consisted of 25 trials: five trials each of two, three, four, five, and six items. Set size was presented at random. Scores were the total number of recalled items in alphabetical order (maximum = 100). Participants were not penalized for incorrect spellings. The task has acceptable internal consistency (Cronbach’s $\alpha = .73$).

In the operation span task (Unsworth, Heitz, Schrock, & Engle, 2005), participants were asked to remember lists of letters while they solved simple arithmetic problems. The task began with a practice section that determined a time limit for the remainder of the task. Time limits for arithmetic operation judgments were set at 2.5 *SD* above each participant’s mean arithmetic operation time. After the practice trials, participants received three sets of each set size (sizes three to seven), totaling 75 to-be-remembered letters and arithmetic problems. Span was the sum of all perfectly recalled sets in correct serial order when arithmetic judgment accuracy was greater than 85%. The task has good internal consistency (Cronbach’s $\alpha = .78$) and test–retest reliability (Cronbach’s $\alpha = .78$) (Unsworth et al., 2005).

Decoding ability was assessed with two tasks: nonword naming and phonological decision. Again, stimuli for both tasks were presented visually. In the nonword naming task, participants were required to pronounce nonword letter strings, such as *plimmed*, as quickly as possible. One hundred nonwords ranging from one to three syllables (four to nine letters) were presented, and a voice key was used to record onset latencies, with longer latencies indicating poorer decoding skill. The internal consistency of this measure was excellent (Cronbach’s $\alpha = .96$). Accuracy in the task was high ($M = 90.38$, $SD = 7.49$, range = 69–100). Mean reaction times were based on correct responses only. For each participant,

reaction times that were 2.5 *SD* above his or her mean reaction time were considered outliers and removed from the data set.

In the phonological decision task, two nonword letter strings were presented to participants on either side of a fixation point. Both nonwords were pronounceable, and participants were asked to indicate which would be an English word if it were pronounced aloud. For example, if the options were *brone* and *brane*, the correct option would be *brane*. The position of the correct option was counterbalanced across trials such that the correct choice appeared equally often on the left and right side of fixation. After four practice trials, each participant completed 75 experimental trials. The dependent measure in this task was the mean response time on correct trials only. The internal consistency of the measure was excellent (Cronbach's $\alpha = .95$). Accuracy in the task was high ($M = 69.90$, $SD = 5.79$, range = 47–75). For each participant, reaction times that were 2.5 *SD* above the mean reaction time were considered outliers and removed from the data set.

Results

The analyses that follow are restricted to reaction times from correct trials. Responses to 55 incorrect trials were removed (0.7% of the responses), and reaction times were winsorized according to the following criteria. For each participant, mean reaction times were calculated for each trial type. Reaction times that were 2.5 *SDs* above these means were replaced with those values; 1.9% of the data were replaced in this manner. Alphabet span and operation span were transformed into proportion scores, and a composite was created, consisting of the mean of the two scores. The composite score was used following recommendations by Waters and Caplan (2003), who found that individual measures of WM span are somewhat unreliable and that composite scores are a conservative means of assessing capacity. Participants score high or low on the composite measure only when their performance is either high or low on both of the individual measures.

Similarly, the average of participants' mean phonological decision and nonword naming reaction times was used as our measure of word decoding. Skewness and kurtosis for WM capacity were both normal: Skewness = 0.24 (standard error [*SE*] = 0.33), and kurtosis = 0.597 (*SE* = 0.65). For decoding, kurtosis was normal (0.761, *SE* = 0.65), but skewness was slightly inflated (0.980, *SE* = 0.33), as is common in reaction time tasks. Follow-up Kolmogorov–Smirnov tests indicated that the distributions of both of the composite variables were not significantly different from normal: WM capacity $D(52) = .065$, $p = .200$; decoding $D(52) = .110$, $p = .162$. Descriptive statistics and bivariate correlations are provided in Tables 2 and 3.

