Psychology and Aging
2004, Vol. 19, No. 4, 601–616

0882-7974/04/$12.00 DOI: 10.1037/0882-7974.19.4.601

The Relationship Between Age, Verbal Working Memory, and Language Comprehension

Gayle DeDe
Boston University

David Caplan
Massachusetts General Hospital

Karen Kemtes
University of Nevada, Las Vegas

Gloria Waters
Boston University

A structural modeling approach was used to examine the relationships between age, verbal working memory (vWM), and 3 types of language measures: online syntactic processing, sentence comprehension, and text comprehension. The best-fit model for the online-processing measure revealed a direct effect of age on online sentence processing, but no effect mediated through vWM. The best-fit models for sentence and text comprehension included an effect of age mediated through vWM and no direct effect of age. These results indicate that the relationship among age, vWM, and comprehension differs depending on the measure of language processing and support the view that individual differences in vWM do not affect individuals' online syntactic processing.

It is well documented that older adults have significantly reduced working memory (WM) spans relative to younger adults (see Carpenter, Miyake, & Just, 1994, for a review). This decline in WM has been interpreted as a reduction in the processing capacity that subserves cognitive tasks (Baddeley, 1986; Engle, Tuholski, Laughlin, & Conway, 1999). The focus of the present study is the relationship between age, verbal working memory (vWM), and performance in one particular cognitive domain, language processing.

vWM capacity has been related to performance on various measures of language processing (e.g., Daneman & Carpenter, 1980; Martin & Romani, 1994; Norman, Kemper, Kynette, Cheung, & Anagnopoulos, 1991; Tun, Wingfield, & Stine, 1991), suggesting that age-related changes in language processing may be attributable to declines in vWM. Some researchers (e.g., Just & Carpenter, 1992) have argued that there is a general vWM system that supports span performance, language processing, and other verbally mediated cognitive functions. Caplan and Waters (1999a) termed this the "single processing resource theory." Caplan and Waters (e.g., Caplan & Waters, 1990, 1999a, 1999b; Waters & Caplan, 1996) have argued for the "separate language interpretation resource theory," which states that there is a domain-specific vWM system that supports online language comprehension. In this view, traditional span measures, such as digit span and the Daneman and Carpenter (1980) reading span task, are not associated with all measures of language processing. According to the separate resource theory, vWM span measures require controlled, conscious verbal processing and are therefore predicted to be related to global measures of language processing such as text comprehension and offline measures of comprehension such as plausibility judgments. However, language tasks that tap unconscious, obligatory, online processing are not predicted to be related to vWM span measures. This contrasts with the single resource theory, which predicts that vWM span measures should be related to all measures of language processing.

There is a substantial body of evidence supporting the existence of a relationship between vWM span and text comprehension. Much of the work in this area has focused on text recall. Such studies have suggested that vWM is predictive of recall of both narrative and expository texts (Hartley, 1986; Hultsch, Hertzog, & Dixon, 1984; Stine & Wingfield, 1990; Stine-Morrow, Loveless, & Soederberg, 1996; Tun et al., 1991; Wingfield & Littlefield, 1995). Daneman and Merikle (1996) completed a meta-analysis of 75 studies that investigated the relationship among the Daneman and Carpenter (1980) reading span task and various measures of language comprehension. They found at least moderately high correlations between vWM and global measures such as the Nelson-Denny Reading Comprehension subtest (Nelson & Denny, 1960) and the Wechsler Adult Intelligence Scale Vocabulary subtest (Wechsler, 1981). Older adults also have been demonstrated to perform more poorly than younger adults on measures of sentence comprehension (Emery, 1985; Feier & Gerstman, 1980; Kemper, 1988; Norman et al., 1991), and older adults' vWM has also been shown to correlate with performance on sentence com-

Gayle DeDe and Gloria Waters, Communication Disorders Program, Department of Health Sciences, Boston University; David Caplan, Neuropsychology Laboratory, Department of Neurology, Massachusetts General Hospital; Karen Kemtes, Department of Psychology, University of Nevada, Las Vegas.

