Published in final edited form as:
Cognition. 2014 June ; 131(3): 373–403. doi:10.1016/j.cognition.2014.01.007.

Low working memory capacity is only spuriously related to poor reading comprehension

Julie A. Van Dyke[1], **Clinton L. Johns**[1], and **Anuenue Kukona**[2]

[1]Haskins Laboratories, University of Dundee, UK

[2]School of Psychology, University of Dundee, UK

Abstract

Accounts of comprehension failure, whether in the case of readers with poor skill or when syntactic complexity is high, have overwhelmingly implicated working memory capacity as the key causal factor. However, extant research suggests that this position is not well supported by evidence on the span of active memory during online sentence processing, nor is it well motivated by models that make explicit claims about the memory mechanisms that support language processing. The current study suggests that sensitivity to interference from similar items in memory may provide a better explanation of comprehension failure. Through administration of a comprehensive skill battery, we found that the previously observed association of working memory with comprehension is likely due to the collinearity of working memory with many other reading-related skills, especially IQ. In analyses which removed variance shared with IQ, we found that receptive vocabulary knowledge was the only significant predictor of comprehension performance in our task out of a battery of 24 skill measures. In addition, receptive vocabulary and non-verbal memory for serial order—but not simple verbal memory or working memory—were the only predictors of reading times in the region where interference had its primary affect. We interpret these results in light of a model that emphasizes retrieval interference and the quality of lexical representations as key determinants of successful comprehension.

Keywords

Interference; Individual Differences; Comprehension; Working Memory; Vocabulary; Corsi Blocks

1. Introduction

The centrality of memory operations to language comprehension has long been recognized: it was 50 years ago, for example, that Miller and Chomsky (1963) proposed that there is an endogenous upper bound on the number of noun phrases that can be manipulated in memory during sentence processing. This theoretical perspective – that capacity constrains language comprehension – was reinforced by the subsequent development of Alan Baddeley's model of working memory (e.g., Baddeley & Hitch, 1974; Repovš & Baddeley, 2006), in which a

Please address correspondence to: Julie A. Van Dyke, Ph.D., jvandyke@haskins.yale.edu, Haskins Laboratories, 300 George Street, New Haven CT 06511, Phone: 203-865-6163.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

single, finite pool of processing resources supports both storage and computation. Given the pervasive influence of Baddeley's model, it is unsurprising that most theories of comprehension skill incorporate working memory capacity, often in a central way (e.g., Engle, Cantor, & Carullo, 1992; Gibson, 1998; Just & Carpenter, 1992; see Long, Johns, & Morris, 2006, for a review). According to these accounts, humans possess a limited supply of neural "resources" with which to support cognitive operations during sentence processing. As the computational demands of ongoing comprehension increase, the resources available to keep items active in working memory decrease; conversely, as memory demands increase, there will be fewer resources available for comprehension processes. Exceeding available resources results in either loss of information from working memory, impaired processing (e.g., syntactic parsing, semantic integration, etc.), or both. The classic demonstration of this is the contrast between subject- and object- extracted relative clauses (RCs), in which the latter are more difficult to process than the former; the reason for this difficulty is thought to derive from the need to actively maintain the initial noun phrase (e.g., *The banker*) in object RCs while processing the embedded clause, after which it can be integrated with its verb phrase (e.g., *climbed*).

(1a) OBJECT RC: The banker that the barber praised climbed the mountain.

(1b) SUBJECT RC: The banker that praised the barber climbed the mountain.

On this account, individual differences in sentence comprehension arise because of intrinsic differences in the total capacity of the resource pool: individuals with smaller total capacity will show impaired comprehension relative to high capacity individuals, especially with complex sentences that require additional computations. Numerous studies have demonstrated the crucial interaction of memory capacity and sentence difficulty: when compared to their high capacity peers, low capacity participants appear to have greater difficulty not only with object RCs (compared to subject RCs), but also with a host of other complex constructions (e.g., Just & Carpenter, 1992; King & Just, 1991; Long & Prat, 2008; MacDonald, Just, & Carpenter, 1992; Nieuwland & Van Berkum, 2006; Traxler, Williams, Blozis, & Morris, 2005).

Despite the prevalence of the idea that a capacity-based memory architecture supports language processing, there is now a broad base of empirical evidence indicating that the amount of information that can be actively maintained in memory during sentence processing is very limited—even for skilled readers. Based on the premise that elements that are maintained in active memory should be accessed more quickly than those passively stored in LTM, a number of studies have utilized precise measures of retrieval speed to determine the size of available, active memory (see McElree, 2006, for a review). For example, in list-learning paradigms, the consistent result is that a speed advantage is only observed for the most recently studied item (McElree, 1996, 1998, 2001, 2006; McElree & Dosher, 1989, 1993; Öztekin & McElree, 2007; Wickelgren et al., 1980). Similarly, in studies of sentence processing, the consistent result is that only the most recently processed linguistic constituent exhibits increased accessibility (McElree, 2000; McElree, Foraker, & Dyer, 2003; Wagers & McElree, 2009). This presents a strong challenge to the capacity view, in which multiple propositions, syntactic structures, or entire interpretations are thought to enjoy increased accessibility by virtue of being actively maintained in working memory.

In addition, there are important theoretical reasons for believing that an emphasis on capacity does not optimally characterize the constraints that the memory system places on language comprehension. Capacity is thought to matter because information that is not maintained is lost—pushed out of active memory by the demands of other processing, and lost because the consequent inattention results in decay (Gibson, 1998; 2000; Just & Carpenter, 1992). However, this approach is problematic in light of extensive research in the memory domain suggesting that interference, and not decay, is the primary source of forgetting (e.g., Underwood & Keppel, 1962; Waugh & Norman, 1965; see Berman, Jonides, & Lewis, 2009, for a more recent assessment). Interference arises when retrieval cues are insufficient to uniquely identify a target item; in such cases, cues are said to be "overloaded," and distracting items, which share some features with the intended target, are erroneously retrieved instead (e.g., Nairne, 2002; Öztekin & McElree 2007; Watkins & Watkins 1975). Although interference effects were originally investigated in the memory domain, there is now a substantial body of evidence demonstrating interference effects in language comprehension (see Van Dyke & Johns, 2012, for a review). For example, in sentence processing, Van Dyke (2007) observed interference effects from a semantically similar distractor (e.g., *neighbor*) when the animate NP (*resident*) must be retrieved as the VP *complained* is parsed (e.g., 2b, as compared with 2a, where the potential distractor *warehouse* is not animate). This occurs despite the presence of syntactic cues that could eliminate the distractor as a potential subject of *complained*.

(2a) The resident who was living near the dangerous warehouse complained about the noise.

(2b) The resident who was living near the dangerous neighbor complained about the noise.

(2c) The resident who declared that the warehouse was dangerous complained about the noise.

(2d) The resident complained about the noise.

Distractors based on the match of syntactic cues alone also produce interference (Van Dyke & Lewis, 2003); thus, (2c) is also more difficult than (2a), because the intervening subject NP *warehouse* matches the syntactic retrieval cues from *complained*, which requires a subject NP to complete the long distance dependency. This finding contrasts sharply with the capacity-based view that complex sentences of this sort are difficult because they consume WM resources. That is, contra the capacity-based account, sentence (2c) is more difficult than (2a) despite having the same amount of intervening material (i.e., identical memory demands) between the dependent subject and verb (*resident-complained*). Further, sentence (2a), which contains neither a syntactic nor a semantic distractor for the subject of *complained*, was found to be no more difficult than sentence (2d), which contains no intervening material at all (Van Dyke & Lewis, 2003).

In addition to interference from semantic and syntactic cue overload, interference as a result of referential cues has also been observed. Gordon and colleagues (Gordon, Hendrick, & Johnson, 2001; 2004; Gordon, Hendrick, Johnson & Lee, 2006) found that sentences whose nouns were of the same referential type (e.g., both descriptive nouns, as in 3a, underlined) were more difficult than those with nouns of mixed type (as in 3b, which includes a proper noun, and 3c, which includes an indexical pronoun), despite having both identical syntactic structure and identical memory demands between the *banker*-*climbed* dependency.

(3a) The banker that the barber praised climbed the mountain.

(3b) The banker that John praised climbed the mountain.

(3c) The banker that you praised climbed the mountain.

The interference effect elicited by the similarity of NP types is highly robust, appearing as decreased accuracy on comprehension questions, slower self-paced reading times at both the main verb (e.g., *climbed* in 3a–c) and the immediately preceding word or region, and longer latencies on both early (gaze duration, right-bounded reading time) and late (rereading time) eye tracking measures in the same critical areas. Notably, this disadvantage is not predicted by capacity-based accounts, because the number of referents, number of propositions, number of syntactic relations, and all other possible units typically used to index memory load are constant across conditions.

Together, these findings suggest that emphasis should shift away from questions about the *quantity* of information that can be maintained in memory during comprehension, and refocused to investigate how the specific *content* of the information in memory affects retrievals that must occur when computing linguistic relationships. The current research follows a series of recent studies that used a dual-task paradigm to directly manipulate the contents of memory during sentence processing. Participants in these studies memorized a short list of words (usually three items) immediately prior to reading a sentence; after the sentence, they answered a comprehension question about sentence content, and then recalled the words from the memory list. This experimental paradigm is interesting not just because it affords control over the contents of memory, but also because it supports an examination of whether the mechanisms utilized for remembering a list of words are the same as those used for language processing. If language and memory processes draw on the same pool of resources, then interactions between either the size or the contents of memory and the sentence reading task are expected. If, on the other hand, language processes have access to a separate domain-specific memory resource (as proposed by Caplan and Waters, 1999), then no interaction between measures of reading behavior and the contents of memory are expected. A number of researchers have reported the predicted interaction (e.g., Fedorenko, Gibson, & Rohde, 2006; Gordon, Hendrick, & Levine, 2002; Van Dyke & McElree, 2006), lending support to the former position. For example, Gordon and colleagues (2002; see also Federenko et al., 2006) found that memorizing a short word list impaired processing of sentences containing object-relative clauses relative to those with subject-relative clauses; however, the effect depended on the *semantic* content of the list items. When the type of list item (e.g., names, Joel-Greg-Andy; or descriptions, poet-cartoonist-voter) differed from the type of NP in the subsequent sentence (e.g., names: *It was Tony that liked Joey before the* *argument began*; or descriptions: *It was the dancer that the fireman liked before the argument began*), accuracy improved. These results demonstrate that it is not simply the presence of the memory load that affects processing, but the specific content of the memory list vis-à-vis the sentence itself. However, these studies did not identify the locus of the interference effect, which could have resulted from either encoding or retrieval operations.

Using a slightly different dual-task paradigm, Van Dyke & McElree (2006) demonstrated that the influence of the memory load was specific to the retrieval operation that was required to resolve the linguistic dependency in the reading task. They asked participants to memorize a word list (e.g., table-sink-truck) prior to reading sentences such as (4a) and (4b).

(4a) It was the boat that the guy who lived by the sea sailed in two sunny days.

(4b) It was the boat that the guy who lived by the sea fixed in two sunny days.

The "Memory Load" conditions were contrasted with "No Load" conditions in which participants read the sentences without first memorizing a word list. The critical manipulation was the relation between the matrix verb in the sentence (e.g., *fixed* or *sailed*) and the memory list items. Interference was expected when features of the list items matched the semantic demands of the verb looking to retrieve its direct object. As predicted, longer reading times were observed at the matrix verb when memory items could serve as objects of the verb relative to when they could not: that is, (4b), in which *table, sink, truck*, and *boat* are all potentially "fixable" objects is more difficult than (4a), in which the only "sail-able" object is *boat*. Moreover, this reading time difference was not present in either "No Load" condition (identical to 4a and 4b, but without the word list), demonstrating that the difference could be solely attributed to the presence of the memory words, and their match to the retrieval cues of the sentences' main verbs.

2. Overview of current study

The present study seeks to replicate and extend this research by examining individual differences in susceptibility to retrieval interference in a traditionally understudied population. Although previous studies of comprehension difficulty have mainly utilized the college "subject-pool" population, we recruited a community-based sample of non-college-bound adolescents (ages 16–24), a population that the National Center for Education Statistics in the US estimated as including approximately 39% of high school seniors in 2004, the most recently studied cohort (Ingels et al., 2008). Based on previous experience with this population, we expected large skill differences (Braze, Tabor, Shankweiler, & Mencl, 2007; Braze, Mencl, Tabor, Pugh, Constable, Fulbright, Magnuson, Van Dyke, & Shankweiler, 2011; Magnuson, Kukona, Braze, Johns, Van Dyke, Tabor, Mencl, Pugh, & Shankweiler, 2011; Shankweiler, Mencl, Braze, Tabor, Pugh, & Fulbright, 2008; Van Dyke & Kuperman, 2011). For example, mean reading skill for the sample in Kuperman and Van Dyke (2011) was at the 10th grade level, and ranged from 4th grade to college level (SD = 3.1). The fact that the current sample is age-matched to the college subject-pool population permits comparisons with our previously published work demonstrating retrieval interference. Extending our research to the community-based sample is important because it affords an investigation of comprehension difficulty in a sample that is more representative of the population at large.

In order to thoroughly characterize the individual cognitive abilities of our sample, we administered an extensive battery of cognitive measures, including working memory capacity. This, combined with the dual-task paradigm used in Van Dyke & McElree (2006), allowed us to contrast the predictions of both capacity-based and interference-based accounts of individual differences in sentence comprehension. Because only the main verb differs across conditions, the sentences contain the same grammatical structure, thus controlling the computational demands of sentence parsing. Because the size of the memory load is constant within condition (either three items or zero items), the amount of information to be maintained is also constant and controlled. Accordingly, capacity-based accounts suggest that we should observe a main effect of the memory load, such that reading times will be slowed at the main verb because limited resources are being diverted to support active maintenance of the memory items. Further, this effect should vary according to differences in individual working memory capacity, with greater performance decrements in lower capacity readers. Critically, capacity-based accounts do not predict individual processing or comprehension differences on the basis of the match between the content of the memory lists and the semantics of the sentential verbs, because this factor affects neither the computational nor the storage/maintenance demands of the stimuli.

In contrast, a retrieval interference account makes very different predictions. As in Van Dyke and McElree (2006), we expect an interaction such that the memory load will impair processing only when the memory load items are semantic distractors for the object of the sentential verbs' retrieval cues (e.g., when the verb is *fixed*, as in 4b). Thus, performance is not expected to vary as a function of working memory capacity *per se*. Rather, it is the *contents* of memory, and their relationship to the retrieval cues that will determine participants' ability to interpret the sentence. The presence of similar items creates retrieval interference because the shared features reduce the distinctiveness of the target (perhaps via a process of feature overwriting, as proposed by Nairne, 1990; and Oberauer & Kliegl, 2006) so that fewer aspects of the target's feature structure can serve as unambiguous retrieval cues. Hence, the probability of retrieving a similar, but incorrect, item is increased. In addition, the efficiency of retrieval is further reduced when the quality of the target memory representation is reduced, making it more difficult to retrieve. A variety of factors may affect representation quality, including improper initial encoding, which may arise due to a range of linguistic or cognitive deficits, which may be present in low skill readers (i.e., poor phonological skills, word knowledge, or insufficient attention). Our comprehensive individual differences battery will enable us to identify which of these deficits is the greatest determinant of susceptibility to interference in poor readers. Thus, through emphasizing the crucial role of representation quality, which reflects the *contents* of memory rather than the *capacity* of memory, the retrieval interference approach affords an alternative means for understanding the mechanism through which poor comprehension arises.

3. Method

3.1 Participants

The participants were 65 young people (ages 16–24) who were paid $15/hour. We recruited participants from the local community in a number of ways, including presentations at adult education centers; advertisements in local newspapers; posters/flyers placed on adult school and community college campuses, public transportation hubs, local retail and laundry facilities; and from referrals from past and current study participants. All were native English speakers, and none had a diagnosed reading or learning disability. Testing took place in two sessions, each on a separate day. The first session lasted no longer than three hours, including several breaks, and the second (which included the experiment) lasted no longer than two hours, including breaks.

3.2 Materials

3.2.1 Sentences—Varying the factors of *Memory Load* and *Interference* (2 × 2 design) resulted in four experimental conditions. Materials were the 36 experimental object-cleft sentences used in Van Dyke and McElree (2006). There were two versions of each sentence, differing only by a single word (the main verb; see Table 1). Memory load was manipulated such that participants either did or did not have a list of three words to maintain in memory while they read the sentences. When the memory list was present, the verb manipulation created an interference condition, such that the words in the memory list either were or were not plausible direct objects for the manipulated verb. (For details about the norming procedure used to assess verb-object plausibility see Van Dyke & McElree, 2006.) All sentences were followed by a forced choice “yes/no” comprehension question about sentence content.

In addition, participants also saw 144 filler sentences, also from Van Dyke & McElree (2006). Of these, 36 were subject cleft sentences and 108 were non-cleft sentences with right-branching structure. Half of the filler items were accompanied by a three-word memory list. In contrast to the experimental items, no words in these lists were related to the main verbs in the filler sentences. All filler items were followed by a “yes/no” comprehension question, in which responses were evenly split between the two answers.

3.2.2 Tests of individual cognitive abilities—We administered a battery of 24 tests that measured print mapping, reading skill, oral language use, memory, and intelligence. Wherever possible, we chose standardized instruments that are well established through large-scale psychometric studies as having high construct validity and test-retest reliability, and are often widely used for clinical assessment and diagnosis. Standardized assessments for four skills (working memory, visual memory, print experience, and spelling) were not available; however, we employed instruments for assessing these that have been widely used in the literature, and characteristic citations for each of these are given below. Testing was distributed across both sessions. Battery data from 15 participants were incomplete: 14 were missing a single test, and one participant was missing two tests. We note these below where applicable. Data from participants with incomplete measures are included in all mixed-effects analyses when possible. Exclusions for specific analyses are noted below where they occurred. Tests included:

- *Comprehensive Test of Phonological Processing* core subtests (CTOPP; Wagner, Torgeson, & Rashotte, 1999), which provided composite measures of (1) phonological awareness (subtests: elision, blending words), (2) phonological memory (subtests: memory for digits, nonword repetition), and (3) rapid naming (subtests: rapid digit naming, rapid letter naming). We also administered the rapid color-naming subtest, which is not included in the rapid naming composite. Two participants were missing rapid naming tests due to experimenter error.
- *Weschler Abbreviated Scale of Intelligence* (WASI; Psychological Corp., 1999) was used to calculate full-scale IQ scores. These scores were derived from the vocabulary and matrix reasoning subtests.
- *Woodcock-Johnson-III Tests of Achievement*, reading and oral comprehension area subtests (WJ-III; Woodcock, McGrew, & Mather, 2001) including (1) word attack (reading a list of pseudowords aloud), (2) word identification (naming words from a list), (3) reading fluency (speed of reading sentences and answering yes/no questions about each), (4) passage comprehension (orally providing words missing from printed sentences/paragraphs), and (5) oral comprehension (orally providing words missing from auditory sentences). Two participants were missing scores from the passage comprehension subtest due to experimenter error.
- *Test of Word Reading Efficiency* (TOWRE; Torgeson, Wagner, & Rashotte, 1999), including the (1) sight word efficiency and (2) phonemic decoding efficiency subtests.
- Multiple tests of reading/listening comprehension, including the *Gray Oral Reading Test*, fourth edition (GORT, passages 5, 7, and 9; Wiederholt & Bryant, 2001); the *Stanford Diagnostic Reading Test*, fourth edition (SDRT, subtest 7: fast reading; Karlson & Gardner, 1995), the *Gates-MacGinitie Reading Tests*, fourth edition (GMRT; MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000); and the *Peabody Individual Achievement Test-Revised* (PIAT-R; Markwardt, 1998; only odd numbered items were administered to assess reading comprehension; even numbered items were recorded and presented aurally in order to assess listening comprehension, as in Spring and French, 1990). Nine participants were missing scores on the Gates-MacGinitie due to a comparatively late adoption of this test into our testing protocol.
- We additionally used the following measures to assess various related abilities: receptive vocabulary skill was assessed by the *Peabody Picture Vocabulary Test-Revised* (PPVT-R; Dunn & Dunn, 1997); verbal working memory was assessed by an auditory version of the Daneman and Carpenter (1980) Sentence Span task; print exposure was assessed by magazine and author checklists based on Cunningham and Stanovich (1990); and spelling ability was assessed using items from the experimental spelling tests in Shankweiler, Lundquist, Dreyer, & Dickinson (1996). Finally, memory for serial order was assessed by a non-verbal task (Corsi Block-

tapping; Corkin, 1974; Berch, Krikorian, & Huha, 1998) in which participants had to reproduce increasingly long visuospatial patterns by tapping on an irregular arrangement of 9 circles displayed on a touch-sensitive computer screen. The patterns occur in blocks of five at each of the lengths from three through ten. The participant's score is the longest sequence that he or she can successfully reproduce three out of five times. Two participants were missing data from the Corsi Block-tapping task due to equipment failure; one participant was missing data from the spelling test due to experimenter error.