Our first analysis was conducted to determine whether our decoding manipulation influenced inference priming. A repeated-measures analysis of variance indicated that passage condition differentially affected target reaction times in both subject analyses, $F(2, 102) = 37.68$, $p < .001$; and item analyses, $F(2, 94) = 31.11$, $p < .001$. Post hoc tests using the Bonferroni adjustment revealed that responses to targets were significantly faster in the SAE condition than in the unrelated condition (668.51, $SD = 147.30$ vs. 751.58, $SD = 166.31$, $p < .001$), that responses to targets were significantly faster in the pseudohomophone

condition than in the unrelated condition (704.76, $SD = 158.65$ vs. 751.58, $SD = 166.31$, $p < .001$), and that responses to targets were significantly faster in the SAE condition than in the pseudohomophone condition (668.51, $SD = 147.31$ vs. 704.76, $SD = 158.65$, $p < .001$). These effects are depicted in Figure 1.

As can be seen in Figure 1, participants appeared to make the bridging inference in both the SAE and the pseudohomophone conditions, but inference priming was larger when sentences were written in SAE. The priming difference reflects the degree of difficulty that was introduced by the decoding manipulation. To examine individual differences in inference priming, an SAE facilitation effect was calculated by subtracting reaction times to targets in the SAE condition from reaction times to targets in the pseudohomophone condition. The larger the value, the greater the facilitation in the SAE condition relative to the pseudohomophone condition. It is important to note that this measure does not reflect the amount of inference priming between related and unrelated conditions; rather, it reflects the extent to which responses in the inference condition only were affected by the decoding manipulation. In other words, readers who were relatively unaffected by the decoding manipulation should show small differences between the SAE and pseudohomophone conditions. In contrast, readers whose responses were slowed by the decoding manipulation should show a large difference between the two conditions. The facilitation values were then regressed onto WM span, decoding, and a WM span $\times$ decoding interaction term. The model accounted for a significant proportion of the variance in SAE facilitation, $F(3, 48) = 3.29$, $p = .03$; and the interaction was statistically significant ($\beta = -0.43$, $t = -2.67$, $p = .01$). The full model is presented in Table 4.¹

Although WM span and decoding were analyzed as continuous variables, we divided participants based on a tripartite split of decoding skill to illustrate the interaction. Figure 2 depicts SAE facilitation as a function of WM capacity for fast, medium, and poor decoders. As can be seen in Figure 2, WM span had little effect on SAE facilitation among fast and medium decoders. The pattern was different for the slow decoders. Slow decoders who scored low on the span tests were faster to respond in the SAE condition than in the pseudohomophone condition, suggesting that the combination of poor decoding ability and low span made these readers more susceptible to difficulties arising from the decoding manipulation. This difficulty was ameliorated for slow decoders who scored high on the span tests.²

The relation between SAE facilitation and inference priming can be seen in Figure 3.³ The figure depicts response times to targets in the SAE, pseudohomophone, and unrelated

¹Given the wide age range among our participants, we conducted an analysis with age as a predictor variable. The pattern of results was identical to that reported previously.

²We conducted our analyses using composites to create measures of WM capacity and word decoding because multiple tasks tend to yield more reliable assessments of constructs than do individual ones (Waters & Caplan, 2003). In order to determine whether our findings were specific to particular measures of WM capacity and word decoding, we also analyzed the data using all combinations of the individual tasks. We found the same results in all analyses with one exception. In the analysis that included phonological decision and alphabet span, the interaction term was significant, but the overall regression model was only marginally so. The strongest effects were found in the analysis that included nonword naming and operation span. The effects in that analysis were stronger than the effects that we found using the composite scores.

³We also conducted an analysis of variance to examine the interaction among passage condition, decoding skill, and WM capacity. We created groups based on median splits of word decoding and WM capacity. The three-way interaction was significant, and follow-up interactions revealed that the only group to be affected by passage condition was the poor-word decoding/low-WM capacity group.

conditions. Fast decoders (panel A) and medium decoders (panel B) showed faster responses in the SAE and pseudohomophone conditions relative to the unrelated condition. The amount of priming, however, was not influenced by WM capacity. In contrast, slow decoders (panel C) showed inference priming that depended on the decoding manipulation and WM capacity. They showed inference priming when passages were written in SAE, but inference priming in the pseudohomophone condition depended on WM capacity. Higher scores on the span tasks were associated with more inference priming.