We thank Tim Brown for his helpful comments on a draft of the manuscript. This work was supported by Grant AG09661 from the National Institute on Aging to Gloria Waters.

Correspondence concerning this article should be addressed to Gloria Waters, Department of Health Sciences, Boston University, 635 Commonwealth Avenue, Boston, MA 02215. E-mail: gwaters@bu.edu

prehension measures (Davis & Ball, 1989; Kemper, 1986; Kemtes & Kemper, 1997; Light, 1990; Obler, Fein, Nicholas, & Albert, 1991).

Evidence of a relationship among vWM and sentence comprehension and text comprehension/recall is consistent with both the single and separate resource theories, because both theories predict that vWM capacity will be correlated with offline measures of language processing. However, it is not clear whether this relationship extends to online linguistic-processing abilities such as the initial assignment of syntactic structure and semantic meaning. Such abilities are best measured by online methods, which are sensitive to effects of syntactic complexity as they occur in capacity-demanding portions of sentences (e.g., Cohen & Mehler, 1996; Ford, 1983; Frauenfelder, Segui, & Mehler, 1980; see MacDonald, Pearlmutter, & Seidenberg, 1994, for review). As reviewed above, the single and separate resource theories make different predictions about the relationship of vWM to online measures of syntactic processing.

A number of studies have examined the relationship between vWM and online syntactic processing in younger adults. Evidence that vWM influences online processing would come from finding that online performance is more affected by local increases in processing load for individuals with lower vWM capacity (see Caplan and Waters, 1999a, for a discussion). In general, the literature suggests that this is not the case. King and Just (1991) reported that young adults with lower vWM spans had longer self-paced reading times for sentences than their peers with higher vWM spans. However, these differences were observed at all points in the sentences rather than only at the capacity-demanding phrases (see Waters & Caplan, 1996, and Just, Carpenter, & Keller, 1996, for discussions). MacDonald, Just, and Carpenter (1992) found that high-span individuals spent more time reading the final word of sentences containing temporary syntactic ambiguities than did low-span individuals. This was interpreted as evidence that high-span individuals maintained multiple possible resolutions of the ambiguity, whereas low-span individuals maintained only one possible resolution. If the high-span group was paying a cost for maintaining multiple interpretations of the sentences, the single resource theory would predict that the span groups' reading times would differ in the region that included the syntactic ambiguity (see Waters & Caplan, 1996, and Just et al., 1996, for discussions). However, the span groups' reading times did not differ in that capacity-demanding region of the sentence. Waters and Caplan (2003a) also presented evidence that high- and low-span young adults did not differ in their online syntactic processing of the capacity-demanding portions of syntactically complex sentences. In general, the results of these studies support the view that individuals with lower vWM spans process sentences more slowly overall and differ from high-span individuals on offline measures of sentence comprehension, but that high- and low-span individuals do not significantly differ in online syntactic processing.

Several studies have shown that online syntactic processing does not significantly differ for older and younger adults. Baum and colleagues (Baum, 1991; Waldstein & Baum, 1992) used a word-monitoring task to measure older and younger adults' online detection of ungrammatical sentences. They found no evidence that older participants were more reliant on sentential context or that they were less sensitive to ungrammaticalities. In their studies of online text processing, Stine-Morrow and colleagues (Stine, 1990; Stine-Morrow et al., 1996; Stine-Morrow, Soederberg-Miller, & Leno, 2001) showed that for both younger and older adults, text recall was greater for participants who spent more time reading syntactically complex sentences. The absence of age effects related to syntactic complexity suggests that the older adults were not differentially affected by that factor (Stine-Morrow et al., 1996).