3.3 Procedure

We created four counterbalanced lists, within which each experimental item occurred only once, and across which each experimental item occurred in all conditions. All stimuli were presented using the E-Prime experimental package (Schneider, Eschman, & Zuccolotto, 2002). For each item, participants viewed two separate screens. In the Memory Load conditions, participants viewed a screen containing a three-word memory list prior to sentence reading. This screen appeared for 3 s, during which participants were instructed to read the words aloud and to maintain them in memory. The memory words appeared simultaneously, centered on a single line and separated by dashes (e.g., *table* ----- *sink* ----- *truck*). In the No Load conditions, participants viewed a screen containing the words "No Memory Load" for 3 s prior to sentence reading. Following the 3 s memory load screen, participants read sentence items in a self-paced reading paradigm. Participants read by pressing a button that revealed the sentences phrase by phrase, according to the demarcation scheme shown in Table 1. Pressing the button to reveal the next phrase caused the current phrase to revert to a series of dashes.

After reading the final phrase, participants answered a comprehension question, indicating "yes" or "no" by using the "1" and "3" on the keyboard number pad, respectively. In the Memory Load conditions, participants were then prompted to type the three-word lists into the computer via the keyboard. They were asked to type each word in its correct position, and to leave blank any position corresponding to a word they did not recall. If participants could not recall the serial order of the three words, they were allowed to type them in any order. The next trial began immediately after the recall task. In the No Load conditions, the next trial began immediately after answering the comprehension question.

3.4 Data Analysis

We used mixed-effect modeling (Baayen, 2004, 2008; Baayen, Davidson, & Bates, 2006) to analyze the data. Analyses of recall and comprehension accuracy used mixed effects models with a logit link function, because these are binomial outcomes (Jaeger, 2008). This method eliminates the need for separate analyses of random effect variables (i.e., separate ANOVAs testing subjects (F_1) and items (F_2) in order to derive F_{MIN}) by accounting for their potential interaction, and is robust in the face of missing data. Particularly important for the current study, these models do not require that continuous variables (such as working memory capacity, reading skill, or any of the individual differences variables included here) be artificially categorized. In addition, statistical power is improved compared to standard ANOVA analyses (see Baayen, 2004, for simulation results). All statistical analyses were carried out with the R statistical software, version 2.15.2 (R Development Core Team, 2004), using package *lme4*. Mixed-effects models included fixed effects of Load, Interference, and the interaction of Load and Interference. Deviation coding was used for both Load (No Load = −0.5; Load = .5) and Interference ("sailed" = −0.5; "fixed" = .5), in order for their effects to be interpretable as "main effects." Models also included fixed effects of the individual differences measures (plus their interactions with the experimental variables; see below). Individual difference measures were converted to standard scores ($M = 0$, $SD = 1$). Finally, models also included random effects of participants and items.

4. Results

4.1 Descriptive Summary of Skill Measures

Range, means and standard deviations for each battery measure are shown in Table 2. To aid interpretability, we also provide grade or age equivalents where available; CTOPP composite scores do not have grade or age equivalents, so we include percentile ranks. Correlations among the measures are shown in Table 3, below the diagonal. Correlations among the measures after adjustment for IQ are shown above the diagonal (see Section 6.1 for discussion).

4.2 Experimental data

Data from the following dependent measures were collected during the experiment: phrase reading time, comprehension question accuracy, and memory list recall accuracy. For phrase reading time, each sentence was divided into 6 separate regions (see example, Table 1). Of these, Region 5 contains the main verb, and is therefore the region of greatest interest for the interaction of memory load and interference effects; all other regions are identical across conditions. Reading time data in each region of interest was trimmed prior to analysis. Reading times that were lower than 100 ms were excluded from analysis. In addition, following Baayen (2008), individual outliers were identified for each participant and item (using quantile-quantile plots) and removed manually from the data set. This procedure eliminated the possibility of arbitrarily trimming outlying participant/item data points that were nonetheless part of regular distributions. These procedures together led to the exclusion of 3.3% of the data.

4.3 Working memory capacity

Our initial analysis followed the common practice of investigating individual differences as indexed by working memory capacity (WMC; see, e.g., Just & Carpenter, 1992; Daneman & Marikle, 1996; Waters & Caplan, 1996). Although we conducted additional analyses (see below) using the full set of individual differences measures, we conducted this analysis focusing on WMC in order to connect this work with previous research focusing solely on WMC. Each of our experimental dependent measures was submitted to a mixed-effects model that included fixed effects of WMC and the interactions of WMC × Memory Load, WMC × Interference, and Memory Load × WMC × Interference.

Table 4 shows the modeling results for all dependent measures.

4.3.1 Reading Times—We analyzed reading times for each of the six phrase regions. Scatter plots showing the interaction of condition and WMC are reported for Regions 1 – 6 in Figures 1A – 1L. A main effect of readers' WM capacity is evident in all regions; the positive values indicate that higher WMC was associated with longer reading times. In addition, a main effect of Memory Load appears in all regions, with the presence of memory items yielding shorter reading times (Panel A vs. B; C vs. D, E vs. F, G vs. H, I vs. J, and K vs. L). There were also significant interactions of Memory Load × WMC in Regions 1 and 3; this interaction also approached conventional significance in Region 5 (p = .0527). In each case, the load manipulation elicited longer reading times in those participants with higher WMC. There are no significant main effects of Interference; however, there was a small but reliable interaction of Memory Load × Interference in Region 4. This interaction cannot be interpreted, however, because the conditions are identical in Region 4. Of more interest is the small but reliable interaction of Interference × WMC, which appeared in the spillover region (Region 6). A scatter plot showing this interaction appears in Figure 2, collapsing across Memory Load condition. This figure suggests that higher span participants read the region after the interfering verb (*fixed*) more slowly than after the non-interfering verb (*sailed*), while the lower span readers showed the opposite pattern. This is also difficult to interpret, however, because the interaction collapses over the memory manipulation, which would have created interference in the *fixed* conditions but not in the *sailed* conditions when the memory words were present. This is discussed further below. Notably, the three-way interaction of primary interest was not significant for reading times in any region.

4.3.2 Recall—We computed recall results by both a strict criterion (all words correctly recalled in the correct serial order) and a lenient criterion (all words correctly recalled irrespective of order). No differences between these were evident: the proportions of correct responses in the interference condition were .66 (lenient) and .65 (strict), while the proportions in non-interference conditions were .64 (lenient) and .63 (strict). The difference between lenient and strict recall criteria was not significant, t (2323.867) = −0.6061, p = 0.5445. Given this, we only report analyses of strict recall performance in Table 4. A significant main effect of working memory capacity is evident, with low span readers performing more poorly than higher span readers. A scatter plot showing the interaction of condition and WMC is reported for Recall in Figure 1N. There is no significant main effect of interference, and the interaction between interference and WMC is also non-significant.

4.3.3 Comprehension accuracy—We observed a reliable main effect of WMC, such that high span participants were more accurate overall. Scatter plots showing this interaction appear in Figures 1O (No Load conditions) and 1P (Load conditions). In addition, we observed a significant main effect of Memory Load, indicating that the presence of the memory words decreased comprehension accuracy overall. The main effect of Interference was not significant; however, we observed a reliable three-way interaction, which is the primary effect of interest (see Table 4). Scatter plots with panels split by Interference condition, are shown in Figure 3. When the verb creates interference from the memory words (*fixed*, right panel), those with lower WMC are more affected, but when there is no interference caused by the verb (*sailed*, left panel) WMC matters less.

4.3.4 Other individual difference measures—In addition to working memory capacity, we also modeled the effects of the other individual difference (ID) measures from our battery on participant performance. These measures were modeled singly (i.e., we did not enter all the measures into a single model), in order to separately evaluate the influence of each variable while protecting against inaccurate estimates that may arise due to the likely collinearity of many of our tests (cf. Table 3; see Kuperman & Van Dyke, 2011, for a similar approach). As with Sentence Span, we observed numerous significant main effects on reading time and recall accuracy. The results of these models are available as Supplementary Materials, Tables 13–36. The most important observation from these analyses is that nine (9) of the other battery measures showed the same critical 3-way interaction with Memory Load and Interference on the comprehension measure as was observed with WMC: these are summarized in Table 5.[1] In reading time measures, we observed a Memory Load × ID interaction in the critical region with 13 measures other than sentence span (summarized in Table 6), with the same pattern described previously: less skilled participants' reading speed increases with load. The implications of these results are discussed further below.

4.4 Discussion

Because the original Van Dyke and McElree (2006) study was not concerned with individual differences, the ID interactions observed here are novel. Nevertheless, a number of similarities with the original study are present. As in the original study, we observed a main effect of Memory Load in all regions, suggesting that the presence of the memory words encouraged participants to read more quickly, perhaps so that they could get to the recall task that followed the sentence before they forgot the words. This main effect is qualified here by interactions of Memory Load × WMC in several regions, including the marginally significant interaction in the critical region, suggesting that participants responded differently to the load manipulation: low span readers read more quickly than higher span readers when there was also a memory list. This result could be interpreted as consistent with traditional capacity-based explanations: longer reading times for higher WMC readers may be due to the "additional resources" they can apply to the dual task. However, our observation that the same interaction was present and, unlike sentence span, significant in the critical region for 13 other skill measures suggests that WMC *per se* is unlikely to be the primary factor influencing reading times in this task. In addition, the capacity approach has difficulty accounting for the primary finding in this study: the significant three-way interaction of Memory Load × Interference × WMC on comprehension accuracy, because the memory load is identical in both the interfering sentences and the non-interfering sentences. That is, the capacity model provides no basis for predicting that the interference manipulation would affect low and high WMC readers differently, yet the interference effect caused by the verb was more pronounced for lower skilled readers. In addition, this interaction was also observed with nine other skill measures, casting further doubt on WMC as the primary determinant of difficulty in our comprehension measure. We

[1]In addition, three measures produced marginal three-way interactions that approached conventional significance: magazine recognition, $p = .0537$, spelling (words), $p = .0508$, and Woodcock-Johnson Word Attack, $p = .0738$. These are not shown in Table 5.

examine the uniqueness of WMC's contribution to the observed effects in three additional analyses, reported below.

5. Working memory capacity as a spurious determinant of poor comprehension

The results reported above point to substantial individual differences both in locus and degree of susceptibility to interference from the items in memory; however, the theoretical discrepancies discussed above, and the shared variance among the measures (cf. Table 3, below the diagonal) raises significant questions about the role of WMC. We sought to clarify this issue using several methods, described in Sections 5.1–5.3. First, we present the results of an exploratory factor analysis (EFA) conducted on all our individual differences measures in order to extract latent factor estimates, which we then used together with the Memory Load and Interference fixed effects as predictors of our experimental dependent measures. The goal of this analysis was to discover whether a "WMC factor" would emerge as a separate predictor of our critical three-way interaction. Second, and based on the results of the EFA, our next analysis sought to more directly address the issue of shared variance by analyzing the data after accounting for variance shared between WMC and IQ, a more general measure of cognitive ability. Finally, we conducted an analysis in which we created composite scores for the abilities measured in our battery, and simultaneously entered all these composites as predictors in mixed-effect models of our dependent variables. Like the EFA, this analysis allows us to simultaneously consider all our skill measures as predictors; however, unlike the EFA, this approach preserves maximal interpretability of the individual factors in the model so that it is possible to draw conclusions about how specific skills relate to our dependent measures.

5.1 Exploratory Factor Analysis

Individual difference measures were submitted to an exploratory factor analysis with oblimin rotation and principal axis factoring. Participants with missing data on any of the individual differences measures were excluded from this analysis (N = 50). Converging results from Horn's parallel analysis, the very simple structure (VSS) criteria, and the minimum average partial (MAP) criteria suggested a two-factor solution, which accounted for 51% of the variance. We also examined a three-factor solution, although we do not report the results of this analysis because it accounted for only 5% more of the variance, and it included an additional factor with only one loading (for a measure of print exposure, author recognition). Pattern matrix loadings for the two-factor solution are reported in Table 7.

Measures concerned with phonological awareness (e.g., 1), various measures of memory, including working memory (e.g., 2, 18, 19), sentence and passage comprehension (e.g., 7, 8, 9, 13, 14, 15, 16; but see 12), vocabulary (e.g., 17), and IQ (e.g., 24) loaded more strongly onto Factor 1. Measures concerned with rapid naming (e.g., 3, 4), word and non-word reading (e.g., 5, 6, 10, 11), and spelling (e.g., 22, 23) loaded more strongly onto Factor 2. Measures of print exposure (e.g., 20, 21) were split across the two factors.

To assess the role of these factors, we extracted factor score estimates using Thurstone's (1935) regression approach. Estimates were computed for each participant for each factor based on the EFA loadings and participants' scores on each of the individual differences measures. These scores were then used as predictors in mixed-effects models for each dependent variable, together with fixed effects of Memory, Interference, and Memory Load × Interference, as well as their interactions with the Factor 1 (F1) and Factor 2 (F2) scores. Scatter plots showing the relationship between the factor scores and reading times (ms), recall, and comprehension accuracy are plotted in Figure 4 for Factor 1 and Figure 5 for Factor 2. Modeling results are reported in Table 8.

The models revealed a reliable main effect of Factor 1 across a range of dependent measures, including reading times in Regions 2–6, and comprehension accuracy. There was also a reliable interaction between Factor 1 and Memory Load in Regions 2, 3, and 5, a finding reminiscent of that reported above for the WMC measure. Similarly, the critical three-way interaction between Factor 1, Memory Load, and Interference was reliable on comprehension accuracy, and not for any other dependent variables. There was only a reliable interaction between Factor 2 and Memory Load on recall, and no reliable 3-way interactions with Factor 2.

Attempts to interpret our factors are speculative, due to the multiple loadings on each (hence, we adopt the nondescript labels "Factor 1" and "Factor 2"). Nevertheless, we notice a pattern consistent with the "simple view of reading" (Gough & Tunmer, 1986; Hoover & Gough, 1990; Tunmer & Chapman, 2012), which is an approach to understanding reading comprehension that has been highly influential among clinically oriented reading researchers. This model argues that the variance in reading comprehension can be described as the product of oral language skill (viz. our Factor 1) and word-level decoding (viz. our Factor 2) when these abilities are measured appropriately. The pattern of loadings we observed on our Factors 1 and 2 is consistent with loadings in other larger-scale factor-analytic studies aimed at evaluating the simple view (e.g., Kendeou, Savage, & van den Broek, 2009; Protopapas, Mouzaki, Sideridis, Kotsolakou, & Simos, 2013; Protopapas, Simos, Sideridis, & Mouzaki, 2012; Tunmer & Chapman, 2012), in which many of the identical skill measures were employed. Of particular importance, WMC did not emerge as an independent or separable factor contributing to poor comprehension. Rather, the fact that it loaded together with other higher-level reading measures, suggests that the predictive power of WMC in studies where it is the only index of variability may be through its shared variance with other measures.

Although these aspects of the EFA are compelling, we emphasize that they should be interpreted cautiously. We acknowledge that EFA may not be an ideal approach to addressing intercorrelations in our measures: one "rule-of-thumb" for this kind of analysis recommends a minimum of 10–15 participants per variable (Field, 2000). Furthermore, because we used listwise deletion in our EFA (i.e., we excluded participants with missing data), the sample size for this analysis is smaller than for the preceding analyses (N = 50). Nevertheless, we include this analysis because there are some circumstances in which smaller sample sizes yield reliable factor analyses, such as when factor loadings are very high (Stevens, 2002; Rietveld & Van Hout, 1993).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

5.2 IQ-partialled skill measures

As in previous research (e.g., Tunmer & Chapman, 2012), many of the measures in our test battery showed significant correlations. As discussed above, these correlations make it difficult to assess the contribution of any particular skill (viz. working memory) towards susceptibility to interference. The current analysis addresses this issue by partialling out variance shared between WM and IQ. Partialling was also conducted on every other battery measure so that we could additionally assess the relationship of these individual skills to our dependent measures once shared variance was reduced. We chose this method because IQ is a domain-general construct that accounts for a large amount of variance in human performance in general, and has not been reliably or discriminatively associated with reading ability (e.g., Stanovich, 1991; Shaywitz, Holford, Holahan, Fletcher, Stuebing, Francis, & Shaywitz, 1995; Ferrer, McArdle, Shaywitz, Holahan, Marchione, & Shaywitz, 2007). In addition, there is a broad consensus that, despite their shared variance, the constructs of IQ and WMC are not identical. Meta-analyses (e.g., Ackerman, Beier, & Boyle, 2005; Kane, Hambrick, & Conway, 2005; Oberauer, Schulze, Wilhelm, & Süβ, 2005) report that the two constructs share no more than 50% of their variance, leaving a great deal of variance left for each measure to explain separately. Finally, the results of the EFA showed that IQ loaded most heavily on our first factor, suggesting that IQ accounts for a significant amount of variance in the factor on which WMC also loaded most strongly. Hence, we deemed this approach to be a simple and expedient strategy for decreasing collinearity among our battery measures, and for determining whether the residual variance explained by WMC is a significant predictor of sensitivity to interference in our comprehension task. Moreover, this analysis enabled us to include as many participants as possible in our models, unlike the EFA analysis in which we had to drop any participant with missing data from the entire analysis, leaving only N=50. Here, participants were dropped only from the analysis that included the specific individual difference measure for which they had no data. Thus, all models had an N=65 except for the model for spelling, which had N=64, models for rapid naming, oral comprehension, and corsi blocks, which had N=63, and the model for Gates McGinitie comprehension, which had N=56. (See Section 3.2.2. for more details about missing data.)

As expected, partialization caused the correlations among the residualized battery measures to be reduced, in many cases substantially. These are shown in Table 3, above the diagonal. Using these new scores, we reanalyzed all the mixed effect models with each skill measure as a predictor. Tables 16–39 in Supplementary Materials show the complete modeling results for each of the other IQ-partialled ID measures. Previously, we found that WMC, along with nine other individual difference measures (including IQ; see Tables 4 and 5), entered into the critical three-way interaction with the Memory Load and Interference variables on comprehension accuracy. Here, we find that after partialling out variance shared with IQ, WMC is no longer a significant predictor of this effect. Table 9 reports the full modeling results for WMC. Table 10 shows parameter estimates for the interaction term in each of the models for the other nine individual difference measures that had previously shown the critical three-way interaction on comprehension accuracy. After partialling out variance shared with IQ, only a single measure entered into the significant three-way interaction with Interference and Memory Load on the comprehension measure: receptive vocabulary. Table 11 shows the full modeling results for this measure and Figure 6 presents scatter plots of the 3-way interaction split by interference condition. This figure indicates that, when the verb creates interference vis-à-vis the words held in memory ("*fixed*" conditions), readers with lower receptive vocabulary skill answered comprehension questions less accurately than did readers with greater vocabulary skill.