Discussion

The aim of this study was to determine the extent to which word decoding and WM capacity interact to influence inference generation when sentence-processing demands are high. We placed a targeted load on decoding processes by writing inference-eliciting sentences in pseudohomophones. We found inference priming in both the SAE and pseudohomophone conditions; readers responded faster to targets when they were related to inferences that were necessary to integrate the final sentence with the preceding context than when they were unrelated. Critically, inference priming was stronger in the SAE condition than in the pseudohomophone condition. This suggests that more processing was required to access lexical representations from the pseudohomophones than from the SAE forms. This difficulty reduced facilitation for inference targets relative to unrelated ones. We also found that SAE facilitation was affected by an interaction between word-decoding skill and WM capacity. The advantage of SAE contexts over pseudohomophone contexts was reduced as a function of capacity; this was especially pronounced for readers with slow responses on the decoding tasks.

Our findings can be interpreted in the context of verbal efficiency theory. When readers devote substantial resources to word decoding, fewer resources are available to execute high-level comprehension processes, such as inference generation. In our study, inference priming was affected by the interaction of word-decoding skill and WM capacity. Inference priming was reduced in the pseudohomophone condition but only among the subset of readers who were slow decoders and had low WM capacity. Thus, when lexical access was impeded, WM capacity provided readers with the cognitive resources that were necessary to generate bridging inferences (Perfetti, 1985). Our observation that the relation between word decoding and WM capacity was found only when the demands on word decoding were high (i.e., the pseudohomophone condition) is consistent with Bell and Perfetti's (1994) claim that effects of word decoding on comprehension in proficient adult readers may be seen primarily when texts contain numerous low-frequency words, as occurs in technical or scientific writing.

Our results are less consistent with the simple view of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990). According to this view, comprehension is the product of word-decoding skill and language comprehension abilities (e.g., vocabulary, domain knowledge, higher order thinking skills). Our finding that poor decoders/low span readers showed less inference priming than good decoders/high span readers is consistent with this view. However, the simple view predicts that poor decoders/high span readers should also show less inference priming than good decoders/high span readers. This is because good decoding

is necessary for good reading comprehension; the product of poor decoding and high WM capacity yields a reading comprehension score that is lower than the product of good decoding and high WM capacity. Thus, our results are more consistent with verbal efficiency theory than the simple view of reading partly because the simple view is more specific about exactly how word decoding and other reader characteristics combine to influence comprehension. It is important to note, however, that the simple view of reading describes how word decoding and language comprehension influence performance on overall reading comprehension. Our dependent measure was inference priming and not comprehension performance, per se. Further studies will be necessary to determine whether our results are specific to inference priming or generalize to other comprehension skills and measures.

Although verbal efficiency theory can accommodate the pattern of results that we found, so can other theories with additional assumptions. Several theories have been proposed to account for the relation between WM span and comprehension. These theories generally fall into two classes. The first class consists of those in which the relation between span and comprehension is a direct one. That is, limitations in the ability to simultaneously maintain and process information affect the types of relations that readers are able to construct during comprehension. Two such theories are Just and Carpenter's (1992) capacity theory and Caplan and Waters's (1999) separate sentence interpretation resource theory. Although the two theories differ in significant ways, both conceptualize WM as consisting of a finite pool of cognitive resources and that individuals vary in the resources they have available to process language. Our results can be accommodated by these theories with the assumption that word-decoding processes consume cognitive resources. When word decoding is challenging, more resources will be directed to word identification, leaving fewer resources for other comprehension processes. When the pool of available resources is small, the difficulties in decoding will consume those resources that are needed for high-level comprehension processes, such as inference generation.

The second class of theories emphasizes the role of experience in accounting for the relation between WM span and comprehension. MacDonald and Christiansen (2002) proposed a connectionist-based account of the relation. In this account, the capacity of a system arises from its architecture (e.g., the number of processing units, how activation passes through the weights) and the system's experience (e.g., how often it has processed similar input in the past). Thus, capacity is not a separate pool of resources; it is a property of the processing system itself. The relation between span tasks and comprehension arises from variation in two factors. First, individuals vary with respect to basic sensory/perceptual abilities, primarily the ability to represent and process phonological information. Second, individuals vary in reading experience. This variation gives rise to individual differences in practice with linguistic stimuli.