Kemtes and Kemper (1997) investigated how vWM span and age interacted with online processing of syntactically ambiguous sentences, and they found no evidence that age or vWM span had a differential effect on word-by-word reading times of ambiguous sentences. Kemtes (1999) used the "stop making sense paradigm" to assess the relationship between vWM and online grammaticality judgments in older and younger adults. In this task, participants are asked to press a button when a sentence becomes grammatically incorrect. Kemtes found that older adults had smaller vWM spans and made slower judgments but that the speed of their judgments was not differentially affected by syntactic complexity. Waters and Caplan (2001, in press) used the auditory moving windows paradigm to demonstrate that older adults were slower and less accurate than younger adults in listening to and comprehending syntactically complex sentences. However, these differences were not disproportionately larger at capacity-demanding phrases of the sentences. All of these studies support the view that vWM does not determine online syntactic-processing ability.

Other studies have been interpreted as providing support for a role of vWM in older adults' online syntactic processing. One study that used the cross-modal lexical priming paradigm demonstrated that older adults' reactivation of traces at gap sites was influenced by the number of intervening words, suggesting that storage capacity played a role in their syntactic processing (Zurif, Swinney, Prather, Wingfield, & Brownell, 1995). However, Zurif et al. (1995) did not test younger adults on the same materials, and the vWM capacity of the older participants was not directly measured, making it difficult to assess the precise role of vWM in the older adults' performance. Kemper and Kemtes (1999) used a word-by-word reading paradigm to assess older and younger adults' online processing. Older adults with low vWM spans had longer reading times for syntactically ambiguous segments of Wh— questions than older adults with high vWM spans. This effect was not observed in younger adults, therefore Kemper and Kemtes (1999) concluded that older and younger adults were differentially affected by syntactic ambiguity. However, the same pattern was observed at unambiguous segments of the sentence. This suggests that older adults with low vWM capacities may have longer reading times overall for Wh— questions, but that they do not show specific effects at capacity-demanding portions of sentences. Another study compared older and younger adults' word-by-word reading times for subject- and object-relativized sentences (Stine-Morrow, Ryan, & Leonard, 2000). They found that older adults were less accurate in comprehension of the object-relativized sentences and failed to show an increase in reading time observed in the younger adults at capacity-demanding portions of the syntactically complex sentences. Although they concluded that older adults were less sensitive to the demands of thematic role assignment in the more complex sentences, they did not demonstrate that there was a significant difference between the older and younger adults' reading times exclusively at the capacity-demanding portions of the syntactically complex sentences (i.e., there was no three-way interaction).

In the present study, a structural equation modeling (SEM) approach was used to study the interrelationships among age, vWM, and several measures of language processing. A small number of studies have used regression and SEM approaches to study the relationships among these constructs. Although these studies do not use online-processing measures, they illustrate the utility of the SEM approach and have at least suggestive value regarding the issues that we raise.

Kwong See and Ryan (1995) used a hierarchical regression approach to study the relationship between vWM, inhibition, speed of processing, and language processing. The language-processing measures included sentence comprehension, text comprehension, and recall. A composite score of language processing was calculated by averaging standardized scores on the language measures. Kwong See and Ryan found that when their measures of processing speed and inhibition were entered into the regression equation before their measure of vWM, vWM no longer accounted for a significant portion of the age-related variance in language processing. This indicates that vWM's effect on language processing may be mediated by other factors, as is the case for other cognitive measures.

One limitation of these results was that vWM, speed, and inhibition were represented by composite scores (see Kwong See & Ryan, 1995, and Van der Linden et al., 1999, for discussions). Creating a composite score may have decreased the sensitivity of the model because individual measures were not allowed to vary in their contribution to the composite score. The SEM approach permits use of multiple indicators of one latent variable (e.g., several measures of vWM), which is advantageous because multiple indicators may provide a more reliable measure of the underlying construct.

Van der Linden et al. (1999) used an SEM approach to examine the relationship between vWM, inhibition, speed of processing, and language processing, and they obtained results that differed from Kwong See and Ryan (1995). The measures of language processing included text comprehension, sentence and story recall, and the California Verbal Learning Test (Delis, Freeland, Kaplan, & Ober, 1987). Van der Linden et al. demonstrated that the influence of age on the language and verbal memory construct was mediated by the constructs of processing speed, inhibition, and vWM. In contrast to Kwong See and Ryan (1995), they found that the contributions of speed and inhibition on language processing were indirect and mediated by vWM, suggesting that vWM plays a critical role in differences in language performance observed in younger and older adults. In addition to the difference in methodologies, Van der Linden et al. suggested that the contrasting results could reflect differences in the types of language-processing measures used, with Van der Linden et al. incorporating more measures of delayed comprehension and Kwong See and Ryan (1995) emphasizing immediate comprehension. Although not relevant to online processing, these results show that vWM may affect different measures of language processing differently.