Reading time results for the Receptive Vocabulary measure were similar to those described in Section 4.3 for the unpartialled WMC measure in several ways. The Memory Load effects in reading times, which were observed in Van Dyke & McElree (2006) are replicated; and the Memory Load x Receptive Vocabulary interaction is present in most regions, most notably in the critical region (Region 5, which contains the main verb). The interaction showed the same pattern as described in Section 4.3: participants with low vocabulary skill sped up more in the Memory Load conditions than did those with higher vocabulary skill. We note, however, that five other measures were better predictors of this effect than receptive vocabulary (see Table 12). We discuss this further in Section 5.3. As with the WMC measure, the significant three-way interaction was not significant in any region for the Receptive Vocabulary measure. This interaction was significant in the spillover region (Region 6) for two other measures, however: serial order memory ($p = .0284$) and magazine recognition ($p = .0202$). Full results for these two measures are presented in Supplemental Tables 33 and 35, respectively. Figure 7 shows a scatterplot of these interactions split by interference condition. We reserve the discussion of these effects for Section 5.3, where a similar result is obtained in our final analysis.

The most important outcome of this analysis is that only a single skill measure (receptive vocabulary) remained as a predictor of the Memory Load x Interference x Skill interaction for comprehension accuracy after removing variance shared with IQ. This interaction suggested that participants with lower receptive vocabulary scores, relative to their higher scoring peers, were more susceptible to interference from the memory words in conditions in which the verb did not uniquely distinguish its proper direct object (e.g., *fixed*); in contrast, vocabulary skill mattered little when a verb which uniquely distinguished its proper direct object (e.g., *sailed*) eliminated interference from the memory words. Further, the analogous interaction that we observed in the unpartialled analysis in Section 4.3, which indexed individual differences in comprehension ability by WMC, was no longer significant when variance shared with IQ was removed. This provides additional support for our suggestion that there is nothing unique about the relationship between WMC and performance in our comprehension task or participants' sensitivity to interference. Of additional interest in this analysis is the observation of the three-way interaction in reading times, which might be expected based on the original finding in Van Dyke & McElree (2006), where the effect of interest was observed in reading times and not in comprehension scores. We discuss this further in Section 5.3 below.

5.3 Composite Measures

The previous analysis presents a straightforward demonstration that the initial finding reported in Section 4.3, in which WMC predicts performance on our comprehension task through its role in the Memory Load x Interference x Skill interaction, was not due to any

NIH-PA Author Manuscript

NIH-PA Author Manuscript

specific contribution that WMC *per se* makes in comprehension processes. The fact that this effect disappeared when variance shared with IQ was removed, and an entirely different measure—receptive vocabulary—emerged as the primary predictor of the Memory Load x Interference x Skill interaction suggests that previous findings emphasizing WMC may be spurious due to its shared variance with many other abilities. We now turn to the question of which of our individual differences measures *best* predict our dependent measures. To address this question, it is necessary to simultaneously include all individual differences measures within a single model. However, models that include all of our individual difference measures as individual predictors are likely to suffer from problems of overfitting and multicollinearity (e.g., unstable estimates, inflated SEs; indeed, this was confirmed by our preliminary analyses). The latter issue is exacerbated by the fact that our battery includes multiple measures that address overlapping theoretical constructs. For example, the Gray Oral Reading Test, the Woodcock-Johnson passage comprehension subtest, Stanford Diagnostic Reading test, the Gates-MacGinitie Reading Test, and the Peabody Individual Achievement Test passage comprehension subtest are all correlated at or above $r = 0.50$, and all address aspects of sentence- and passage-level reading. The inclusion of multiple measures of the same construct is common in clinical investigations of reading ability, with the aim of using the similar measures to create composite variables, which more robustly represent the particular construct being assessed (e.g., Braze et al., 2007; Guo, Roehrig, & Williams, 2011; Sabatini, Sawaki, Shore, & Scarborough, 2010; Shankweiler et al., 2008). We follow this approach in the current analysis. Participants with missing data on any of the individual differences measures were excluded from this analysis (N = 50).

Composite measures were created for the following constructs: phonological processing (CTOPP phonological awareness, phonological memory, rapid naming, and rapid color naming); word/nonword reading (WJ-III word reading and identification); word/nonword fluency (WJ-III reading fluency, TOWRE reading words and nonwords); listening sentence/passage comprehension (WJ-III listening comprehension and PIAT-R speech passage comprehension); reading sentence/passage comprehension (WJ-III reading comprehension, Gray Oral Reading Test, SDRT, Gates-MacGinitie, and PIAT-R print passage comprehension); print experience (recognition authors and magazines); and word/nonword spelling (spelling words and nonwords). We generated composite scores by averaging z-scores of the respective individual difference measures. The following individual difference measures were also included in our mixed-effects models, but were not included in a composite because each uniquely addressed a separate construct: receptive vocabulary, working memory, visual memory, and IQ. To further address multicollinearity in our models, we partialled all measures on IQ (see the IQ-partialled analysis above), thus reducing correlations among the predictors. This permitted the IQ measure itself to bear all of its associated variance in the mixed-effect models so that we could evaluate its contribution to our dependent measures directly. Table 13 presents simple correlations among all the composite and non-composite measures used in this analysis.

Each of our experimental dependent measures was submitted to a mixed-effects model with fixed effects of Load, Interference, and Load x Interference, as well as their interactions with each of the skill measures (composite and remaining non-composite measures). Examination of the condition number kappa in our models (all kappa < 10), and the variance inflation factor among the predictors in our models (all VIF < 5) indicates that multicollinearity is not problematic in these analyses. As our experimental design aimed at investigating the 3-way interaction of Memory Load x Interference x ID, we report the results for this interaction for each dependent measure in Table 14. Lower-order interactions with Interference are not interpretable because the Interference is created through the simultaneous presence of the Memory Load words and the verb that matches them. Nevertheless, we present all the significant lower-order interactions and main effects for each dependent measure in Table 15 for completeness.

The main result from this analysis converges with the results of the EFA presented in section 5.1 and the IQ-partialled analysis presented in section 5.2. Receptive vocabulary emerged as the only additional significant predictor of the 3-way interaction on the comprehension measure, even when all individual differences measures were considered simultaneously. Figure 8 shows scatterplots of these interactions, broken out by Load condition. As discussed previously, those with poorer scores on the Receptive Vocabulary measure showed increased susceptibility to Interference from the memory words in the Load condition.

Also consistent with the IQ-partialled analysis presented in section 5.2, we found that serial order memory entered into a significant three-way interaction with Memory Load and Interference in the reading times in Region 6 (the spillover region). Unlike in the previous analysis, where magazine recognition also entered into a significant three-way interaction in Region 6 with Memory Load and Interference, here when magazine recognition was combined with author recognition to provide a more stable composite indicator of print experience, the three-way interaction was not significant. This suggests that the previous effect arose due to shared variance among other measures. Finally, we found that receptive vocabulary entered into an additional significant three-way interaction, this time in reading times for Region 6. Figure 9 shows scatter plots of these interactions. These plots reveal an important difference in the three-way interaction observed in the reading times as compared with the interaction in the comprehension measure; namely, that it is the *better* scorers who are more affected by the interference during online reading.

This contrast in the direction of the interference effect appears to stem from differences in how individuals responded to the dual task, as evident from the pervasive Memory Load x Skill interaction. Although a number of skills interacted with Memory Load in various reading regions (see Table 15), the Load x Print Experience interaction was the most pervasive, shown in Figure 10.[2] Participants with less overall reading experience sped up more in the load conditions than did those with more reading experience, possibly because they chose to prioritize the recall task over the sentence reading task in our procedure. Hence, they paid little attention to the Interference manipulation in the sentence, which was inconsequential until the comprehension task forced them to explicitly recall the meaning of the sentence. At this point the receptive vocabulary measure, out of all verbal skill measures,

[2]The Load x Print Experience interaction occurred in all regions save Region 4, where it showed only a trend towards significance at p = .11.

best indexed sensitivity to interference: readers with poor receptive vocabulary had more difficulty answering the comprehension question when interference was present. This relationship is readily explained if we understand the receptive vocabulary measure not as an index simply of whether participants know the meaning of words (since all of the words in our sentences were quite common), but rather as an index of the *quality* of an individual's lexical representations. It is well known that poor readers exhibit diminished ability to discriminate orthographic, phonological, and semantic features, suggesting that they have lower quality lexical representations than skilled readers (Landi & Perfetti, 2007; Nation & Snowling, 1998; 1999; 2004; Nation, 2009; Perfetti & Hart, 2001; Perfetti, 2007; Yang & Perfetti, 2006). Low quality representations have also been specifically linked to less reading experience (e.g., Balass, Nelson, & Perfetti, 2010; Cain & Oakhill, 2011). Thus, our findings suggest that less distinct lexical representations make it more difficult to *reconstruct* the actual dependency in the sentence during offline comprehension (e.g., was it the *boat* that was fixed, or one of the other three "fixable" nouns they had focused attention on while reading?).

In contrast, individuals with *higher* receptive vocabulary scores showed sensitivity to the interference manipulation in the *online* reading task. This is consistent with the results of the original Van Dyke & McElree (2006) study, which was conducted with college students who (presumably) would be comparable to the more skilled readers in the current study. As in that study, the interference manipulation caused the relatively skilled readers to slow down when retrieval was necessary to resolve the grammatical dependency online, while reading the sentence. Online use of retrieval cues from the manipulated verb, and the higher quality lexical representations in these individuals, resulted in a more immediate semantic interference effect due to the semantic overlap among the "fixable" items. Thus, the slower reading times reflect an online interference effect – and one that the more skilled readers could resolve fairly quickly, resulting in better eventual comprehension compared to their less skilled peers.

Of additional interest is our finding of an interaction of Memory Load x Interference x serial order memory in the reading times in Region 6. Although our measure of serial order memory (the Corsi Blocks Test) is widely used in psychometric studies of reading, to our knowledge it has not previously been associated with any online reading comprehension measures, or with interference effects of any sort. Hence, this result must be treated with caution, but it is nevertheless intriguing. The Corsi Blocks Test is a non-verbal measure in which participants tap out a remembered sequence of increasing length on a visuospatial array. While it is often used as a measure of visual memory, it has also been reliably linked to reading behavior. A number of previous studies have demonstrated a positive relationship between serial order memory and reading ability (Corkin, 1974; Gould & Glencross, 1990; Katz, Shankweiler, & Liberman, 1981; Katz, Healy, & Shankweiler, 1983; Torgesen, 1978) and movement sequencing and reading ability (Birkett & Talkcott, 2012; Carello, LeVasseur, & Schmidt, 2002; Denckla, 1985; Gladstone, Best, & Davidson, 1989; Kuperman, Ally, & Van Dyke, 2013; Wolff, Michel, Ovrut, & Drake, 1990), both of which are implicated in the Corsi Blocks task. Our current findings suggest that individuals with better serial order memory were more affected by interference (Figure 9), which is consistent with a positive correlation between reading ability and performance on the Corsi task. This finding requires further investigation, however, we believe that the sequencing component of the task is likely most relevant to the current findings. We speculate that this association may be related to syntactic parsing ability, as the ability to process the displaced ordering of elements in the clefted constructions examined here (see Table 1) is crucial for correctly integrating the manipulated verb into the sentence.

6. General Discussion

The current study makes three contributions to the study of reading comprehension. First, it points to retrieval interference as the primary determining factor for accurate comprehension; the implication of this is that differences in reading skill may best be characterized in terms of susceptibility to interference, rather than the size of an individual's working memory capacity. Second, and consequently, it provides support for an alternative model of the relationship between memory and language, which more clearly articulates the mechanisms through which the two interact. Finally, our third contribution is to help clarify the relationship between working memory span and language comprehension, while emphasizing alternative factors that contribute to poor comprehension. We discuss each of these in more detail below.

Our first contribution not only provides further support for retrieval interference effects in language comprehension, it provides important evidence for wide variability in susceptibility to interference in our population. Van Dyke & McElree (2006) demonstrated retrieval interference in a more homogeneous population of college students in the same age range as investigated here, thus providing important evidence for cue-based retrieval in sentence processing. However, that study did not investigate individual differences in sensitivity to interference. The present study shows that Memory Load, Interference, and individual cognitive abilities interact, with considerable differences in the time course and extent to which individuals are affected by interference, thus highlighting its significant role in determining comprehension.

Further, our analysis of sensitivity to interference in relation to our battery of skill measures provides an indication as to the mechanisms that determine this sensitivity. Receptive vocabulary, and not working memory capacity, was most consistently associated with vulnerability to interference. That receptive vocabulary was the best predictor of interference effects is in line with a number of other recent findings that implicate vocabulary as a pivotal measure in assessing individual differences in linguistic performance (e.g., Braze, Tabor, Shankweiler, & Mencl, 2007; Prat & Just, 2011; Traxler & Tooley, 2007; Tunmer & Chapman, 2012). Importantly, we are not making the trivial claim that comprehension fails because readers do not know the meanings of words that they are reading. Rather, we believe that the vital role of vocabulary signifies that it is a fundamental element in an architectural account of comprehension difficulty (cf. Tunmer & Chapman, 2012), in which the memory retrieval mechanism plays a primary role. Decades of memory research (e.g., Dosher, 1976, 1981; McElree & Dosher, 1989; Ratcliff, 1978; Wickelgren, 1977) have established that the probability of retrieving particular items depends on the strength or distinctiveness of the representation itself. Thus, it follows that one critical factor for comprehension should be the robustness of the lexical representations themselves. If the to-be-retrieved lexical representations are “noisy” – that is, weak and representationally indistinct – then the probability of accessing a target item that is necessary to complete the long distance dependencies investigated here decreases, and comprehension suffers. This is particularly apparent in the case of poor reading ability, where individuals are likely to have a greater proportion of poor quality lexical representations. Indeed, a particularly troublesome aspect of such representations is their potential to spread spurious activation to neighboring representations, by virtue of their inexact or lower dimensional feature structures. This would result in the activation of irrelevant information, which will interfere with retrieval and comprehension. Evidence for this phenomenon has been observed by Gernsbacher and colleagues, who showed that poor readers were less able to inhibit the context-irrelevant meanings of ambiguous words during sentence comprehension than skilled readers (Gernsbacher, Varner, & Faust, 1990; Gernsbacher & Robertson, 1995; see also Gernsbacher, 1990; Gernsbacher & Faust, 1991, 1995; Long, Oppy, & Seely, 1999). Thus, retrieval operations on low-quality lexical representations will necessarily be less efficient, and more vulnerable to interference, resulting in poorer comprehension in general, and greater difficulty when comprehension requires retrieving distal information to complete grammatical dependencies in particular.

Our second contribution, which is to more clearly articulate the relationship between memory and language, follows on a long line of experimental work suggesting that the size of active memory is actually quite limited, perhaps only to the single most recently processed constituent *even for skilled readers* (see McElree, 2006, for a review). Such a restricted capacity calls the explanatory utility of capacity-based approaches into question. Further, it suggests a new conceptualization of the architecture that supports language comprehension, in which a cue-based retrieval process provides the computational power necessary to create dependencies in real time (for a reviews, see Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Johns, 2012; for a computational implementation of such a system, see Lewis and Vasishth, 2005). The plausibility of this approach is also supported by mathematical analyses of reaction time distributions (Ratcliff, 1978) and evidence from the speed-accuracy tradeoff paradigm (e.g., McElree, 2001), which suggest that humans are able to restore items to active memory in as little as 80–90ms. Such rapid retrieval speeds effectively enable the parsing mechanism to compensate for the severe limit on the size of active memory, permitting parsing decisions to be made in ~250 ms, which is typical for real-time language processing. The result is a model in which accurate and efficient language comprehension can occur even in the face of a highly restricted memory capacity. Critical for the current results, only a retrieval-based model predicts the Memory Load x Interference interaction observed here (and elsewhere, e.g., Van Dyke & McElree, 2006); capacity-based models are silent about potential interactions between properties of the head of a dependency, its filler, and any other content in memory. Such models only make predictions about detrimental effects related to the *quantity* of information that is actively maintained in memory (which was held constant in the current experiment), and do not address potential effects related to the *contents* of memory. The current results show that successful comprehension depends upon the efficient use of retrieval cues to distinguish target items from a field of distractors, which could be highly related to the target on various dimensions (e.g., semantic, syntactic, pragmatic).

Our strategy of connecting a particular skill assessment (here, receptive vocabulary knowledge) to a particular component of a well-articulated model of language processing (here, retrieval interference) is a departure from most clinically inspired reading research, which often focuses on factor-analytic analyses that predict reading comprehension as assessed by off-line standardized instruments (e.g., Woodcock-Johnson Reading Comprehension, Nelson-Denny Reading Comprehension, Stanford Achievement Test, or other governmentally-sanctioned standardized assessments). The simple view of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) has been highly influential in this community, providing the theoretical starting point for data modeling. It is noteworthy that even with a modest sample size, our exploratory factor analysis extracted two factors that are consistent with this view. At the same time, it should be noted that much recent research in this tradition has argued that this simple model should be augmented by a number of other factors, including processing speed (Joshi & Aaron, 2000; Tiu, Thompson, & Lewis, 2003), rapid naming speed (Johnston & Kirby, 2006), reading fluency (Adlof, Catts, & Little, 2006; Silverman, Speece, Harring, & Ritchey, 2013) and, of special interest in light of our current results, IQ (Tiu et al., 2003) and vocabulary (Braze et al., 2007; Tunmer & Chapman, 2012). We believe that a greater understanding of the factors that contribute to poor reading comprehension can be gained through clearly articulated process models that specify the mechanism(s) through which each of these components (and possibly others) may interact. Our retrieval model, which we believe has provided a new understanding of how low vocabulary skill may contribute to poor comprehension ability, represents an initial step in this direction. At the same time, we acknowledge that the current dataset is not comprehensive in its assessment of factors that may lead to increased sensitivity to interference. For example, it is possible that poor vocabulary ability may have its effect *in conjunction with* poor inhibitory mechanisms, the latter of which were not independently assessed in our current battery. This would be consistent with research showing a relationship between low reading ability and increased interference in the Stroop task (Booth & Boyle, 2009; Long & Prat, 2002; Protopapas, Archonti, & Skaloumbakas, 2007). Future research will seek to disentangle the separate roles of vocabulary knowledge and inhibition as contributors to increased sensitivity to retrieval interference.

Finally, our third important contribution is our assessment of the utility of the working memory capacity construct, its relation to other skill measures, and its overall role in individual differences research. Working memory capacity, as measured by “complex span” tasks (such as the “sentence span” task; Daneman and Carpenter, 1980), is often the only individual difference measure obtained in studies of adult language processing and comprehension. Our finding that this measure is strongly and significantly correlated with many other measures of language ability emphasizes the difficulty that overreliance on this measure presents for understanding the causes of poor comprehension. Although the use of capacity measures has been motivated and encouraged by their close relationship to Baddeley’s construct of Working Memory, the evidence presented here suggests that the WM model is not adequate to explain the mechanism by which memory systems and complex language systems interact. This dovetails with a growing consensus – among those who investigate WMC *outside* of the domain of language – that WMC plays its greatest role in situations where it is necessary to overcome automatic, prepotent responses (e.g., Bunting, Conway, & Heitz, 2004; Conway & Engle, 1994; Unsworth & Engle, 2007). This contrasts sharply with sentence processing, which is one of the most reflexive tasks humans can perform in most natural contexts (i.e., those not manipulated experimentally). In this context WMC might be expected to be inconsequential, at least until normal processing mechanisms have failed and deliberate reanalyses or reinterpretation becomes necessary. In contrast, word knowledge will always matter, because the dimensionality of representations of meaning and grammatical function can affect the *probability* of retrieving words in general, and the *discriminability* of the cues contributed by individual words in particular.