The role of experience is also prominent in a theory proposed by Ericsson and Kintsch (1995), called long-term working memory (LTWM). According to their theory, WM consists of the standard limited-capacity system that they call short-term working memory and one based on skilled storage and retrieval in long-term memory that they call LTWM. The amount of information that can be actively maintained in LTWM is not limited by a fixed

capacity. As individuals become skilled at a task, they develop mechanisms for encoding and retrieving information from long-term memory that meet the demands of the task.

Our results can be accommodated by the connectionist-based and LTWM accounts with the assumption that poor decoders read less frequently than do good decoders. In the connectionist-based account, this would mean that the structure of their network and the efficiency with which it operates would be affected by their limited linguistic experience. In the LTWM model, poor decoders would be poorer than good decoders at encoding and retrieving information from long-term memory, effectively reducing the functional capacity of LTWM.

Limitations and Conclusions

Our goal in this study was to examine the possibility that word decoding and WM capacity interact to influence inference generation, a possibility that is consistent with verbal efficiency theory. One limitation of our study is that we assessed only two of the many reader characteristics that are known to influence reading comprehension. Most reader characteristics are correlated; thus, the effects that we reported here may be secondary to the influence of variables that we did not include. For example, we chose to assess word-decoding skill as a measure of lexical processing because our manipulation involved the difficulty of grapheme–phoneme conversion processes. If we had included measures of vocabulary, we may have found that the interaction between decoding and WM capacity was secondary to an interaction between vocabulary and WM capacity. This would be consistent with the broader lexical quality hypothesis, which emphasizes the depth and breadth of readers' knowledge about words in comprehension.

A second limitation of our study is that performance on the nonword naming and phonological decision tasks that we used are affected by several lexical abilities: word decoding, word fluency, and phonological awareness. We used the label *word decoding* because it has been used to describe these measures in previous research on reading in adult populations. However, more research using a broader range of lexical ability measures will be necessary to determine which of these abilities is central in the interaction between lexical ability and WM capacity.

Finally, we had only two measures of our word-decoding and WM capacity constructs in this study, and we assessed performance in a relatively small sample of participants. A larger set of measures and participants would have allowed us to conduct more sophisticated analyses, such as structural equation modeling. One advantage of our sample, however, is that it included a broader range of adult readers than typically occurs in university studies. Our sample contained only proficient readers, but it was more varied in age and educational level than is typical.

The results of our study are consistent with those of Hamilton et al.'s (2013) study in suggesting that the influence of word decoding in proficient readers may be seen primarily in its interaction with WM capacity. Readers with poor decoding skill had difficulty integrating concepts and ideas in a text when the text was challenging and the readers had

small WM capacities. We found that WM capacity can compensate for poor decoding skill by providing extra resources than can be devoted to decoding when readers encounter new or unfamiliar words, allowing readers to engage in higher order reading processes, such as inference generation.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (5R01HD073948-02).