Kemper and Sumner (2001) used an SEM approach to investigate the relationship between measures of vWM, vocabulary, and processing fluency in younger and older adults. Processing fluency included comprehension tasks (reading rate) and verbal production measures such as mean length of utterance, propositional density, and categorization and retrieval (e.g., category and initial letter fluency). The factor structures for the constructs of vWM, vocabulary, and processing fluency were similar for the two age groups. All of the factors were intercorrelated in the group of younger adults, but in the older adults the processing factor did not covary with the vWM factor. This suggested that the relationship between vWM and some measures of language processing becomes less important as people age. However, other language-processing measures (e.g., a measure of reading comprehension and an index of grammatical complexity in verbal production) loaded on the vWM factor along with reading span and backward digit span in both older and younger adults. This also demonstrates that the relationship between vWM and language processing varies as a function of the measure of language processing.

Each of these studies (Kemper & Sumner, 2001; Kwong See & Ryan, 1995; Van der Linden et al., 1999) has attempted to describe the relationships of various cognitive processes, including vWM, with age and various measures of language processing. However, all of these studies group different types of measures of language processing. It is possible that the contributions of factors such as age and vWM differ for different types of language processing, as has been observed in studies that did not use an SEM approach.

We used an individual differences approach to examine the relationship between age, vWM, and three aspects of language processing: online syntactic processing, sentence comprehension, and text comprehension. Note that the data have been previously published in several articles that used other analytic tools (Waters & Caplan, 2001, 2002, 2003a, 2003b, in press). The present analysis was undertaken to provide a more sensitive test of the separate resource theory by using a structural modeling approach.

The relationship among these variables was modeled separately for each aspect of language processing, resulting in three sets of SEMs. We predicted that measures of vWM would mediate performance on measures of sentence and text comprehension but not online sentence processing.

Method

Participants

Participants were drawn from a number of studies conducted in our laboratories at Boston University and Massachusetts General Hospital over the past several years. Participants' entries into the structural models were dependent on the tasks they completed. Hence, different, but partially overlapping, sets of participants were entered into each set of models. Participant characteristics are presented for older (60–90 years of age) and younger (18–30 years of age) adults separately, as well as for the group as a whole. For models investigating the relationship between vWM and online syntactic processing, there were a total of 215 participants ($n = 72$ younger adults, $n = 143$ older adults; see Table 1, under "Models of vWM and online sentence processing," for mean age and education levels). For models of vWM and sentence-level comprehension, there were a total of 216 participants ($n = 72$ younger adults, $n = 144$ older adults; see Table 1, under "Models of vWM and offline sentence comprehension"). Finally, for models of vWM and text comprehension, there were a total of 248 participants ($n = 118$ younger adults, $n = 130$ older adults; see Table 1, under "Models of vWM and text comprehension").

Younger adults were university students recruited through electronic mailing and posted announcements. Older adults were community dwelling and recruited through the Harvard Cooperative Program on Aging as well as through ads posted in churches, synagogues, and newsletters to seniors.