One lingering question raised by the current work involves the relationship between WMC and IQ. The high correlation between these measures, as observed here, is consistent with numerous studies reporting extremely high associations between WMC and fluid intelligence (e.g., Ackerman et al., 2002; Colom, Rebollo, Palacious, Juan-Espinosa, & Kyllonen, 2004; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2002; Kyllonen & Christal, 1990; Suß, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). The close relationship between WMC and IQ provides the potential for reconciling previous results that have been interpreted as supporting a critical role for WMC as a determinant of poor comprehension. Recent factor analytic research suggests that the majority of the variance shared between WMC and IQ is associated with neural activity in brain areas responsible for interference control (e.g., in an n-back task; for details see Burgess, Gray, Conway, & Braver, 2011). Although this work did not investigate language comprehension processes *per se*, it is likely that both success in processing non-adjacent grammatical dependencies and success in completing the n-back task (which is computationally similar because related items are temporally separated) hinge on retrieval processes. Hence, we interpret this as additional support for the view that sensitivity to interference, derived from faulty retrieval processes, shows great promise as an explanatory factor for poor comprehension ability.

7. Conclusion

The current study is a novel approach to the study of individual differences in adult language comprehension, grounded in the cue-based retrieval framework, which focuses on the content and quality of memory representations, rather than the quantity of information that can be actively maintained in memory. We provide evidence for *retrieval interference* as a key determinant of poor comprehension, and show that out of our battery of 24 verbal skill measures, vocabulary knowledge, and not working memory capacity, is the most consistent predictor of susceptibility to this interference. We suggest that poor vocabulary knowledge has its affect on retrieval through the increased noise associated with low-dimensional lexical representations, which will be more difficult to discriminate from competitors. Poor readers in particular will have a higher proportion of low-dimensional lexical representations, due to lack of reading experience or difficulties in word-level subskills such as phonological decoding.

Finally, we would like to again emphasize a unique aspect of the current work: the investigation of comprehension ability in a community-based sample. This approach represents an advance over the majority of research on adult reading comprehension, which typically utilizes the university subject-pool population. Our observation of broad variability in a variety of ability measures in our adult population, even in skills such as decoding and fluency, which are often assumed to be mastered by late adulthood, is consistent with previous work (e.g., Cunningham et al., 1990; Shankweiler et al., 1996), and points to the need for a more comprehensive understanding of the factors that contribute to poor reading comprehension in adults.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the following grants from the NIH National Institute of Child Health and Human Development: HD 058944 and HD 073288 to Julie Van Dyke (PI), HD 001994 to Jay Rueckl, HD 067364 to Ken Pugh, HD 056200 to Brian McElree (PI), and HD 040353 to Donald Shankweiler (PI). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors are grateful to Erica Davis and Josh Coppola for assistance with data collection; Dave Braze and Donald Shankweiler for consultations regarding skills testing; and Victor Kuperman for consultations regarding statistical analyses.

References

Abney SP, Johnson M. Memory requirements and local ambiguities of parsing strategies. Journal of Psycholinguistic Research. 1991; 20:233–250.

Ackerman PL, Beier ME, Boyle MO. Working Memory and Intelligence: The Same or Different Constructs? Psychological Bulletin. 2005; 131(1):30–60.10.1037/0033-2909.131.1.30 [PubMed: 15631550]

Adlof SM, Catts HW, Little TD. Should the simple view of reading include a fluency component? Reading and Writing. 2006; 19(9):933–958.

Baayen, RH. Mental Lexicon Working Papers. Vol. 1. Edmonton; 2004. Statistics in psycholinguistics: A critique of some current gold standards; p. 1-45.

Baayen, RH. Analyzing linguistic data: A practical introduction to statistics using R. New York, NY: Cambridge University Press; 2008.

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Langauge. 2006; 59(4):390–412.

Baddeley, AD. Working Memory. Oxford, England: Oxford University Press; 1986.

Baddeley AD. The episodic buffer: A new component of working memory? Trends in Cognitive Science. 2000; 4(11):417–423.

Baddeley, AD.; Hitch, GJ. Working memory. In: Bower, GA., editor. Recent advances in learning and motivation. Vol. 8. New York: Academic Press; 1974. p. 47-90.

Balass M, Nelson JR, Perfetti CA. Word learning: An ERP investigation of word experience effects on recognition and word processing. Contemporary educational psychology. 2010; 35(2):126–140. [PubMed: 22399833]

Berman M, Jonides J, Lewis RL. In search of decay in verbal short-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:317–333.

Birkett EE, Talcott JB. Interval timing in children: effects of auditory and visual pacing stimuli and relationships with reading and attention variables. PloS one. 2012; 7(8):e42820. [PubMed: 22900054]

Booth JN, Boyle JM. The role of inhibitory functioning in children's reading skills. Educational Psychology in Practice. 2009; 25(4):339–350.

Braze D, Mencl WE, Tabor W, Pugh K, Constable RT, Fulbright RK, Magnuson JS, Van Dyke JA, Shankweiler DP. Unification of sentence processing via Ear and Eye: An fMRI study. Cortex. 2011; 47:416–431. [PubMed: 20117764]

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):226–243.10.1177/00222194070400030401 [PubMed: 17518215]

Bunting MF, Conway ARA, Heitz RP. Individual differences in the fan effect and working memory capacity. Journal of Memory and Language. 2004; 51(4):604–622.10.1016/j.jml.2004.07.007

Burgess GC, Gray JR, Conway ARA, Braver TS. Neural Mechanisms of Interference Control Underlie the Relationship Between Fluid Intelligence and Working Memory Span. Journal of Experimental Psychology: General. 2011; 140(4):674–692. [PubMed: 21787103]

Burnham K, Anderson D. Multimodel inference. Sociological Methods & Research. 2004; 33(2):261.

Cain K, Oakhill J. Matthew Effects in Young Readers Reading Comprehension and Reading Experience Aid Vocabulary Development. Journal of learning disabilities. 2011; 44(5):431–443. [PubMed: 21772058]

Caplan D, Waters GS. Verbal working memory and sentence comprehension. Behavioral and Brain Sciences. 1999; 22:77–94. [PubMed: 11301522]

Carello C, LeVasseur VM, Schmidt RC. Movement sequencing and phonological fluency in (putatively) nonimpaired readers. Psychological Science. 2002; 13(4):375–379. [PubMed: 12137142]

Clark HH. The language-as-fixed-effect fallacy: A critique of language statistics is psychological research. Journal of Verbal Learning and Verbal Behavior. 1973; 12:335–359.

Colom R, Rebollo I, Palacios A, Juan-Espinosa M, Kyllonen PC. Working memory is (almost) perfectly predicted by g. Intelligence. 2004; 32:277–296.10.1016/j.intell.2003.12.002

Conway ARA, Engle RW. Working Memory and Retrieval: A Resource-Dependent Inhibition Model. Journal of Experimental Psychology: General. 1994; 123(4):354–373. [PubMed: 7996121]

Conway ARA, Cowan N, Bunting MF, Therriault DJ, Minkoff SRB. A latent variable analysis of working memory capacity, short term memory capacity, processing speed, and general fluid intelligence. Intelligence. 2002; 30:163–184.10.1016/S0160-2896(01)00096-4

Corkin S. Serial-ordering deficits in inferior readers. Neuropsychologia. 1974; 12:347–354. [PubMed: 4417664]

Crowder, RG. Principles of learning and memory. Hillsdale, NJ: Erlbaum; 1976.

Cunningham AE, Stanovich KE. Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. Journal of Educational Psychology. 1990; 82:733–740.

Cunningham, A.; Stanovich, K.; Wilson, M. Cognitive variation in adult college students differing in reading ability. In: Carr, TH.; Levy, B., editors. Reading and its development: Component skills approaches. San Diego, CA: Academic Press; 1990. p. 129-159.

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19:450–466.

Daneman M, Merikle PM. Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review. 1996; 3:422–433. [PubMed: 24213976]

Denckla MB. Motor coordination in dyslexic children: Theoretical and clinical implications. Dyslexia: A neuroscientific approach to clinical evaluation. 1985:187–195.

Dosher BA. The retrieval of sentences from memory: A speed-accuracy study. Cognitive Psychology. 1976; 8:291–310.

Dosher BA. The effects of delay and interference: A speed-accuracy study. Cognitive Psychology. 1981; 13(4):551–582.

Dunn, LM.; Dunn, LM. Peabody picture vocabulary test. 3. American Guidance Services; Circle Pines, MN: 1997.

Engle RW, Cantor J, Carullo JJ. Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1992; 18(5):972.

Engle RW, Tuholski SW, Laughlin JE, Conway ARA. Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General. 1999; 128:309–331.10.1037/0096-3445.128.3.309 [PubMed: 10513398]

Fedorenko E, Gibson E, Rohde D. The nature of working memory capacity in sentence comprehension: Evidence against domain-specific working memory resources. Journal of Memory and Language. 2006; 54(4):541–553.10.1016/j.jml.2005.12.006

Ferrer E, McArdle JJ, Shaywitz BA, Holahan JN, Marchione K, Shaywitz SE. Longitudinal models of developmental dynamics between reading and cognition from childhood to adolescence. Developmental Psychology. 2007; 43:1460–1473. [PubMed: 18020824]

Field, A. Discovering Statistics using SPSS for Windows. London: Sage Publications; 2000.

Gernsbacher, MA. Language comprehension as structure building. Hillsdale, NJ: Erlbaum; 1990.

Gernsbacher MA, Faust M. The mechanisms of suppression: A component of general comprehension skill. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1991; 17:245–262.

Gernsbacher, MA.; Faust, M. Skilled suppression. In: Dempster, FN.; Brainerd, CJ., editors. Interference and Inhibition in Cognition. San Diego, CA: Academic Press; 1995. p. 295-327.

Gernsbacher MA, Robertson RRW. Reading skill and suppression revisited. Psychological Science. 1995; 6(3):165–169.

Gernsbacher MA, Varner KR, Faust M. Investigating differences in general comprehension skill. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1990; 16:430–445.

Gibson E. Linguistic complexity: locality of syntactic dependencies. Cognition. 1998; 68:1–76. [PubMed: 9775516]

Gibson, E. The dependency locality theory: A distance-based theory of linguistic complexity. In: Marantz, A., editor. Image, language, brain: Papers from the first mind articulation project symposium. Cambridge, MA: MIT Press; 2000. p. 94-126.

Gladstone M, Best CT, Davidson RJ. Anomalous bimanual coordination among dyslexic boys. Developmental Psychology. 1989; 25(2):236.

Gordon PC, Hendrick R, Johnson M. Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2001; 27:1411–1423.

Gordon PC, Hendrick R, Levine WH. Memory-load interference in syntactic processing. Psychological Science. 2002; 13:425–430. [PubMed: 12219808]

Gordon PC, Hendrick R, Johnson M. Effects of noun phrase type on sentence complexity. Journal of Memory and Language. 2004; 51:97–114.

Gordon PC, Hendrick R, Johnson M, Lee Y. Similarity-based interference during language comprehension: Evidence from eye tracking during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2006; 32(6):1304–1321.10.1037/0278-7393.32.6.1304

Gough PB, Tunmer WE. Decoding, reading, and reading disability. Remedial and special education. 1986; 7(1):6–10.

Gould JH, Glencross DJ. Do children with a specific reading disability have a general serial-ordering deficit? Neuropsychologia. 1990; 28(3):271–278. [PubMed: 2325839]

Guo Y, Roehrig AD, Williams RS. The Relation of Morphological Awareness and Syntactic Awareness to Adults' Reading Comprehension Is Vocabulary Knowledge a Mediating Variable? Journal of Literacy Research. 2011; 43(2):159–183.

Hoover WA, Gough PB. The simple view of reading. Reading and writing. 1990; 2(2):127–160.

Ingels, S.; Dalton, B.; LoGerfo, L. Trends among high school seniors, 1972–2004 (NCES 2008-320). National Center for Education Statistics, Institute for Education Sciences, U.S. Department of Education; Washington, DC: 2008.

Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language. 2008; 59:434–446. [PubMed: 19884961]

Johnston TC, Kirby JR. The contribution of naming speed to the simple view of reading. Reading and Writing. 2006; 19(4):339–361.

Joshi RM. Vocabulary: A critical component of comprehension. Reading & Writing Quarterly. 2005; 21(3):209–219.

Joshi RM, Aaron PG. The component model of reading: Simple view of reading made a little more complex. Reading Psychology. 2000; 21(2):85–97.

Just MA, Carpenter PA. A capacity theory of comprehension: Individual differences in working memory. Psychological Review. 1992; 99:122–149. [PubMed: 1546114]

Kane MJ, Hambrick DZ, Conway ARA. Working Memory Capacity and Fluid Intelligence Are Strongly Related Constructs: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin. 2005; 131(1):66–71.10.1037/0033-2909.131.1.66 [PubMed: 15631552]

Kane MJ, Hambrick DZ, Tuholski SW, Wilhelm O, Payne TW, Engle RW. The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. Journal of Experimental Psychology: General. 2004; 133:189–217.10.1037/0096-3445.133.2.189 [PubMed: 15149250]

Karlson, B.; Gardner, E. Stanford diagnostic reading test. 4. San Antonio, TX: Psychological Corp; 1995.

Katz RB, Shankweiler D, Liberman IY. Memory for item order and phonetic recoding in the beginning reader. Journal of Experimental Child Psychology. 1981; 32(3):474–484. [PubMed: 7320681]

Katz RB, Healy AF, Shankweiler D. Phonetic coding and order memory in relation to reading proficiency: A comparison of short-term memory for temporal and spatial order information. Applied Psycholinguistics. 1983; 4(03):229–250.

Kendeou P, Savage R, Broek P. Revisiting the simple view of reading. British Journal of Educational Psychology. 2009; 79(2):353–370. [PubMed: 19091164]

Kimball. Seven principles of surface structure parsing in natural language. Cognition. 1973; 2:15–47.

King J, Just MA. Individual differences in syntactic processing: the role of working memory. Journal of Memory and Language. 1991; 30:580–602.

Kuperman, V.; Ally, P.; Van Dyke, JA. Finger tapping predicts text comprehension. Abstract submitted to the 17th European Conference on Eye Movements; August 11–16, 2013; Lund, Sweden. 2013.

Kuperman V, Bresnan J. The effects of construction probability on word durations during spontaneous incremental sentence production. Journal of Memory and Language. 2012; 66:588–611.

Kuperman V, Van Dyke JA. Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. Journal of Memory and Language. 2011; 65:45–73.

Kyllonen PC, Christal RE. Reasoning ability is (little more than) working memory capacity?! Intelligence. 1990; 14:389–433.

Landi N, Perfetti CA. An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. Brain and Language. 2007; 102:30–45. [PubMed: 17188346]

Lewis RL, Vasishth S, Van Dyke JA. Computational principles of working memory in sentence comprehension. Trends in Cognitive Sciences. 2006; 10(10):447–454. [PubMed: 16949330]

Long, DL.; Johns, CL.; Morris, PE. Comprehension ability in mature readers. In: Traxler, M.; Gernsbacher, M., editors. Handbook of Psycholinguistics. Burlington, MA: Academic Press; 2006. p. 801-834.

Long DL, Oppy BJ, Seely MR. The strategic nature of less skilled readers' suppression problems. Discourse Processes. 1999; 27:281–302.

Long DL, Prat CS. Working memory and Stroop interference: An individual differences investigation. Memory & Cognition. 2002; 30(2):294–301. [PubMed: 12035891]

Long DL, Prat CS. Individual differences in syntactic ambiguity resolution: Readers vary in their use of plausibility information. Memory & Cognition. 2008; 36(2):375–391. [PubMed: 18426067]

McElree B. Accessing short-term memory with semantic and phonological information: A time-course analysis. Memory & Cognition. 1996; 24:173–187. [PubMed: 8881321]

McElree B. Attended and non–attended states in working memory: Accessing categorized structures. Journal of Memory & Language. 1998; 38:225–252.

McElree B. Sentence comprehension is mediated by content-addressable memory structures. Journal of Psycholinguistic Research. 2000; 29:111–123. [PubMed: 10709178]

McElree B. Working memory and focal attention. Journal of Experimental Psychology: Learning, Memory & Cognition. 2001; 27:817–835.

McElree, B. Accessing recent events. In: Ross, BH., editor. The psychology of learning and motivation. Vol. 46. San Diego: Academic Press; 2006.

McElree B, Dosher BA. Serial position and set size in short-term memory: Time course of recognition. Journal of Experimental Psychology: General. 1989; 118:346–373.

McElree B, Dosher BA. Serial retrieval processes in the recovery of order information. Journal of Experimental Psychology: General. 1993; 122:291–315.

McElree, B. Accessing recent events. In: Ross, BH., editor. The psychology of learning and motivation. Vol. 46. San Diego, CA: Academic Press; 2006. p. 155-200.

McElree B, Foraker S, Dyer L. Memory structures that subserve sentence comprehension. Journal of Memory and Language. 2003; 48:67–91.

MacDonald MC, Just MA, Carpenter PC. Working memory constraints on the processing of syntactic ambiguity. Cognitive Psychology. 1992; 24:56–98. [PubMed: 1537232]

MacGinitie, W.; MacGinitie, R.; Maria, K.; Dreyer, L. Gates-MacGinitie Reading Tests. Itasca, IL: Riverside; 2000. (4E)

MacKinnon, DP. Analysis of mediating variables in prevention and intervention research. In: Cazares, A.; Beatty, LA., editors. Scientific methods in prevention research. Washington, DC: U.S. Govt. Print Office; 1994. p. 127-153.NIDA Research Monograph 139. DHHS Pub. No. 94-3631

Magnuson, JS.; Kukona, A.; Braze, B.; Johns, CL.; Van Dyke, J.; Tabor, W.; Mencl, E.; Pugh, KR.; Shankweiler, D. Phonological instability in young adult poor readers: Time course measures and computational modeling. In: McCardle, P.; Miller, B.; Lee, JR.; Tseng, O., editors. Dyslexia Across Languages: Orthography and the Brain-Gene-Behavior Link. Baltimore: Paul Brookes Publishing; 2011. p. 184-201.

Markwardt, FC, Jr. Peabody individual achievement test–Revised. Circle Pines, MN: American Guidance Service; 1998.

Miller, GA.; Chomsky, N. Finitary models of language users. In: Luce, RD.; Bush, R.; Galanter, E., editors. Handbook of Mathematical Psychology. Vol. II. New York: Wiley; 1963. p. 419-492.Ch. 13

Nairne JS. A feature model of immediate memory. Memory and Cognition. 1990; 18:251–269. [PubMed: 2192233]

Nairne JS. Remembering over the short-term: The case against the standard model. Annual Review of Psychology. 2002; 53:53–81.

Nation, K. Reading comprehension and vocabulary: what's the connection?. In: Wagner; Schatchneider; Phythian-Sence, editors. Reading comprehension: Approaches to understanding its behavioural and biological origins. Guildford Publications; NY: 2009. p. 179-194.

Nation K, Snowling MJ. Beyond phonological skills: broader language skills contribute to the development of reading. Journal of Research in Reading. 2004; 27:342–356.

Nation K, Snowling MJ. Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. Cognition. 1999; 70:B1–13. [PubMed: 10193058]

Nation K, Snowling MJ. Semantic processing skills and the development of word recognition: evidence from children with reading comprehension difficulties. Journal of Memory and Language. 1998; 39:85–101.

Nieuwland MS, Van Berkum JJA. Individual differences and contextual bias in pronoun resolution: evidence from ERPs. Brain Research. 2006; 1(6):155–167. [PubMed: 16956594]

Oberauer K, Kliegl R. A formal model of capacity limits in working memory. Journal of Memory and Language. 2006; 55(4):601–626.