References

- Anderson RC, Ortony A. On putting apples into bottles: A problem of polysemy. *Cognitive Psychology*. 1975; 7(2):167–180.
- Baddeley A, Logie R, Nimmo-Smith I, Brereton N. Components of fluent reading. *Journal of Memory and Language*. 1985; 24(1):119–131.
- Bell LC, Perfetti CA. Reading skill: Some adult comparisons. *Journal of Educational Psychology*. 1994; 86(2):244–255.
- Bohn-Gettler CM, Kendeou P. The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*. 2014; 39(3):206–219. [PubMed: 25018581]
- Boudewyn MA, Long DL, Swaab TY. Cognitive control influences the use of meaning relations during spoken sentence comprehension. *Neuropsychologia*. 2012; 50(11):2659–2668. [PubMed: 22842106]
- Boudewyn MA, Long DL, Swaab TY. Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology*. 2013; 4:1–16. [PubMed: 23382719]
- Braze D, Tabor W, Shankweiler DP, Mencl WE. Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities*. 2007; 40(3):226–243. [PubMed: 17518215]
- Caplan D, Waters GS. Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*. 1999; 22(1):77–94. [PubMed: 11301522]
- Cromley JG, Snyder-Hogan LE, Luciw-Dubas UA. Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*. 2010; 102(3):687–700.
- Cunningham, AE.; Stanovich, KE.; Wilson, MR. Cognitive variation in adult college students differing in reading ability. In: Carr, TH.; Levy, BA., editors. *Reading and its development: Component skills approaches*. San Diego, CA: Academic; 1990. p. 129-159.
- Daneman M, Merikle PM. Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*. 1996; 3(4):422–433. [PubMed: 24213976]
- Dixon P, LeFevre J, Twilley LC. Work knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*. 1988; 80(4):465–472.
- Ehri LC. Phases of development in learning to read words by sight. *Journal of Research in Reading*. 1995; 18(2):116–125.
- Ericsson KA, Kintsch W. Long-term working memory. *Psychological Review*. 1995; 102(2):211–245. [PubMed: 7740089]
- Georgiou GK, Das JP. Reading comprehension in university students: Relevance of PASS theory of intelligence. *Journal of Research in Reading*. 2014; 37(S1):S101–S115.
- Gough P, Tunmer W. Decoding, reading, and reading disability. *Remedial and Special Education*. 1986; 7(1):6–10.
- Gras D, Tardieu H, Nicolas S. Predictive inference activation. *Swiss Journal of Psychology*. 2012; 71(3):141–148.
- Hamilton ST, Freed EM, Long DL. Modeling reader and text interactions during narrative comprehension: A test of the lexical quality hypothesis. *Discourse Processes*. 2013; 50(2):139–163. [PubMed: 23526862]

- Hannon B. Understanding the relative contributions of lower-level word processes, higher-level processes, and working memory to reading comprehension performance in proficient adult readers. *Reading Research Quarterly*. 2012; 47(2):125–152.
- Hannon B. Are there gender differences in the cognitive components of adult reading comprehension? *Learning and Individual Differences*. 2014; 32:69–79.
- Holmes VM. Bottom-up processing and reading comprehension in experienced adult readers. *Journal of Research in Reading*. 2009; 32(3):309–326.
- Hoover W, Gough P. The simple view of reading. *Reading and Writing*. 1990; 2(2):127–160.
- Johnson-Laird, PN. *Mental models*. Cambridge, MA: Harvard University Press; 1983.
- Just MA, Carpenter PA. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*. 1992; 99(1):122–149. [PubMed: 1546114]
- Keefe DE, McDaniel MA. The time course and durability of predictive inferences. *Journal of Memory and Language*. 1993; 32(4):446–463.
- Kintsch W. The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review*. 1988; 95(2):163–182. [PubMed: 3375398]
- Landi N. An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing*. 2010; 23(6):701–717. [PubMed: 21691452]
- Long DL, Prat CS, Johns C, Morris P, Jonathan E. The importance of knowledge in vivid text memory: An individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review*. 2008; 15(3):604–609. [PubMed: 18567262]
- Long DL, Wilson J, Hurley R, Prat CS. Assessing text representations with recognition: The interaction of domain knowledge and text coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006; 32(4):816–827.
- Macaruso P, Shankweiler D. Expanding the simple view of reading in accounting for reading skills in community college students. *Reading Psychology*. 2010; 31(5):454–471.
- MacDonald MC, Christiansen MH. Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*. 2002; 109(1):35–54. [PubMed: 11863041]
- McKoon G, Ratcliff R. Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1986; 12(1):82–91.
- McNamara DS, Kintsch W. Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*. 1996; 22(3):247–288.
- Perfetti, CA. *Reading ability*. New York, NY: Oxford University Press; 1985.
- Perfetti C. Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*. 2007; 11(4):357–383.
- Perfetti, CA.; Hart, L. The lexical quality hypothesis. In: Verhoeven, C.; Elbro, C.; Reitsma, P., editors. *Precursors of functional literacy*. Philadelphia, PA: John Benjamins; 2002. p. 189–213.
- Potts GR, Keenan JM, Golding JM. Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*. 1988; 27(4):399–415.
- Traxler, MJ.; Caplan, D.; Long, DL.; Waters, GS. Working memory in sentence comprehension and production. In: Levin, ES., editor. *Working memory: Capacity, developments and improvement techniques*. Hauppauge, NY: Nova Science; 2011. p. 225–248.
- Unsworth N, Heitz RP, Schrock JC, Engle RW. An automated version of the operation span task. *Behavior Research Methods*. 2005; 37(3):498–505. [PubMed: 16405146]
- Van Dyke JA, Johns CL, Kukona A. Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition*. 2014; 131(3):373–403. [PubMed: 24657820]
- Virtue S, van den Broek P, Linderholm T. Hemispheric processing of inferences: The effects of textual constraint and working memory capacity. *Memory & Cognition*. 2006; 34(6):1341–1354. [PubMed: 17225513]
- Waters GS, Caplan DC. The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*. 2003; 35(4):550–564.