Table 1
Participant Characteristics for All Models

Variable	Younger adults M (SD)	Older adults M (SD)	Group comparison t value	All participants M (SD)	All participants Multiple R^2
Models of vWM and online sentence processing					
N	72	143	—	215	—
Age (years)	21.3 (2.6)	72.9 (7.4)	57.3**	55.6 (25.2)	—
Education (years)	14.5 (1.5)	14.7 (3.1)	0.5	14.7 (2.6)	—
Alphabet span	4.6 (.8)	3.9 (.7)	6.8**	4.2 (.8)	.54
Subtract 2 span	5.9 (1.0)	5.1 (1.1)	5.6**	5.4 (1.1)	.55
Sentence span	3.4 (1.5)	2.4 (1.1)	5.6**	2.7 (1.4)	.48
COV − CSN2	187.7 (330.6)	434.3 (393.8)	4.6**	351.7 (390.9)	.55
SOV1 − OSN2	114.3 (235.2)	173.2 (293.8)	1.6	153.5 (276.4)	.61
SOV2 − OSV1	130.6 (144.9)	99.8 (190.5)	0.6	110.1 (176.8)	.43
Models of vWM and offline sentence comprehension					
N	72	144	—	216	—
Age (years)	21.3 (2.6)	72.7 (7.4)	57.3**	55.6 (25.1)	—
Education (years)	14.7 (1.5)	14.9 (3.1)	0.5	14.8 (2.5)	—
Alphabet span	4.6 (.8)	3.9 (.7)	6.4**	4.1 (.8)	.59
Subtract 2 span	5.9 (1.0)	5.1 (1.1)	5.4**	5.4 (1.1)	.57
Sentence span	3.4 (1.5)	2.4 (1.1)	5.4**	2.7 (1.4)	.48
Acc CO	91.1 (10.0)	86.6 (10.4)	3.0*	88.1 (10.5)	.73
Acc CS	97.7 (2.8)	95.8 (3.8)	3.8*	96.4 (3.6)	.56
Acc SO	87.3 (7.9)	84.7 (7.5)	2.3*	85.6 (7.7)	.69
Models of vWM and text comprehension					
N	118	130	—	248	—
Age (years)	21.0 (2.5)	72.7 (7.2)	74.7**	48.1 (26.5)	—
Education (years)	14.7 (1.5)	14.4 (3.0)	0.7	14.5 (2.4)	—
Alphabet span	4.7 (0.8)	3.9 (0.8)	8.1*	4.2 (0.9)	.60
Subtract 2 span	5.9 (1.0)	5.1 (1.1)	5.5*	5.5 (1.1)	.56
Sentence span	3.6 (1.6)	2.3 (1.1)	7.3*	2.9 (1.5)	.52
Nelson-Denny	63.3 (25.6)	52.5 (29.9)	3.0*	57.6 (28.4)	.39

Note. Dashes indicate that the values were not calculated. COV − CSN2 = Listening times for verb (V) of the cleft object (CO) minus the 2nd noun phrase (N2) of the cleft subject (CS) sentences; SOV1 − OSN2 = the embedded verb (V1) of the subject object (SO) minus N2 of the object subject (OS) sentences; SOV2 − OSV1 = the main verb (V2) of SO minus the main verb of OS sentences. Accuracy (Acc) CO, CS, and SO refer to the accuracy of plausibility judgments on the offline sentence comprehension measure. Nelson-Denny = Nelson-Denny Reading Comprehension Test; vWM = verbal working memory.
* $p < .05$. ** $p < .001$.

All participants were paid for their participation. The participants were native speakers of English, denied a past medical history of neurological disease and language impairment, had completed at least a high school education, and had normal pure tone audiometry.

Experimental Procedure

All participants were tested individually, either in their homes or in the laboratory setting. Testing sessions lasted 1 to 1.5 hr. The number of testing sessions depended on the protocol that each participant was completing and ranged from three to eight sessions over the course of several weeks. Sessions were typically held once per week. Data included in the present study were collected within the first three sessions of any given protocol.

Measures

Background Testing

All older participants were pretested on a battery of neuropsychological tests to rule out the presence of dementia. These background measures included the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), the Logical Memory I and II (LMI, LMII) subtests of the Wechsler Memory scale (Wechsler, 1987), and the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983).

Measures of vWM

The three measures of vWM were selected on the basis of prior work suggesting that these tasks have good internal consistency (Waters & Caplan, 2003b). Additionally, these tasks had the best test–retest reliability in a battery of seven vWM measures (ranging from .65 to .76), and a composite score based on these measures had good test–retest reliability (.85; Waters & Caplan, 2003b).