Oberauer K, Schulze R, Wilhelm O, Süß HM. Working Memory and Intelligence--Their Correlation and Their Relation: Comment on Ackerman, Beier, and Boyle (2005). Psychological Bulletin. 2005; 131(1):61–65.10.1037/0033-2909.131.1.61 [PubMed: 15631551]

Öztekin I, McElree B. Retrieval dynamics of proactive interference: PI slows retrieval by eliminating fast assessments of familiarity. Journal of Memory and Language. 2007

Perfetti, CA.; Hart, L. The lexical quality hypothesis. In: Verhoeven, L.; Elbro, C.; Reitsma, P., editors. Precursors of Functional Literacy. Amsterdam: John Benjamins Publishing Company; 2002.

Perfetti CA. Reading ability: Lexical quality to comprehension. Scientific Studies of Reading. 2007; 11(4):357–383.

Prat CS, Just MA. Exploring the neural dynamics underpinning individual differences in sentence comprehension. Cerebral Cortex. 2011; 21(8):1747– 1760. [PubMed: 21148612]

Protopapas A, Archonti A, Skaloumbakas C. Reading ability is negatively related to Stroop interference. Cognitive Psychology. 2007; 54(3):251–282. [PubMed: 16962090]

Protopapas A, Mouzaki A, Sideridis GD, Kotsolakou A, Simos PG. The role of vocabulary in the context of the simple view of reading. Reading & Writing Quarterly. 2013; 29(2):168–202.

Protopapas A, Simos PG, Sideridis GD, Mouzaki A. The components of the simple view of reading: A confirmatory factor analysis. Reading Psychology. 2012; 33(3):217–240.

Ratcliff R. A theory of memory retrieval. Psychological Review. 1978; 85:59–108.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2004.

Repovš G, Baddeley A. The multi-component model of working memory: Explorations in experimental cognitive psychology. Neuroscience. 2006; 139(1):5–21.10.1016/j.neuroscience. 2005.12.061 [PubMed: 16517088]

Rietveld, T.; Van Hout, R. Statistical Techniques for the Study of Language and Language Behavior. New York: Mouton de Gruyter; 1993.

Sabatini JP, Sawaki Y, Shore JR, Scarborough HS. Relationships among reading skills of adults with low literacy. Journal of Learning Disabilities. 2010; 43(2):122–138. [PubMed: 20179307]

Schneider, W.; Eschman, A.; Zuccolotto, A. E-Prime user's guide. Pittsburgh: Psychology Software Tools; 2002. [P.]

Shankweiler DP, Lundquist E, Dreyer LG, Dickinson CC. Reading and spelling difficulties in high school students: Causes and consequences. Reading and Writing. 1996; 8:267–294.

Shankweiler D, Mencl WE, Braze D, Tabor W, Pugh KR, Fulbright R. Reading differences and brain: Cortical integration of speech and print in sentence processing varies with reader skill. Developmental Neuropsychology. 2008; 33(6):745–775. [PubMed: 19005913]

Shaywitz BA, Holford TR, Holahan JM, Fletcher JM, Stuebing KK, Francis DJ, Shaywitz SE. A Matthew effect for IQ but not for reading: Results from a longitudinal study. Reading Research Quarterly. 1995; 30:894–906.

Silverman RD, Speece DL, Harring JR, Ritchey KD. Fluency Has a Role in the Simple View of Reading. Scientific Studies of Reading. 2013; 17(2):108–133.

Spring C, French L. Identifying children with specific reading disabilities from listening and reading discrepancy scores. Journal of Learning Disabilities. 1990; 23:53–58. [PubMed: 2295871]

Stanovich KE. Discrepancy definitions of reading disability: Has intelligence led us astray? Reading Research Quarterly. 1991; 26:7–29.

Stevens, J. Applied Multivariate Statistics for the Social Sciences. 4. Mahwah, NJ: Erlbaum; 2002.

Süß HM, Oberauer K, Wittman WW, Wilhelm O, Schulze R. Working memory capacity explains reasoning ability—and a little bit more. Intelligence. 2002; 30:261–288.

Thurstone, LL. The vectors of mind. Chicago: University of Chicago Press; 1935. p. 226-231.

Tiu RD, Thompson LA, Lewis BA. The role of IQ in a component model of reading. Journal of Learning Disabilities. 2003; 36(5):424–436. [PubMed: 15497486]

Torgesen JK. Performance of reading disabled children on serial memory tasks: A selective review of recent research. Reading Research Quarterly. 1978:57–87.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of word reading efficiency. Austin, TX: Pro-Ed; 1999.

Traxler MJ, Tooley KM. Lexical mediation and context effects in sentence processing. Brain Research. 2007; 1146:59–74.10.1016/j.brainres.2006.10.010 [PubMed: 17112486]

Traxler MJ, Williams RS, Blozis SA, Morris RK. Working memory, animacy, and verb class in the processing of relative clauses. Journal of Memory and Language. 2005; 53:204–224.

Tunmer WE, Chapman JW. The Simple View of Reading Redux: Vocabulary Knowledge and the Independent Components Hypothesis. Journal of Learning Disabilities. 2012; 45(5):453–466. [PubMed: 22293683]

Underwood BJ, Keppel G. An evaluation of two problems of method in the study of retention. American Journal of Psychology. 1962; 75:1–17.

Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. Psychological Review. 2007; 114(1):104–132.10.1037/0033-295X.114.1.104 [PubMed: 17227183]

Van Dyke JA, Johns CL. Memory Interference as a Determinant of Language Comprehension. Language and Linguistics Compass. 2012; 6(4):193–211.10.1002/lnc3.330 [PubMed: 22773927]

Van Dyke JA, McElree B. Retrieval interference in sentence comprehension. Journal of Memory and Language. 2006; 55(2):157–166. [PubMed: 18209744]

Van Dyke JA. Interference effects from grammatically unavailable constituents during sentence processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33(2): 407–430.

Van Dyke JA, Lewis RL. Distinguishing effects of structure and decay on attachment and repair: A retrieval interference theory of recovery from misanalyzed ambiguities. Journal of Memory and Language. 2003; 49:285–413.

Wagers, M.; McElree, B. Focal attention and the timing of memory retrieval in language comprehension. Talk given at the Architectures and Mechanisms for Language Processing Conference; September 7–9; Barcelona. 2009.

Wagner, R.; Torgesen, J.; Rashotte, CA. Comprehensive test of phonological processing. ProEd; Austin, TX: 1999.

Wanner, E.; Maratsos, M. An ATN approach to comprehension. In: Halle, M.; Bresnan, J.; Miller, GA., editors. Linguistic theory and psychological reality. Cambridge, MA: MIT Press; 1978.

Warner J, Glass AL. Context and distance-to-disambiguation effects in ambiguity resolution: Evidence from grammaticality judgment of garden path sentences. Journal of Memory and Language. 1987; 26:714–738.

Warren T, Gibson E. Effects of NP-type on reading English clefts. Language and Cognitive Processes. 2005; 20 (6):751–767.

Warren T, Gibson E. The influence of referential processing on sentence complexity. Cognition. 2002; 85:79–112. [PubMed: 12086714]

Watkins OC, Watkins MJ. Build-up of proactive inhibition as a cue overload effect. Journal of Experimental Psychology: Human Learning and Memory. 1975; 104:442–452.

Waters, GS.; Caplan, D. The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). 1996.

Waugh NC, Norman DA. Primary memory. Psychological Review. 1965; 72(2):89–104. [PubMed: 14282677]

Wechsler, D. WASI: Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation; 1994. 1999

Wickelgren W. Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica. 1977; 41:67–85.

Wickelgren WA, Corbett AT, Dosher BA. Priming and retrieval from short-term memory: A speed-accuracy tradeoff analysis. Journal of Verbal Learning and Verbal Behavior. 1980; 19:387–404.

Wiederholt, JL.; Bryant, BR. Gray oral reading tests-(GORT-4). Austin, TX: Pro-Ed; 2001.

Wolff PH, Michel GF, Ovrut M, Drake C. Rate and timing precision of motor coordination in developmental dyslexia. Developmental Psychology. 1990; 26(3):349.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III tests of achievement. Itasca, IL: Riverside; 2001.

Yang, CL.; Perfetti, C. Reading skill and the acquisition of high quality representations of new words. Society of the Scientific Study of Reading Annual Meeting; Vancouver, Canada. 2006 Jul.

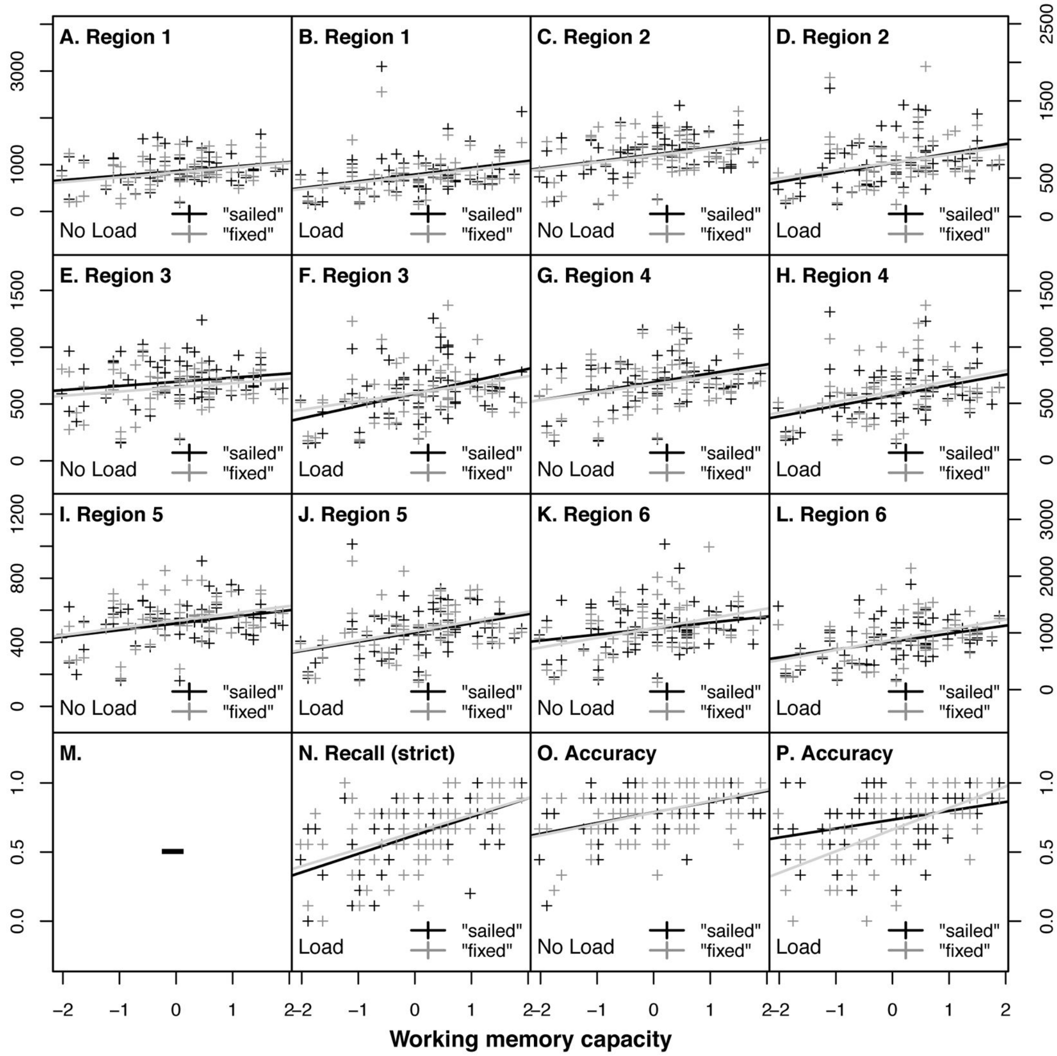

Figure 1.

Scatter plots of reading times (ms), accuracy, and recall by condition and WMC. Interference conditions are denoted by the “fixed” legend entry, and “sailed” corresponds to the “No Interference” conditions.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

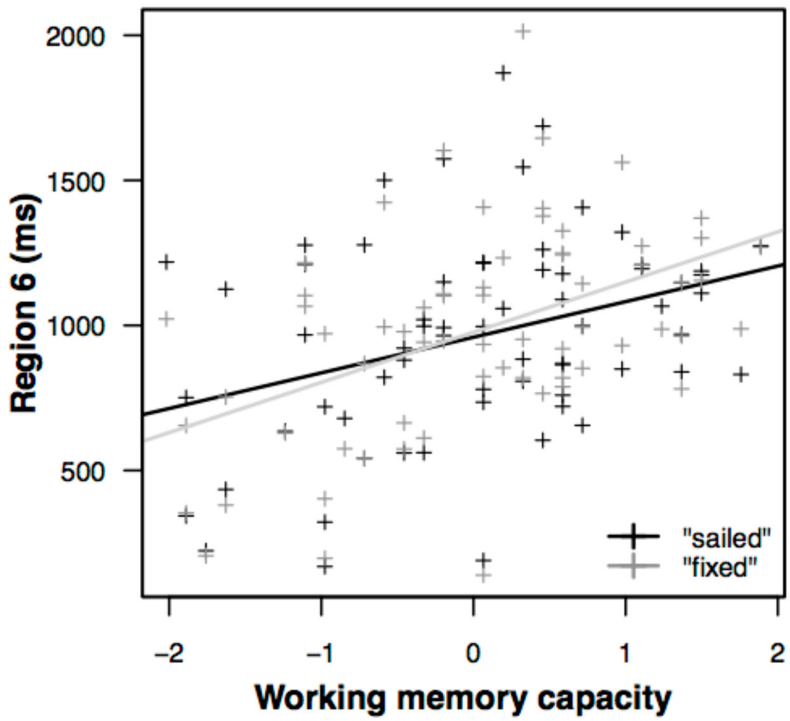

Figure 2.
The interaction of WMC and interference, collapsing across Memory Load condition.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

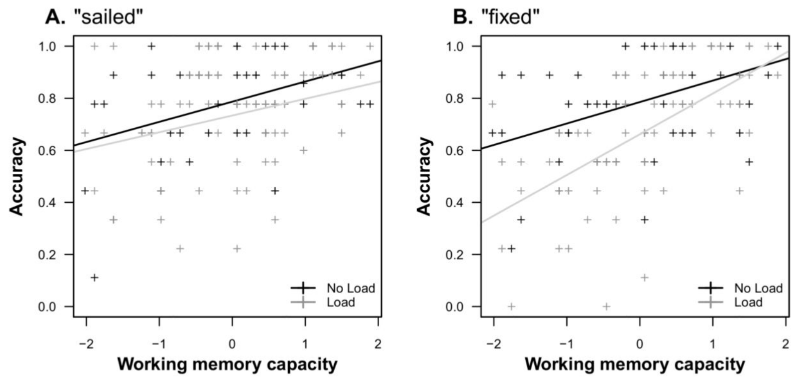

Figure 3.

The interaction of working memory capacity, memory load, and interference. When interference is present (right panel, Load condition), readers with low working memory capacity are more affected.

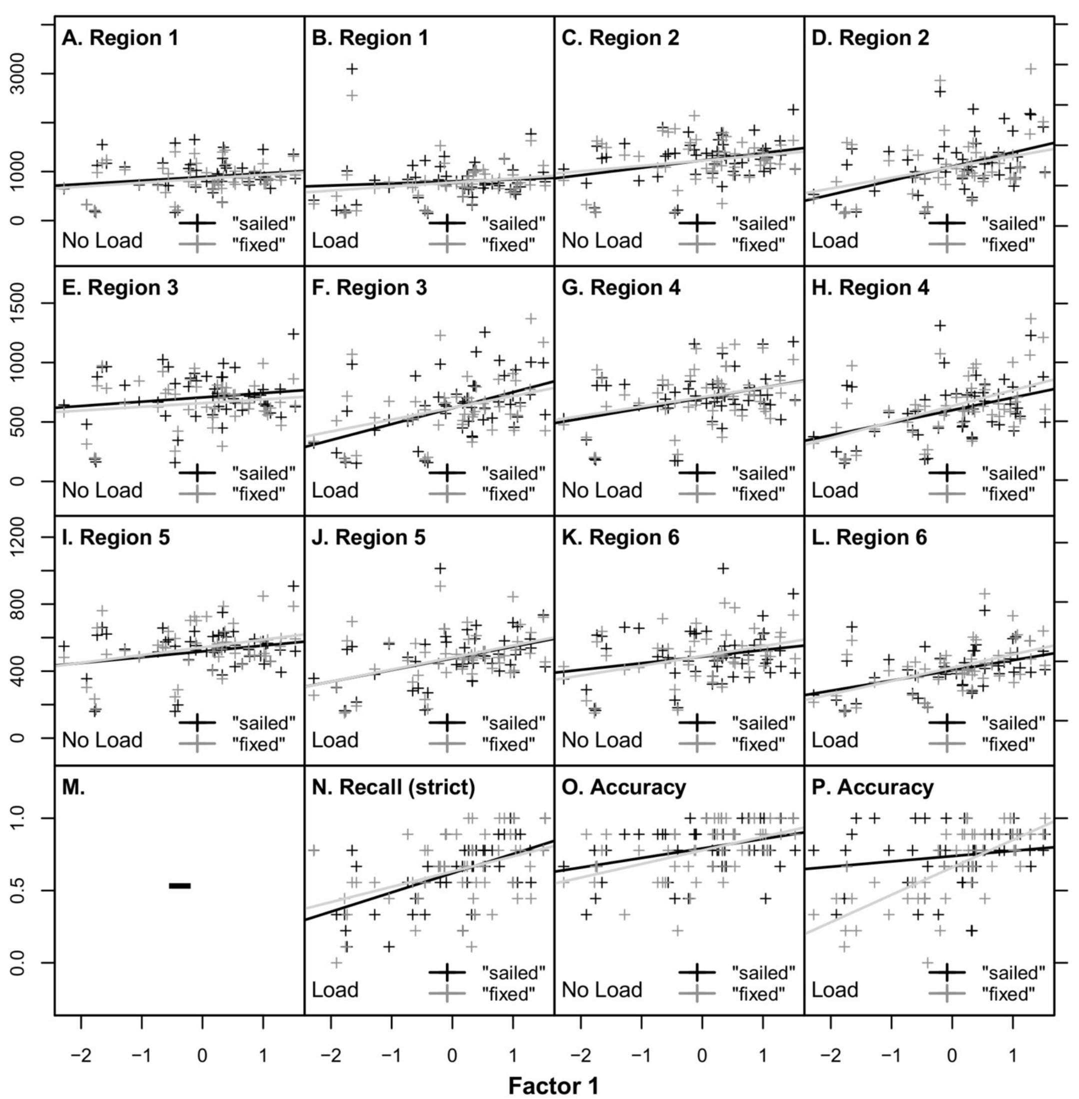

Figure 4.

Scatter plots of reading times (ms), recall, and comprehension accuracy by condition and Factor 1 score.