Zwaan RA, Langston MC, Graesser AC. The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*. 1995; 6(5):292–297.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

FIGURE 1.
Mean Reaction Times (RTs) to Targets in the Standard American English (SAE),
Pseudohomophone (Pseudo), and Unrelated Conditions

FIGURE 2.

Standard American English (SAE) Facilitation as a Function of Working Memory Capacity and Word Decoding

Note. Pseudo = pseudohomophone; RT = reaction time.

FIGURE 3.

(a) Fast Decoders, (b) Medium Decoders, and (c) Slow Decoders

Note. Pseudo = pseudohomophone; SAE = Standard American English.

TABLE 1

Sample Passages With Inference and Unrelated Targets

Condition	Sample passage	Inference target
Standard American English	"Greg decided to try a game of horseshoes in the park today. He had seen other people play horseshoes, but had never actually tried it himself. After arriving at the park, Greg picked up his horseshoe. Greg was amazed that his horseshoe landed right in front of the metal post."	<i>throw</i>
Pseudohomophone	"Greg decided to try a game of horseshoes in the park today. He had seen other people play horseshoes, but had never actually tried it himself. After arriving at the park, Greg picked up his horseshoe. Greg was amaized that his hoarshoo landidd rye in frunt of the mettul poste."	<i>throw</i>
Unrelated	"The movie theatre had been playing the same movie for weeks. Everyone who worked there was sick of it and hated the people coming back for repeat viewings. They decided to play a prank. They locked the doors after they were all in, and played white noise for two hours."	<i>throw</i>

TABLE 2

Descriptive Statistics of the Study

Variable	Range	Mean (standard deviation)
Alphabet span (accuracy)	6–90	44 (14)
Operation span (accuracy)	7–75	48 (15)
Nonword naming (ms)	2,139–3,071	2,515 (232)
Phonological decision (ms)	1,333–4,725	2,592 (751)

TABLE 3

Bivariate Correlations Between the Variables in the Study

Variable	1	2	3	4	5	6	7	8
1. Standard American English target reaction time	—							
2. Pseudohomophone target reaction time	.94***	—						
3. Unrelated target reaction time	.90***	.89***	—					
4. Standard American English facilitation	.03	.37***	.16	—				
5. Alphabet span	.07	.05	.14	-.03	—			
6. Operation span	.14	.07	.12	-.18	.32*	—		
7. Nonword naming	.44***	.39***	.42***	-.07	-.16	-.10	—	
8. Phonological decision	.20	.15	.23	-.12	.17	-.04	.47	—

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

TABLE 4

Regression Model

Predictor variable	<i>B</i>	Standard error <i>B</i>	β	<i>t</i>
Constant	31.68	7.34	—	4.32 ***
Working memory span	−165.75	59.71	−0.44	−2.78 **
Decoding	−0.05	−0.03	−0.21	−1.53
Working memory span × Decoding	−0.73	0.27	−0.43	−2.67 **

Note. $R^2 = .17$.

**
 $p < .01$.

 $p < .001$.