For each measure, testing began at Span Size 2. There were five trials at each span size. *Span* was defined as the longest span-size length at which the participant was able to correctly recall three out of the five trials. An additional .5 point was awarded if the participant was able to recall two of the five trials at the next span-size level. All trials at all list lengths were administered to all participants.

Alphabet span. The participant was required to repeat an auditorily presented list of monosyllabic, unrelated words after rearranging them in alphabetical order. The internal consistency of this measure has ranged from .84 to .86 (Waters & Caplan, 2003b).

Subtract 2 span. The participant was required to repeat a list of single digits (2–9) in the same order in which they were presented but after subtracting 2 from each. Internal consistency has been demonstrated to be .82 (Waters & Caplan, 2003b).

Sentence span. This is a variant of the Daneman and Carpenter (1980) sentence span task developed by Waters and Caplan (1996). Internal consistency for this task has been demonstrated to range from .92 to .95 (Waters & Caplan, 1996). Participants were presented with syntactically complex (subject object relatives) sentences on a computer screen. They read the sentences and were required to make plausibility judgments after each sentence by pressing a button on a button box interfaced with the computer. Half of the items were plausible, and half were implausible. Both reaction time to make the decision and accuracy of the judgment were recorded by the computer. The accuracy and reaction time data were not used in the present study. After reading and making judgments about the last sentences within a span size, an asterisk appeared on the computer screen, indicating that the participants had to recall the last word of each sentence in the correct serial order. The final word-recall measure was used as the indicator of span for each participant.

Measures of Language Processing

Online syntactic processing. Online syntactic processing was measured using the auditory moving windows paradigm (AMW; Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996; Waters & Caplan, 1996, 2001). For each trial of the task, participants listened to sentences that had been digitized and segmented into phrases. Participants paced their way through the sentence by pressing a button interfaced with a computer to hear each successive phrase of the sentence and then made a plausibility judgment about the sentence. Reaction time for each button press and response time and accuracy for the plausibility judgments were recorded by PsyScope experimental software (Cohen, MacWhinney, Flatt, & Provost, 1993).

The sentences were recorded by a male speaker and digitized at a sampling rate of 40 kHz with a 16-bit quantization by SoundEdit software (Dunn, 1994). The waveform files were edited with SoundEdit to place an inaudible marker at each phrase boundary. The markers were placed in areas of the waveform such that intelligibility was maximized, determined by perceptual analysis. In addition, visual and perceptual inspections were used to place tags in areas of low-signal amplitude whenever possible to make smooth phrase-to-phrase transitions. Waveforms were then entered into PsyScope (Cohen et al., 1993) and played out over headphones.

The target stimuli were 104 semantically acceptable and 104 semantically unacceptable sentences divided equally among 4 sentence types. Table 2 presents examples of the stimuli phrase by phrase (indicated by slashes ["/"] in the table). The stimuli were two pairs of sentence types (cleft subject [CS] vs. cleft object [CO], and object subject [OS] vs. subject object [SO]), with one member of each pair consisting of a syntactically more complex object-relativized sentence (CO and SO) and the second consisting of a relatively simpler subject-relativized sentence (CS and OS). In the object-relativized sentences, the head noun of the relative clause must be retained over more words than in subject-relativized sentences. Also, in object-relativized sentences, the parser assigns two thematic roles at the embedded verb compared with one in subject-relativized sentences. Both the memory and computational demands are greater in the object-relativized sentences (Caplan, Hildebrandt, & Waters, 1994; Gibson, 1998). These effects can also spill over to the main verb of SO relativized sentences (King & Just, 1991). Some participants completed a version of the task that contained an additional 52 sentences, which were not analyzed in the present study.