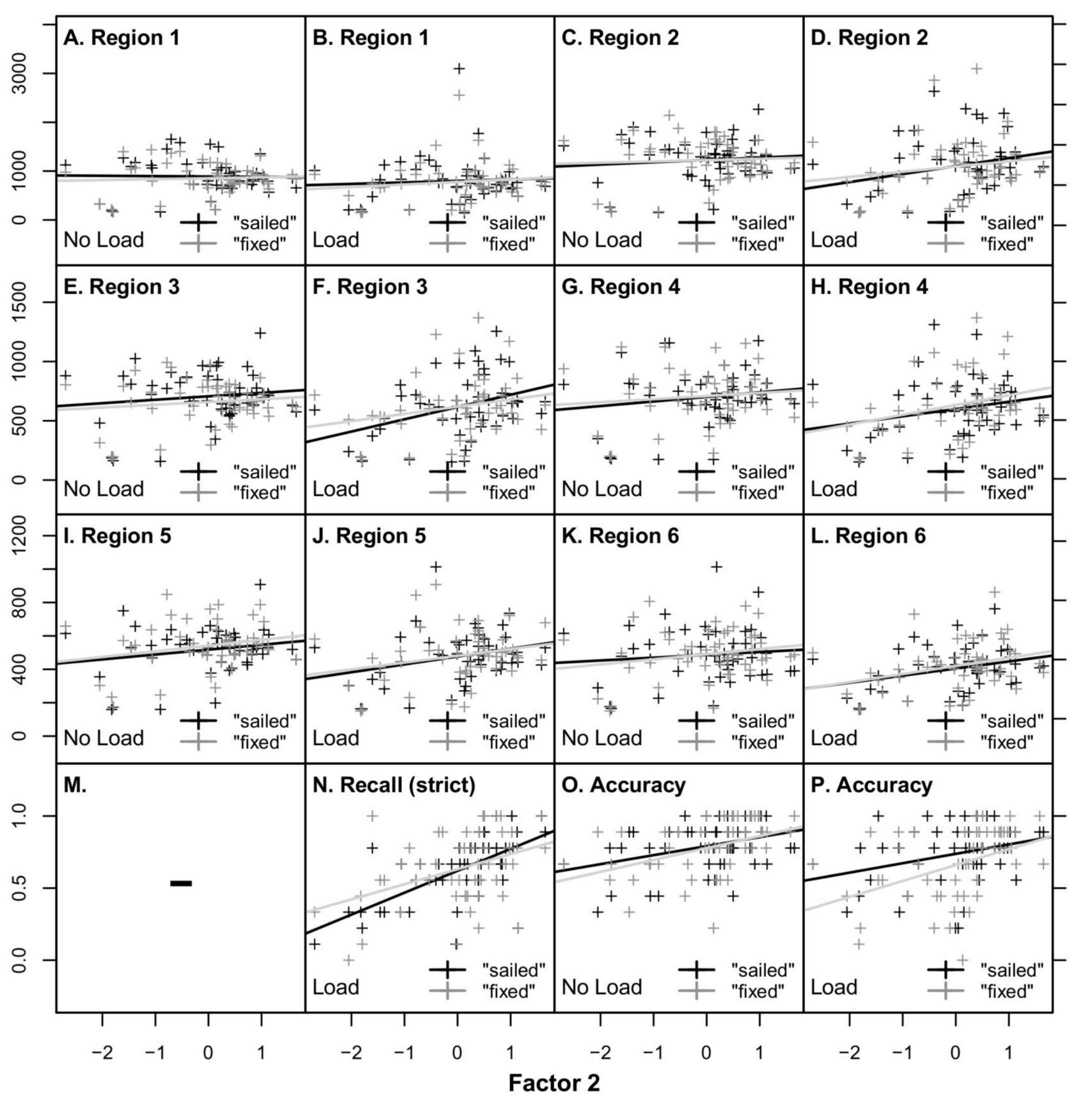

Figure 5.
Scatter plots of reading times (ms), recall, and comprehension accuracy by condition and Factor 2 score.

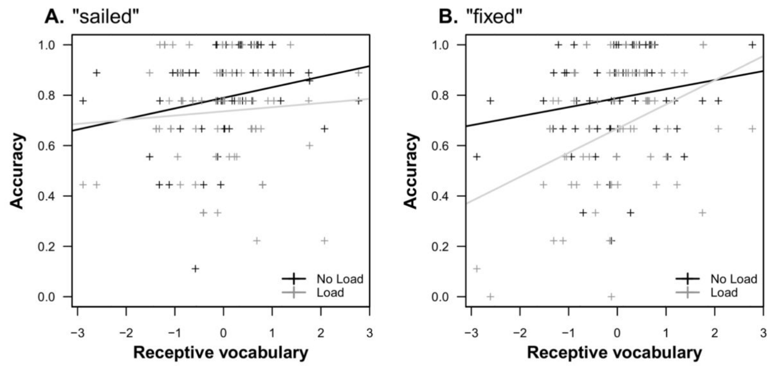

Figure 6.

The interaction of IQ-partialled receptive vocabulary, memory load, and interference. When interference is present (right panel; Load condition), readers with low vocabulary scores are more affected.

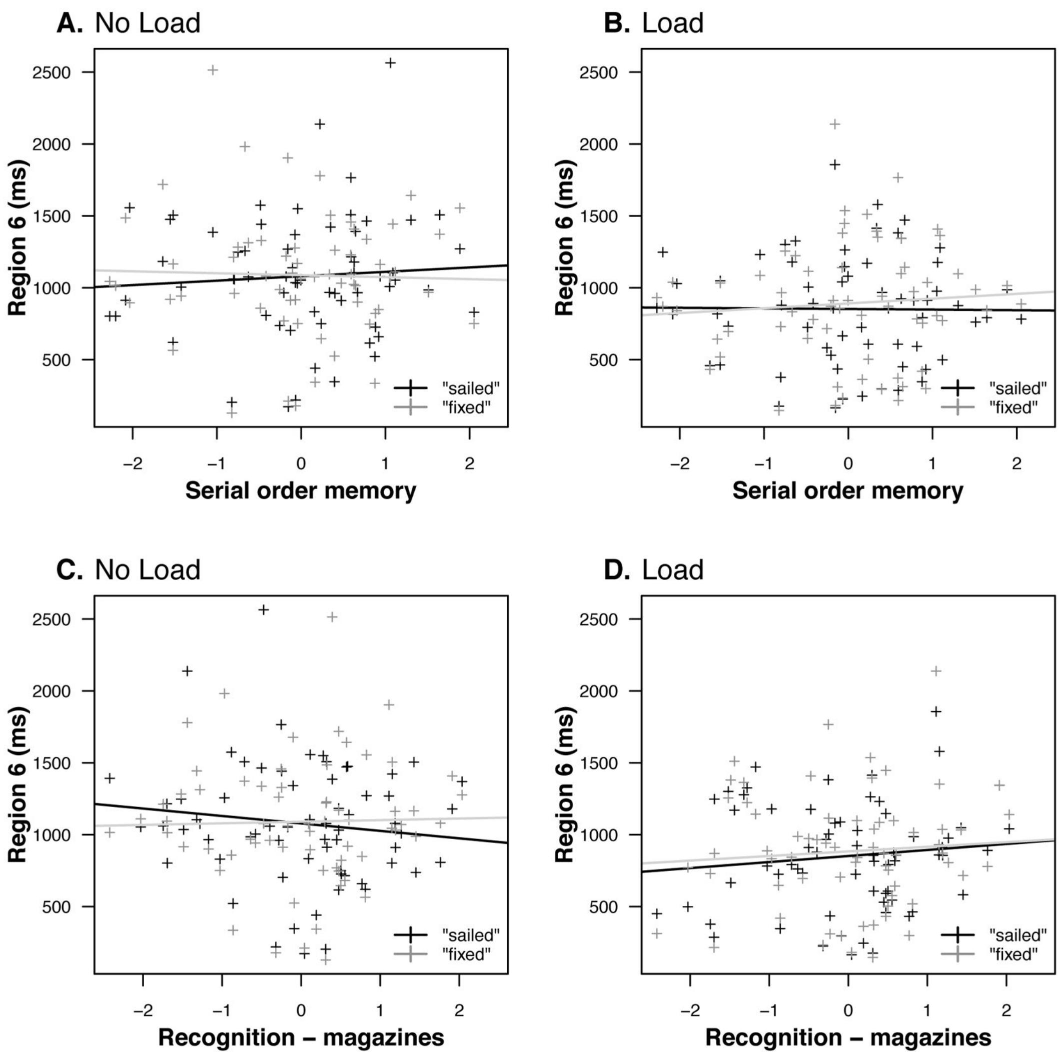

Figure 7.

Scatter plots of reading times (ms) in the spillover region (Region 6) for the Memory Load x Interference x Serial Order Memory (IQ-partialled) interaction (top) and Memory Load x Interference x Magazine Recognition (IQ-partialled) interaction (bottom).

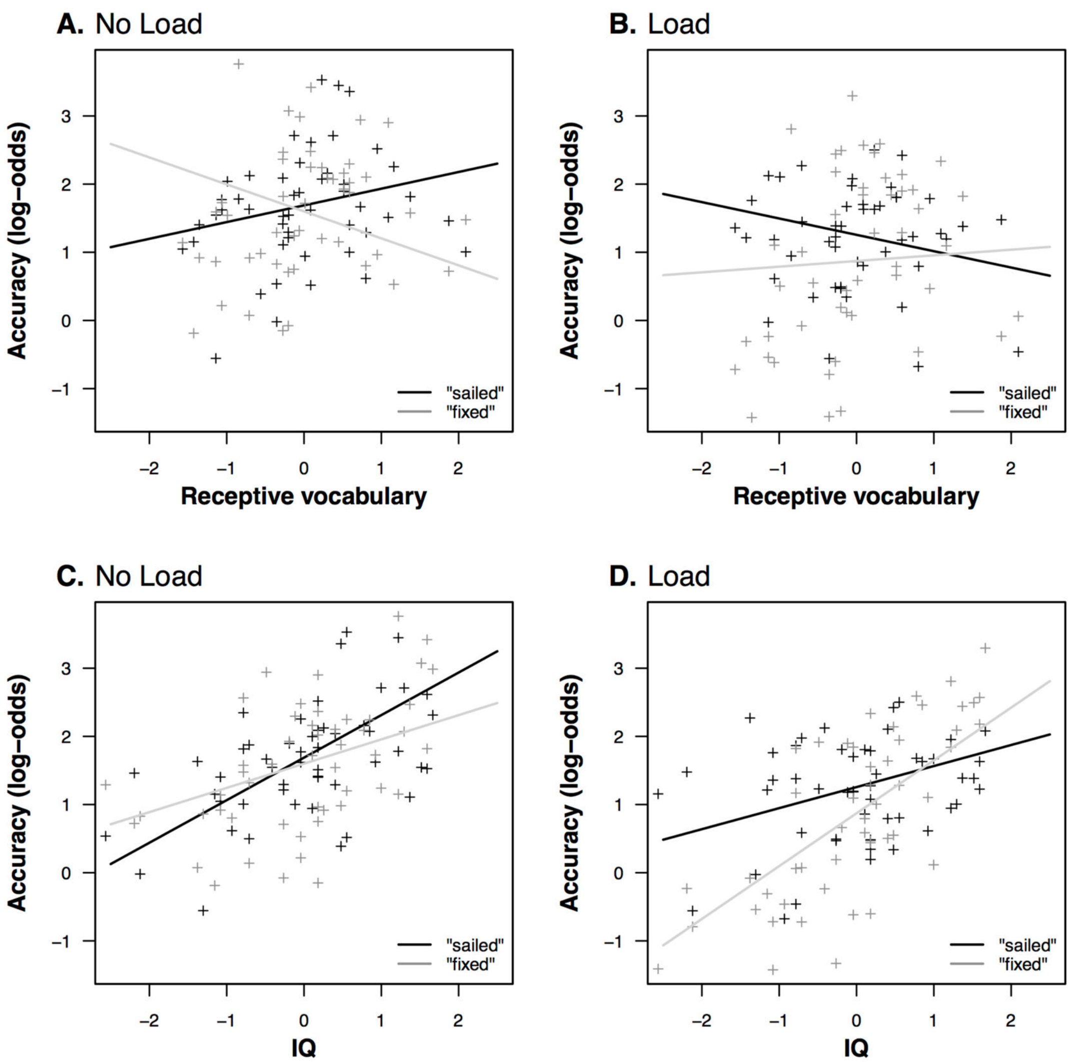

Figure 8.

Scatter plots from the composite analysis of comprehension accuracy. Plots show the interaction of IQ-partialled Receptive Vocabulary (A and B) and IQ (C and D) with Memory Load and Interference. When Interference is present (right panel) readers with lower scores are more affected.

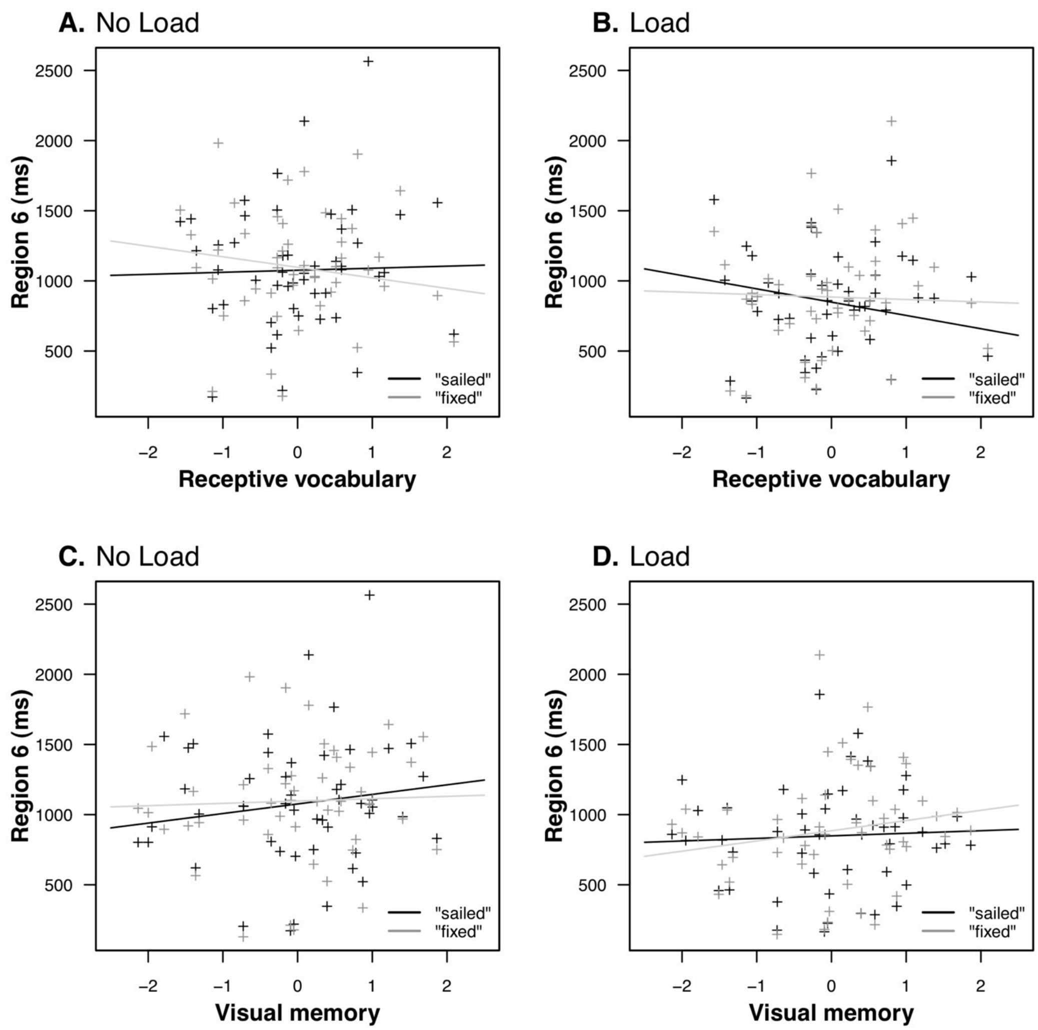

Figure 9.

Scatter plots showing RTs from the composite analysis of Region 6. Plots show the interaction of IQ-partialled Receptive Vocabulary (A and B) and Serial Order Memory (C and D) with Memory Load and Interference. When Interference is present (right panel) readers with higher vocabulary scores are more affected.

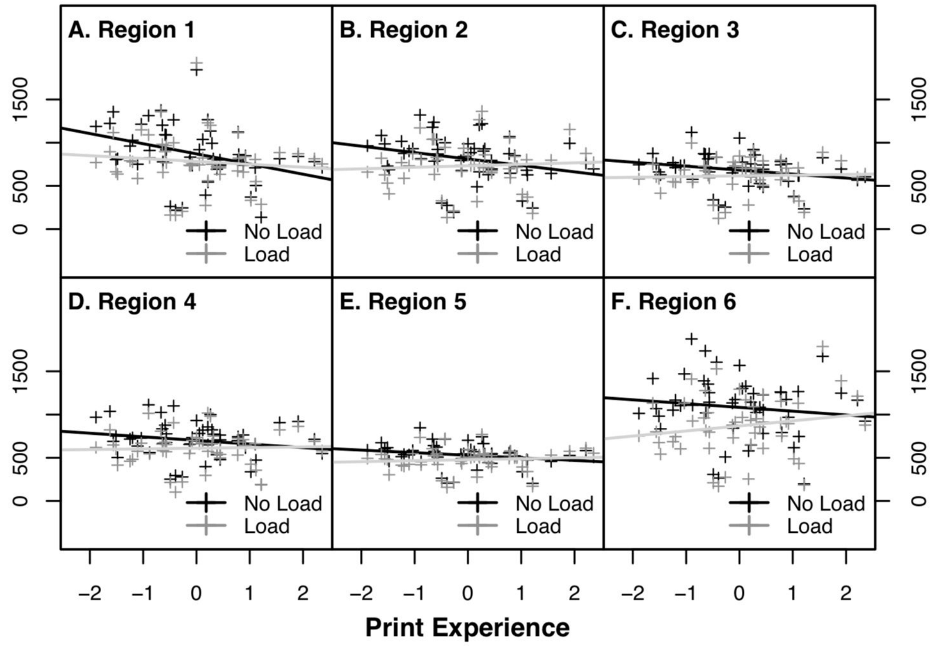

Figure 10.
Load x Print Experience interaction in all reading regions in the composite analysis.

Table 1

Sample Experimental Items. The main verb in the critical region is underlined, but was presented normally to participants.

Conditions	Memory List	Sentence (interference manipulation underlined)	Comprehension Question
Non-interfering/Load	table-sink-truck	It was the boat / that the guy / who lived / by the sea / sailed / in two sunny days.	Did the guy live by the sea?
Interfering/Load	table-sink-truck	It was the boat / that the guy / who lived / by the sea / fixed / in two sunny days.	Did the guy live by the sea?
Non-interfering/No Load	---	It was the boat / that the guy / who lived / by the sea / sailed / in two sunny days.	Did the guy live by the sea?
Interfering/No Load	---	It was the boat / that the guy / who lived / by the sea / fixed / in two sunny days.	Did the guy live by the sea?

Table 2

Range, means, & standard deviations for all cognitive measures

Measure	Range	*M*	*SD*	Max. possible
1. phonological awareness	51 – 118	87.09	17.69	150
Percentile equivalent	<1 – 89	30.62	28.92	99
2. phonological memory	61 – 124	92.24	13.47	150
Percentile equivalent	<1 – 95	35.20	26.14	99
3. rapid naming	67 – 139	101.13	14.78	150
Percentile equivalent	1 – 99	51.5	28.55	99
4. rapid color naming	25.5 – 61.2	39.17	6.41	72
Grade equivalent	4.4 – 9.8	8.92	1.29	9.8
5. word reading (word attack)	7 – 32	25.21	5.44	32
Grade equivalent	1.8 – 19	9.22	4.63	19
6. word identification	44 – 76	64.83	6.69	76
Grade equivalent	3.3 – 19	11.25	4.69	19
7. reading fluency	19 – 98	70.80	17.53	98
Grade equivalent	2.3 – 19	11.93	4.74	19
8. reading comprehension	27 – 43	34.87	3.75	47
Grade equivalent	3.4 – 19	9.50	4.62	19
9. oral comprehension	10 – 31	24.98	4.32	34
Grade equivalent	1.9 – 19	11.67	4.49	19
10. reading – words	63 – 104	89.47	10.55	104
Grade equivalent	4.4 – 12.6	10.13	2.47	12.6
11. reading –nonwords	7 – 62	45.06	12.45	63
Grade equivalent	1.6 – 12.6	8.55	3.29	12.6
12. Gray Oral Reading Test	4 – 29	18.17	6.48	30
13. SDRT	5 – 30	18.53	7.36	30
Grade equivalent	3 – PHS*	9.1	3.22	PHS
14. Gates-MacGinitie	10 – 46	32.00	9.09	48
Grade equivalent	4.9 – PHS	11.68	1.80	PHS
15. print passage comprehension	7 – 39	26.70	7.19	41
16. speech sentence comprehension	9 – 39	30.17	6.70	41
17. receptive vocabulary	115 – 195	167.11	19.85	204
Age equivalent score	8.09 – 22	17.71	4.59	22
18. working memory capacity	26 – 56	41.50	7.68	60
19. serial order memory	2.7 – 7.3	5.33	.94	9
20. recognition – authors	0 – 11	3.23	2.58	80
21. recognition – magazines	0 – 18	7.98	4.52	80
22. spelling – words	9 – 30	22.88	5.41	30
23. spelling – nonwords	0 – 12	4.95	2.28	20
24. IQ	61 – 126	95.88	15.08	–

Note: Measures 1–4: Comprehensive Test of Phonological Awareness (Wagner et al., 1999); 5–9: Woodcock-Johnson-III Tests of Achievement (Woodcock et al., 2001); 10–11: Test of Word Reading Efficiency (Torgeson et al., 1999); 13: Stanford Diagnostic Reading Test (Karlsen & Gardner, 1995); 15–16: Peabody Individual Achievement Test-Revised (Markwardt, 1998); 17: Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1997); 18: listening span (Daneman & Carpenter, 1980); 19: Corsi Blocks (Corkin, 1974); 20–21: Print Exposure (adapted from Cunningham & Stanovich, 1990); 24: Weschler Abbreviated Scales of Intelligence, vocabulary & matrix reasoning subtests (Psychological Corp., 1999).

* PHS indicates Post High School.

Table 3

Simple correlations among measure (N=65).

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. phono. awareness	–	.34	−.02	−.12	.36	.49	.24	.23	.25	.20	.43	.33	.31	.39	.22	.10	.41	.34	.02	.21	.30	.38	.29
2. phono. memory	.57	–	−.07	−.15	.15	.19	.15	.08	−.04	.26	.22	.07	.00	.15	.09	.03	.00	.17	−.07	−.18	.03	.08	.09
3. rapid naming	−.04	−.08	–	−.27	.25	.16	.11	.15	.13	.46	.29	.22	.09	.19	−.05	.08	−.04	.06	.13	.04	−.10	.28	.13
4. rapid color naming	−.12	−.15	−.26	–	−.31	−.28	−.24	−.32	−.14	−.45	−.35	−.37	−.25	−.25	−.23	−.27	−.14	−.33	−.28	−.02	−.21	−.33	−.17
5. word reading	.60	.42	.19	−.29	–	.73	.26	.26	.30	.44	.87	.56	.29	.41	.20	.18	.31	.31	.02	.17	.06	.77	.45
6. word identification	.70	.47	.11	−.25	.83	–	.39	.37	.49	.46	.76	.72	.47	.49	.31	.25	.54	.36	.09	.29	.30	.78	.53
7. reading fluency	.50	.40	.08	−.23	.49	.60	–	.25	.32	.40	.37	.44	.68	.42	.35	.20	.38	.07	.12	.20	.24	.47	.33
8. reading comp.	.57	.42	.08	−.28	.55	.65	.51	–	.47	.34	.35	.29	.47	.55	.30	.44	.53	.34	−.01	.09	.52	.40	.20
9. oral comprehension	.60	.37	.07	−.14	.59	.73	.58	.72	–	.29	.28	.32	.43	.51	.40	.31	.73	.27	−.01	.25	.23	.41	.32
10. reading – words	.33	.37	.43	−.44	.50	.52	.48	.44	.40	–	.53	.51	.44	.48	.41	.27	.22	.23	.13	.20	.19	.47	.27
11. reading –nonwords	.61	.43	.24	−.33	.90	.82	.54	.57	.53	.58	–	.68	.43	.46	.19	.14	.34	.24	.01	.26	.28	.79	.44
12. Gray Oral Reading	.48	.26	.19	−.37	.64	.76	.55	.47	.49	.56	.74	–	.58	.42	.20	.15	.33	.14	.17	.25	.39	.73	.38
13. SDRT	.55	.30	.06	−.24	.52	.66	.78	.66	.65	.51	.59	.66	–	.49	.24	.11	.51	.15	.08	.41	.51	.48	.27
14. Gates-MacGinitie	.61	.39	.10	−.21	.59	.67	.57	.73	.73	.51	.59	.47	.65	–	.47	.44	.60	.28	.05	.17	.38	.58	.39
15. print passage comp.	.50	.36	−.06	−.22	.46	.56	.54	.56	.63	.48	.42	.37	.48	.61	–	.47	.43	.24	.08	.06	.14	.35	.35
16. speech sent. comp.	.51	.40	.04	−.23	.51	.59	.50	.70	.67	.38	.44	.37	.45	.65	.67	–	.34	.19	−.19	.03	.17	.23	.22
17. receptive vocabulary	.67	.38	−.05	−.14	.58	.75	.61	.76	.87	.35	.56	.49	.69	.78	.64	.67	–	.36	−.09	.41	.37	.52	.29
18. working memory	.59	.44	.03	−.30	.55	.61	.37	.62	.59	.35	.47	.34	.43	.53	.50	.53	.63	–	−.03	.18	.17	.32	.07
19. serial order memory	.33	.22	.09	−.27	.30	.38	.36	.32	.36	.25	.26	.33	.34	.33	.34	.24	.29	.28	–	−.10	−.09	.14	.00
20. recog – authors	.39	.06	.03	−.04	.34	.44	.36	.33	.42	.28	.39	.36	.52	.44	.25	.28	.53	.35	.10	–	.30	.27	.00
21. recog – magazines	.37	.15	−.10	−.22	.17	.36	.32	.54	.31	.24	.35	.44	.53	.31	.24	.27	.41	.27	.04	.35	–	.24	.02
22. spelling – words	.58	.34	.22	−.31	.84	.84	.61	.61	.61	.54	.84	.78	.63	.67	.53	.50	.68	.53	.36	.40	.32	–	.48
23. spelling – nonwords	.44	.27	.11	−.18	.56	.62	.46	.40	.48	.35	.54	.47	.42	.50	.48	.41	.45	.28	.19	.14	.11	.57	–
24. IQ	.64	.54	−.03	−.06	.57	.65	.54	.66	.73	.28	.50	.39	.55	.65	.71	.56	.69	.61	.50	.22	.36	.51	.38

$|r| \geq .25$, $p < .05$, $|r| \geq .32$, $p < .01$, $|r| \geq .40$, $p < .001$, $|r| \geq .47$, $p < .0001$

[*] Correlations among measures without adjustment for variance shared with IQ are below the diagonal. Correlations among measures after partialling out variance shared with IQ are above the diagonal.

Note: Measures 1–4: Comprehensive Test of Phonological Awareness (Wagner et al., 1999); 5–9: Woodcock-Johnson-III Tests of Achievement (Woodcock et al., 2001); 10–11: Test of Word Reading Efficiency (Torgeson et al., 1999); 13: Stanford Diagnostic Reading Test (Karlsen & Gardner, 1995); 15–16: Peabody Individual Achievement Test-Revised (Markwardt, 1998); 17: Peabody Picture

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Vocabulary Test-Revised (Dunn & Dunn, 1997); 18: listening span (Daneman & Carpenter, 1980); 19: Corsi Blocks (Corkin, 1974); 20–21: Print Exposure (adapted from Cunningham & Stanovich, 1990); 22–23: spelling tests adapted from Shankweiler et al., 1996; 24: Weschler Abbreviated Scales of Intelligence, vocabulary & matrix reasoning subtests (Psychological Corp., 1999)

Table 4

Mixed-effects modeling results for all dependent measures* – working memory capacity (Sentence Span). Reliable effects are in **bold**.

	Region 1				Region 2				Region 3				Region 4			
	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*
(Intercept)	**806.002**	**37.606**	**21.433**	**0.000**	**748.021**	**32.122**	**23.287**	**0.000**	**632.545**	**22.618**	**27.966**	**0.000**	**633.073**	**25.830**	**24.509**	**0.000**
Load	**−80.825**	**17.955**	**−4.502**	**0.000**	**−101.866**	**17.397**	**−5.855**	**0.000**	**−78.209**	**13.227**	**−5.913**	**0.000**	**−92.363**	**12.069**	**−7.653**	**0.000**
Interference	−30.277	17.949	−1.687	0.092	−4.421	17.394	−0.254	0.799	−23.701	13.229	−1.792	0.073	10.034	12.066	0.832	0.406
WMC	**120.411**	**38.329**	**3.142**	**0.002**	**100.716**	**29.982**	**3.359**	**0.001**	**63.964**	**22.057**	**2.900**	**0.004**	**84.731**	**23.383**	**3.624**	**0.000**
Load × Interference	−0.488	35.912	−0.014	0.989	13.516	34.798	0.388	0.698	47.925	26.457	1.811	0.070	**48.400**	**24.134**	**2.005**	**0.045**
Load × WMC	**44.294**	**18.604**	**2.381**	**0.017**	21.658	18.075	1.198	0.231	**51.228**	**13.677**	**3.746**	**0.000**	21.946	12.518	1.753	0.080
Interference × WMC	−1.113	18.592	−0.060	0.952	−10.663	18.045	−0.591	0.555	−15.855	13.663	−1.160	0.246	−3.054	12.495	−0.244	0.807
Load × Interference × WMC	−10.672	37.209	−0.287	0.774	−12.455	36.129	−0.345	0.730	−31.231	27.360	−1.141	0.254	10.804	25.026	0.432	0.666

	Region 5 (critical region)				Region 6 (spillover region)				Recall (strict)				Comprehension Accuracy			
	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*Z*	*p*	Est. (β)	SE	*Z*	*p*
(Intercept)	**495.270**	**14.872**	**33.303**	**0.000**	**969.463**	**45.380**	**21.363**	**0.000**	**0.71**	**0.17**	**4.16**	**.0000**	**1.333**	**0.184**	**7.240**	**0.000**
Load	**−60.948**	**7.720**	**−7.894**	**0.000**	**−213.664**	**21.212**	**−10.073**	**0.000**					**−0.577**	**0.111**	**−5.196**	**0.000**
Interference	12.283	7.721	1.591	0.112	18.236	21.213	0.860	0.390	0.11	0.13	0.82	0.4112	−0.179	0.111	−1.605	0.108
WMC	**51.141**	**15.243**	**3.355**	**0.001**	**147.083**	**40.093**	**3.669**	**0.000**	**0.39**	**0.15**	**2.56**	**0.0104**	**0.626**	**0.091**	**6.909**	**0.000**
Load × Interference	−10.200	15.444	−0.660	0.509	39.436	42.431	0.929	0.353					−0.321	0.221	−1.452	0.146
Load × WMC	15.454	7.975	1.938	0.053	18.109	21.957	0.825	0.409					0.009	0.114	0.076	0.940
Interference × WMC	3.906	7.976	0.490	0.624	**44.797**	**21.901**	**2.045**	**0.041**	0.06	0.15	0.38	0.7065	0.173	0.115	1.507	0.132
Load × Interference × WMC	−2.364	15.955	−0.148	0.882	−24.174	43.901	−0.551	0.582					**0.500**	**0.228**	**2.198**	**0.028**

*Absent values for Recall reflect the fact that the task was completed only in the Memory Load conditions, when the memory items were present

Table 5

Individual difference measures showing a significant three-way interaction with Memory Load and Interference on comprehension accuracy.

Memory Load × Interference ×	Estimate	Std. Error	*Z*	*p*
1. receptive vocabulary	0.713	0.210	3.406	.0007
2. oral comprehension	0.614	0.206	2.977	.0029
3. Gates-MacGinitie	0.639	0.231	2.768	.0056
4. IQ	0.594	0.222	2.682	.0073
5. reading comprehension	0.561	0.232	2.420	.0155
6. SDRT	0.536	0.222	2.409	.0160
7. sentence span	0.500	0.228	2.198	.0280
8. reading - words	0.495	0.226	2.192	.0284
9. word identification	0.462	0.216	2.137	.0326
10. reading fluency	0.436	0.220	1.986	.0471

Table 6

Individual difference measures showing a significant two-way interaction with Memory Load in the critical region only (Region 5). See Supplemental Materials for the same interaction with multiple skill measures in other regions.

Memory Load x	Estimate	Std. Error	*t*	*p*
1. recognition – magazines	39.584	7.622	5.193	.0000
2. SDRT	38.796	7.662	5.064	.0000
3. rapid naming	−40.388	8.088	−4.994	.0000
4. reading fluency	38.087	7.734	4.925	.0000
5. recognition - articles	31.591	7.812	4.045	.0001
6. receptive vocabulary	27.902	7.674	3.634	.0003
7. oral comprehension	22.861	7.706	2.967	.0030
8. speech sentence comp	22.663	7.677	2.952	.0032
9. IQ	22.495	7.730	2.910	.0036
10. Gates-MacGinitie comp	22.528	8.298	2.715	.0066
11. Gray Oral Reading Test	21.188	7.834	2.705	.0068
12. print sentence comp	19.413	7.691	2.524	.0116
13. Phonological awareness	19.053	7.782	2.448	.0144

Table 7

Factor loadings from the exploratory factor analysis.

Measure	Factor 1	Factor 2
1. phono. awareness	**0.51**	0.30
2. phono. memory	**0.47**	−0.03
3. rapid naming	−0.35	**0.45**
4. rapid color naming	0.03	**−0.39**
5. word reading	0.21	**0.69**
6. word identification	0.37	**0.62**
7. reading fluency	**0.48**	0.25
8. reading comp.	**0.65**	0.18
9. oral comprehension	**0.81**	0.08
10. reading – words	0.05	**0.60**
11. reading –nonwords	0.04	**0.86**
12. Gray Oral Reading	−0.13	**0.94**
13. SDRT	**0.51**	0.37
14. Gates-MacGinitie	**0.69**	0.24
15. print passage comp.	**0.75**	−0.08
16. speech sent. comp.	**0.61**	0.08
17. receptive vocabulary	**0.87**	0.08
18. working memory	**0.62**	0.05
19. serial order memory	**0.32**	0.07
20. recog – authors	0.05	**0.35**
21. recog – magazines	**0.36**	0.26
22. spelling – words	0.13	**0.80**
23. spelling – nonwords	0.27	**0.33**
24. IQ	**0.95**	−0.21

Table 8

Mixed-effects modeling results for all dependent measures – Factor 1 (F1) and 2 (F2) scores. Significant effects in **bold**.

	Region 1				Region 2				Region 3				Region 4			
	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*
(Intercept)	**826.470**	**46.142**	**17.911**	**0.000**	**771.983**	**36.660**	**21.058**	**0.000**	**649.951**	**26.315**	**24.699**	**0.000**	**655.860**	**29.749**	**22.046**	**0.000**
Load	**−87.821**	**20.377**	**−4.310**	**0.000**	**−82.667**	**20.401**	**−4.052**	**0.000**	**−67.69**	**15.478**	**−4.373**	**0.000**	**−87.103**	**14.306**	**−6.089**	**0.000**
Interference	−28.830	20.350	−1.417	0.157	−2.897	20.365	−0.142	0.887	−24.82	15.447	−1.607	0.108	19.836	14.278	1.389	0.165
F1	106.667	62.366	1.710	0.087	**142.918**	**46.685**	**3.061**	**0.002**	**68.259**	**34.490**	**1.979**	**0.048**	**120.315**	**36.039**	**3.339**	**0.001**
F2	−46.835	62.556	−0.749	0.454	−42.738	46.831	−0.913	0.361	9.143	34.581	0.264	0.791	−26.551	36.150	−0.734	0.463
Load × Interference	20.015	40.757	0.491	0.623	−4.281	40.822	−0.105	0.916	42.963	30.946	1.388	0.165	35.083	28.602	1.227	0.220
Load × F1	−40.059	27.943	−1.434	0.152	**63.159**	**28.008**	**2.255**	**0.024**	**77.064**	**21.284**	**3.621**	**0.000**	28.346	19.630	1.444	0.149
Load × F2	**65.195**	**28.084**	**2.321**	**0.020**	25.612	28.301	0.905	0.365	7.578	21.361	0.355	0.723	24.744	19.849	1.247	0.213
Interference × F1	10.169	28.682	0.355	0.723	−17.032	29.560	−0.576	0.564	−1.522	22.201	−0.069	0.945	3.832	20.983	0.183	0.855
Interference × F2	16.259	28.210	0.576	0.564	−14.874	28.627	−0.520	0.603	−24.097	21.575	−1.117	0.264	1.855	20.162	0.092	0.927
Load × Interference × F1	64.050	55.791	1.148	0.251	6.345	55.982	0.113	0.910	−4.631	42.564	−0.109	0.913	38.822	39.189	0.991	0.322
Load × Interference × F2	−54.862	56.145	−0.977	0.328	−19.712	56.591	−0.348	0.728	−28.597	42.730	−0.669	0.503	−13.063	39.686	−0.329	0.742

	Region 5 (critical region)				Region 6 (spillover region)				Recall (strict)*				Comprehension Accuracy			
	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*Z*	*p*	Est. (β)	SE	*Z*	*p*
(Intercept)	**503.269**	**17.084**	**29.458**	**0.000**	**976.689**	**50.698**	**19.265**	**0.000**	0.648	0.168	3.856	0.000	**1.331**	**0.187**	**7.105**	**0.000**
Load	**−51.613**	**8.814**	**−5.856**	**0.000**	**−220.868**	**23.410**	**−9.435**	**0.000**					**−0.580**	**0.127**	**−4.555**	**0.000**
Interference	12.475	8.810	1.416	0.157	26.924	23.355	1.153	0.249	0.034	0.154	0.223	0.824	**−0.267**	**0.127**	**−2.095**	**0.036**
F1	**54.471**	**23.271**	**2.341**	**0.019**	**178.000**	**62.350**	**2.855**	**0.004**	0.332	0.181	1.831	0.067	**0.493**	**0.122**	**4.038**	**0.000**
F2	2.032	23.335	0.087	0.931	−18.631	62.569	−0.298	0.766	**0.470**	**0.184**	**2.557**	**0.011**	0.195	0.121	1.606	0.108
Load × Interference	−19.990	17.632	−1.134	0.257	17.793	46.825	0.380	0.704					−0.280	0.254	−1.104	0.270
Load × F1	**41.370**	**12.105**	**3.418**	**0.001**	34.871	32.080	1.087	0.277					0.104	0.166	0.625	0.532
Load × F2	−14.957	12.082	−1.238	0.216	32.783	32.629	1.005	0.315					−0.082	0.165	−0.500	0.617
Interference × F1	13.276	12.225	1.086	0.277	25.272	34.244	0.738	0.461	0.041	0.217	0.187	0.851	0.200	0.179	1.119	0.263
Interference × F2	−9.419	12.107	−0.778	0.437	9.657	33.073	0.292	0.770	−0.248	0.215	−1.154	0.249	0.009	0.168	0.056	0.955
Load × Interference × F1	1.909	24.211	0.079	0.937	10.486	64.076	0.164	0.870					**1.090**	**0.331**	**3.292**	**0.001**
Load × Interference × F2	−12.231	24.167	−0.506	0.613	−1.981	65.276	−0.030	0.976					−0.438	0.329	−1.330	0.184

Table 9

Mixed-effects modeling results for all dependent measures – IQ partialed WMC (Sentence Span)

WMC	**Region 1**				**Region 2**				**Region 3**				**Region 4**			
IQ partialled	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***
(Intercept)	**806.73**	**39.27**	**20.54**	**0.000**	**749.36**	**33.90**	**22.10**	**0.000**	**633.55**	**23.75**	**26.67**	**0.000**	**634.14**	**27.37**	**23.17**	**0.000**
Load	**−80.76**	**17.96**	**−4.50**	**0.000**	**−101.04**	**17.41**	**−5.80**	**0.000**	**−77.61**	**13.26**	**−5.85**	**0.000**	**−92.21**	**12.08**	**−7.64**	**0.000**
Interference	−30.63	17.95	−1.71	0.088	−4.61	17.40	−0.26	0.791	−23.48	13.26	−1.77	0.077	9.91	12.07	0.82	0.412
ID	79.98	40.80	1.96	0.050	49.13	32.58	1.51	0.132	27.60	23.69	1.17	0.243	41.65	25.67	1.62	0.105
Load × Interference	−0.66	35.92	−0.02	0.985	14.00	34.81	0.40	0.688	48.67	26.52	1.84	0.066	**49.27**	**24.15**	**2.04**	**0.041**
Load × ID	**42.82**	**18.94**	**2.26**	**0.024**	−2.69	18.48	−0.15	0.884	27.41	14.03	1.95	0.051	11.05	12.83	0.86	0.389
Interference × ID	5.95	18.92	0.31	0.753	0.27	18.43	0.01	0.988	−14.75	14.00	−1.05	0.292	−1.64	12.79	−0.13	0.898
Load × Interference × ID	−10.52	37.90	−0.28	0.781	−17.73	36.96	−0.48	0.631	−40.94	28.08	−1.46	0.145	−20.66	25.68	−0.80	0.421

WMC	**Region 5**				**Region 6**				**Recall**				**Comp. Acc.**			
IQ partialled	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***Z***	***p***	**Est. (β)**	**SE**	***Z***	***p***
(Intercept)	**495.82**	**15.82**	**31.35**	**0.000**	**971.23**	**48.10**	**20.19**	**0.000**	**0.71**	**0.17**	**4.16**	**0.000**	**1.34**	**0.20**	**6.85**	**0.000**
Load	**−60.87**	**7.73**	**−7.88**	**0.000**	**−213.27**	**21.22**	**−10.05**	**0.000**					**−0.57**	**0.11**	**−5.29**	**0.000**
Interference	12.44	7.73	1.61	0.107	18.50	21.22	0.87	0.383	0.11	0.13	0.82	0.411	−0.19	0.11	−1.72	0.085
ID	26.93	16.53	1.63	0.103	75.41	44.00	1.71	0.087	**0.39**	**0.15**	**2.56**	**0.010**	**0.33**	**0.11**	**2.88**	**0.004**
Load × Interference	−10.29	15.46	−0.67	0.506	40.07	42.44	0.94	0.345					−0.40	0.22	−1.84	0.066
Load × ID	9.48	8.19	1.16	0.247	−0.42	22.41	−0.02	0.985					0.08	0.12	0.73	0.468
Interference × ID	−0.94	8.18	−0.11	0.909	37.11	22.33	1.66	0.097	0.06	0.15	0.38	0.707	**0.24**	**0.12**	**2.10**	**0.036**
Load × Interference × ID	−4.66	16.39	−0.28	0.776	−53.16	44.85	−1.19	0.236					0.15	0.23	0.66	0.511

Table 10

Interactions of Memory Load × Interference × Individual differences measures on Comprehension Accuracy after partialling out IQ (compare to Table 5).

Memory Load × Interference ×	Estimate	Std. Error	*Z*	*p*
1. receptive vocabulary	**0.485**	**0.210**	**2.302**	**.0213**
2. oral comprehension	0.342	0.206	1.656	.0977
3. Gates-MacGinitie	0.258	0.236	1.094	.2738
4. reading comprehension	0.183	0.225	0.812	.4170
5. SDRT	0.241	0.215	1.116	.2643
6. sentence span	0.151	0.230	0.657	.5114
7. reading - words	0.348	0.229	1.520	.1285
8. word identification	0.130	0.222	.5878	.5567
9. reading fluency	0.135	0.218	0.619	.5359

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 11

Mixed-effects modeling results for all dependent measures – IQ partialled receptive vocabulary (Peabody Picture Vocabulary Test). Significant effects in **bold**.

Receptive vocabulary	Region 1				Region 2				Region 3				Region 4			
IQ partialled	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*
Intercept	**809.820**	**40.340**	**20.077**	**0.000**	**751.650**	**33.800**	**22.237**	**0.000**	**634.870**	**23.790**	**26.692**	**0.000**	**635.950**	**27.320**	**23.276**	**0.000**
Load	**−79.520**	**17.950**	**−4.430**	**0.000**	**−100.450**	**17.360**	**−5.785**	**0.000**	**−75.960**	**13.230**	**−5.742**	**0.000**	**−91.560**	**12.050**	**−7.599**	**0.000**
Interference	−29.710	17.950	−1.655	0.098	−5.040	17.360	−0.290	0.772	−24.170	13.230	−1.827	0.068	10.080	12.050	0.837	0.403
ID	9.480	40.000	0.237	0.813	48.890	30.960	1.579	0.114	23.230	22.590	1.028	0.304	39.980	24.460	1.634	0.102
Load x Interference	−0.690	35.910	−0.019	0.985	13.140	34.730	0.378	0.705	46.510	26.460	1.758	0.079	**48.010**	**24.100**	**1.992**	**0.047**
Load x ID	12.310	18.060	0.682	0.495	**49.740**	**17.500**	**2.842**	**0.005**	**45.910**	**13.330**	**3.445**	**0.001**	**28.680**	**12.200**	**2.350**	**0.019**
Interference x ID	**41.100**	**18.150**	**2.264**	**0.024**	−2.810	17.750	−0.159	0.874	−7.020	13.480	−0.521	0.603	22.640	12.390	1.827	0.068
Load x Interference x ID	17.810	36.140	0.493	0.622	11.310	35.030	−0.323	0.747	−29.850	26.680	−1.119	0.263	−12.680	24.420	−0.519	0.604

Receptive vocabulary	Region 5 (critical region)				Region 6 (spillover region)				Recall (strict)*				Comprehension Accuracy			
IQ partialled	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*t*	*p*	Est. (β)	SE	*Z*	*p*	Est. (β)	SE	*Z*	*p*
Intercept	**497.070**	**15.830**	**31.405**	**0.000**	**974.660**	**47.990**	**20.308**	**0.000**	0.722	0.175	4.128	0.000	**1.350**	**0.200**	**6.905**	**0.000**
Load	**−60.120**	**7.700**	**−7.805**	**0.000**	**−213.040**	**21.220**	**−10.040**	**0.000**	-	-	-	-	**−0.570**	**0.110**	**−5.262**	**0.000**
Interference	12.400	7.700	1.609	0.108	19.660	21.220	0.926	0.354	0.111	0.135	0.822	0.411	**−0.210**	**0.110**	**−1.925**	**0.054**
ID	25.500	15.750	1.619	0.106	71.000	41.960	1.692	0.091	0.172	0.147	1.167	0.243	**0.290**	**0.110**	**2.668**	**0.008**
Load x Interference	−10.610	15.410	−0.689	0.491	38.770	42.450	0.913	0.361	-	-	-	-	0.350	0.220	−1.590	0.112
Load x ID	**26.510**	**7.710**	**3.437**	**0.001**	20.150	21.550	0.935	0.350	-	-	-	-	0.060	0.110	0.538	0.590
Interference x ID	8.230	7.740	1.063	0.288	18.730	21.900	0.856	0.392	0.069	0.136	0.503	0.615	0.020	0.110	−0.204	0.838
Load x Interference x ID	−0.834	15.440	−0.054	0.957	−0.330	43.110	0.008	0.994	-	-	-	-	**0.485**	**0.210**	**2.302**	**0.021**

* Absent values for Recall reflect the fact that the task was completed only in the Memory Load conditions, when the memory items were present.

Table 12

Interactions of Memory Load x Individual difference measures in the critical region after partialling out variance shared with IQ (Compare to Table 6).

Memory Load x	Estimate	Std. Error	*t*	*p*
1. SDRT	37.749	7.640	4.940	.0000
2. Rapid Naming	−39.965	8.090	−4.940	.0000
3. recognition- magazines	37.749	7.662	4.927	.0000
4. reading fluency	36.878	7.745	4.762	.0000
5. recognition – authors	29.395	7.821	3.759	.0002
6. receptive vocabulary	26.507	7.712	3.437	.0006
7. oral comprehension	19.746	7.741	2.551	.0107
8. speech comprehension	18.807	7.687	2.447	.0144
9. Gray Oral Reading Test	17.573	7.875	2.231	.0257
10. serial order memory	−17.070	7.875	−2.168	.0302
11. Gates MacGinitie	17.889	8.329	2.148	.0317

Table 13

Simple correlations among composite and non-composite measures (N=50).

Measure	1	2	3	4	5	6	7	8	9	10	11
1. phono. awareness	-	.22	.29	.00	.11	.08	.22	.13	−.02	−.15	.00
2. word/non-word reading	.42	-	.69	.47	.64	.38	.77	.60	.31	.03	.00
3. fluency	.44	.76	-	.56	.73	.41	.63	.47	.25	.08	.00
4. oral comprehension	.31	.65	.67	-	.64	.19	.56	.69	.32	.11	.00
5. reading comp.	.38	.76	.79	.79	-	.55	.63	.69	.32	.06	.00
6. print experience	.22	.47	.50	.34	.60	-	.18	.38	.34	−.12	.00
7. spelling	.37	.82	.70	.66	.71	.29	-	.50	.12	.08	.00
8. receptive vocabulary	.41	.74	.62	.83	.83	.48	.62	-	.36	−.05	.00
9. working memory	.25	.52	.44	.56	.56	.45	.32	.60	-	.05	.00
10. serial order memory	.10	.27	.26	.36	.33	.04	.25	.28	.29	-	.00
11. IQ	.47	.55	.46	.64	.65	.32	.42	.69	.56	.45	-

$|r| \geq .28$, $p < .05$, $|r| \geq .37$, $p < .01$, $|r| \geq .46$, $p < .001$, $|r| \geq .53$, $p < .0001$

Note: Measure 1: composite derived from CTOPP subtests for phonological awareness, phonological memory, rapid naming, rapid color naming; Measure 2: composite derived from Woodcock-Johnson word reading and word identification subtests; Measure 3: composite derived from Woodcock-Johnson reading fluency subtest and TOWRE reading words and nonwords; Measure 4: composite derived from Woodcock-Johnson listening comprehension and Peabody speech passage comprehension; Measure 5: composite derived from Woodcock-Johnson reading comprehension, Gray Oral Reading Test, Stanford Diagnostic Reading Test, Gates-MacGinitie reading comprehension, and Peabody print passage comprehension; Measure 6: composite derived from author and magazine recognition tests; Measure 7: composite derived from spelling words and nonwords. All other measures are non-composites.

Table 14

Mixed-effects modeling results for all dependent measures – composite analysis. Effects are reported for the 3-way interaction of Load x Interference x ID predictor for Regions 1–5 and Comprehension Accuracy, and for the 2-way interaction of Interference x ID predictor for Recall.

	Region 1				Region 2				Region 3				Region 4			
(Load x) Interference x ID predictor	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***
phonological processing	−15.464	46.224	−0.335	0.738	−0.213	46.138	−0.005	0.996	23.561	35.009	0.673	0.501	4.490	32.477	0.138	0.890
word/nonword reading	41.685	77.010	0.541	0.588	−143.932	76.778	−1.875	0.061	−8.636	58.589	−0.147	0.883	47.271	54.061	0.874	0.382
word/nonword fluency	−33.940	71.164	−0.477	0.633	−11.546	71.894	−0.161	0.872	−91.711	54.614	−1.679	0.093	−79.101	50.617	−1.563	0.118
listening pass/sent comp.	22.775	67.962	0.335	0.738	−56.185	68.208	−0.824	0.410	−48.072	51.765	−0.929	0.353	35.591	47.868	0.744	0.457
reading pass/sent comp.	−60.506	80.087	−0.756	0.450	34.469	80.388	0.429	0.668	87.585	60.821	1.440	0.150	65.620	56.531	1.161	0.246
print experience	21.231	55.067	0.386	0.700	66.702	55.549	1.201	0.230	−18.994	42.122	−0.451	0.652	−10.571	39.060	−0.271	0.787
word/nonword spelling	−45.742	70.768	−0.646	0.518	110.080	70.823	1.554	0.120	43.269	53.601	0.807	0.420	−0.962	49.626	−0.019	0.985
receptive vocab	75.298	75.703	0.995	0.320	25.196	75.092	0.336	0.737	5.990	57.113	0.105	0.916	−48.396	52.847	−0.916	0.360
WMC	−20.134	46.790	−0.430	0.667	−4.464	47.064	−0.095	0.924	−33.971	35.524	−0.956	0.339	−34.078	33.098	−1.030	0.303
Serial order memory	8.149	39.236	0.208	0.835	17.225	39.040	0.441	0.659	−14.583	29.821	−0.489	0.625	−11.974	27.282	−0.439	0.661
IQ	20.446	47.401	0.431	0.666	−22.734	47.621	−0.477	0.633	−25.271	36.224	−0.698	0.485	45.752	33.379	1.371	0.170

	Region 5 (critical region)				Region 6 (spillover region)				Recall (strict)*				Comprehension Accuracy			
(Load x) Interference x ID predictor	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***t***	***p***	**Est. (β)**	**SE**	***Z***	***p***	**Est. (β)**	**SE**	***Z***	***p***
phonological processing	2.737	19.863	0.138	0.890	64.111	52.772	1.215	0.224	−0.102	0.180	−0.566	0.571	0.219	0.294	0.743	0.457
word/nonword reading	8.965	33.165	0.27	0.787	−66.34	86.6	−0.766	0.444	−0.347	0.290	−1.198	0.231	−0.466	0.458	−1.016	0.310
word/nonword fluency	−36.04	30.773	−1.171	0.242	−91.201	82.35	−1.107	0.268	−0.162	0.277	−0.585	0.559	0.313	0.462	0.677	0.498
listening pass/sent comp.	−2.621	29.031	−0.09	0.928	−72.04	77.457	−0.93	0.352	−0.026	0.252	−0.103	0.918	−0.288	0.438	−0.658	0.511
reading pass/sent comp.	56.103	34.958	1.605	0.109	17.935	92.351	0.194	0.846	−0.096	0.309	−0.309	0.757	−0.51	0.522	−0.978	0.328
print experience	−36.894	23.863	−1.546	0.122	−69.469	63.984	−1.086	0.278	−0.075	0.213	−0.352	0.725	0.005	0.358	0.013	0.989
word/nonword spelling	−12.483	30.768	−0.406	0.685	123.568	80.318	1.538	0.124	0.344	0.262	1.311	0.190	0.141	0.437	0.323	0.747
receptive vocab	5.018	32.27	0.155	0.876	**172.127**	**84.928**	**2.027**	**0.043**	0.274	0.289	0.949	0.343	**0.995**	**0.468**	**2.125**	**0.034**
WMC	−0.729	20.437	−0.036	0.972	−21.142	53.58	−0.395	0.693	0.152	0.182	0.835	0.404	0.137	0.285	0.48	0.631
Serial order memory	1.201	16.78	0.072	0.943	**97.536**	**44.693**	**2.182**	**0.029**	−0.139	0.156	−0.894	0.371	0.024	0.254	0.096	0.924
IQ	−13.007	20.516	−0.634	0.526	−41.948	54.775	−0.766	0.444	−0.201	0.203	−0.991	0.322	**0.716**	**0.307**	**2.334**	**0.020**

Table 15

All significant main effects and lower order interactions observed in the composite analysis, reported by dependent measure. Note that interactions with Interference alone are uninterpretable, as Interference depends on the presence of the Memory Load and the Interference conditions are identical prior to Region 5. These uninterpretable interactions are presented in italics.

Region 1					**Region 2**				
Factor	**Est**	**SE**	**t**	**p**	**Factor**	**Est**	**SE**	**t**	**p**
(Intercept)	864.124	38.750	22.300	0.000	(Intercept)	800.618	32.696	24.487	0.000
Load	−85.628	21.383	−4.005	0.000	Load	−76.761	21.409	−3.585	0.000
Receptive Vocab	−152.588	66.386	−2.298	0.022	IQ	152.516	33.435	4.562	0.000
IQ	153.128	42.351	3.616	0.000	Phonological processing	67.615	32.325	2.092	0.036
Word reading	146.604	67.866	2.160	0.031	Fluency	−160.576	49.766	−3.227	0.001
Fluency	−204.188	62.996	−3.241	0.001	Load x Print Experience	102.684	27.779	3.696	0.000
Load x Print Exp.	98.973	27.535	3.594	0.000	Load x WMC	−57.854	23.537	−2.458	0.014
					Load x IQ	63.984	23.806	2.688	0.007
					Interference x Print Experience	*−61.351*	*27.531*	*−2.228*	*0.026*
					Interference x Serial Order Mem	*−41.746*	*19.601*	*−2.130*	*0.033*

Region 3					**Region 4**				
Factor	**Est**	**SE**	**t**	**p**	**Factor**	**Est**	**SE**	**t**	**p**
(Intercept)	666.386	22.240	29.964	0.000	(Intercept)	674.150	26.602	25.342	0.000
Load	−66.049	16.291	−4.054	0.000	Load	−84.418	15.011	−5.624	0.000
IQ	99.581	23.155	4.301	0.000	IQ	124.291	25.546	4.865	0.000
Fluency	−133.847	34.470	−3.883	0.000	Fluency	−133.641	38.019	−3.515	0.000
Load x Print Exp.	50.145	21.073	2.380	0.017	Load x Serial Order Memory	−29.840	13.639	−2.188	0.029
Load x IQ	66.960	18.115	3.696	0.000	Load x Fluency	90.754	25.319	3.584	0.000
Load x Word Reading	−58.802	29.281	−2.008	0.045	*Interference x Receptive Vocabulary*	*67.495*	*26.859*	*2.513*	*0.012*
Load x Fluency	74.907	27.283	2.746	0.006					

Region 5					**Region 6**				
Factor	**Est**	**SE**	**t**	**p**	**Factor**	**Est**	**SE**	**t**	**p**
(Intercept)	516.608	14.395	35.889	0.000	(Intercept)	997.613	44.675	22.330	0.000

Region 5					Region 6				
Factor	**Est**	**SE**	**t**	**p**	**Factor**	**Est**	**SE**	**t**	**p**
Load	−54.066	9.282	−5.825	0.000	Load	−222.609	24.452	−9.104	0.000
IQ	66.817	15.909	4.200	0.000	IQ	190.075	43.013	4.419	0.000
Reading Comprehension	54.118	26.788	2.020	0.043	Fluency	−266.251	63.970	−4.162	0.000
Fluency	−99.225	23.670	−4.192	0.000	Load x Print Experience	108.681	31.983	3.398	0.001
Load x Print Exp.	47.789	11.932	4.005	0.000	Load x Phonological Processing	−67.349	26.388	−2.552	0.011
					Load x Fluency	104.483	41.291	2.530	0.011
					Interference x Fluency	*88.588*	*40.753*	*2.174*	*0.030*
					Load x Interf x Receptive Vocab.	**172.127**	**84.928**	**2.027**	**0.043**
					Load x Interf x Serial Order Mem.	**97.536**	**44.693**	**2.182**	**0.029**

Recall (strict)*					Comprehension Accuracy				
Factor	**Est**	**SE**	**t**	**p**	**Factor**	**Est**	**SE**	**t**	**p**
(Intercept)	0.774	0.163	4.736	0.000	(Intercept)	1.435	0.188	7.613	0.000
WMC	0.424	0.138	3.073	0.002	Load	−0.595	0.146	−4.076	0.000
IQ	0.686	0.146	4.694	0.000	WMC	0.265	0.090	2.945	0.003
					IQ	0.607	0.095	6.386	0.000
					Interference x WMC	0.332	0.146	2.270	0.023
					Load x Interference x Receptive Vocabulary	**0.995**	**0.468**	**2.125**	**0.034**
					Load x Interference x IQ	**0.716**	**0.307**	**2.334**	**0.020**