The listening times for the phrases are the measure of online syntactic processing (Waters & Caplan, 2001, 2003a, in press). Because the response time for each segment included the duration of the segment in addition to the time required by the participant to comprehend it, listening times were calculated by subtracting the segment's duration from the response time.[1] Segment duration was calculated separately for each instance of the lexical items in the sentences. The effect of word frequency on listening times was controlled by regressing listening times against word frequency to calculate the predicted listening times for each word frequency regardless of position and by then subtracting each listening time from the average to create residual listening times. Word frequency for content words only was determined by using Francis and Kucera's (1982) frequency analysis. This technique was originally proposed by Ferreira and Clifton (1986) to control for effects of letter length and word frequency in a word-by-word reading experiment. Residual listening times greater and less than 3 *SD*s from the mean for each condition for each participant were removed as outliers.

Difference scores were calculated as a measure of the difference in processing complexity at segments of interest in more and less complex

[1] Two potential issues with our stimuli are whether factors such as the part of speech and uniqueness points of the lexical items were adequately controlled. With respect to uniqueness points, an anonymous reviewer pointed out that "listeners actively process the speech stream as it is presented (e.g., Marslen-Wilson's demonstrations that word recognition occurs within 300 ms), so subtracting out presentation time is probably producing underestimates of processing time; furthermore . . . the magnitude of this error probably covaries with length." Marslen-Wilson (e.g., see Marslen-Wilson & Welsch, 1978) has identified two factors that allow participants to recognize spoken words before their acoustic end: (a) Context allows people to make intelligent guesses regarding word identity; and (b) long words can become phonologically unique before their ends, allowing people to identify them bottom-up. The first factor, context, is noncontributory in our materials. Particular words are not more predictable in one sentence type than another on the basis of the meaning of the sentences or the probability of occurrence of particular words in one or another sentence type. To see if differences in uniqueness points of critical words across sentence types could have produced effects that we take to be syntactic, we examined the distribution of uniqueness points in the critical words in each sentence pair. Distance of uniqueness points from the ends of words, calculated in terms of number of phonemes and syllables, were determined by identifying the point at which each critical word diverged from all other lexical items listed in the Medical Research Council Psycholinguistic database (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). Uniqueness points were calculated from the ends of words because the distance from the uniqueness point to the end of the word is relevant to the question of how much of the corrected listening time might be due to syntactic processing. Corrected listening times are the times between button presses minus segment duration. Consider two button press times that are the same (to facilitate this discussion, assume they are longer than the segment durations) in two words with different uniqueness points—one at the end of the word and one 100 ms before the end of the word. The button press for the word with the earlier uniqueness point might reflect more time spent in syntactic processing. For these reasons, what is most important is where lexical items become unique relative to the button press. Uniqueness points were systematically further from the ends of the critical words in complex sentences (CO and SO) than in the critical words in simple sentences (CS and OS). Accordingly, our measurement may have underestimated, not overestimated, the magnitude of the effect of syntax.

Table 2
Examples of Sentence Stimuli

Sentence type	Sentence
	Plausible sentences
Cleft subject	**N2**
Simple	It was/the food/that nourished/the child.
Cleft object	**V**
Complex	It was/the woman/that the toy/amazed.
Object subject	**V1 N2**
Simple	The father/read/the book/that terrified/the child.
Subject object	**V1 V2**
Complex	The man/that the fire/injured/called/the doctor.
	Implausible sentences
Cleft subject	
Simple	It was/the car/that drove/the woman.
Cleft object	
Complex	It was/the coffee/that the secretary/disappointed.
Object subject	
Simple	The girl/drank/the boy/that entered/the hospital.
Subject object	
Complex	The secretary/that the camera/met/drove/the car.

Note. Slashes depict phrase segments. N2 = 2nd noun phrase; V1 = 1st verb phrase; V2 = 2nd verb phrase.

sentences.[2] Segments of interest were chosen on the basis of previous work by Waters and Caplan (2001, 2003a, in press). These difference scores included (a) the verb (V) of CO sentences minus the second noun phrase (N2) of CS sentences (COV − CSN2), (b) the embedded verb (V1) of SO sentences minus the second noun phrase (N2) of OS sentences (SOV1 − OSN2), and (c) the main verb (V2) of SO sentences minus the main verb (V1) of OS sentences (SOV2 − OSV1).[3] The first and second difference scores compared the verb of the more complex sentence type with a noun in clause- or sentence-final position. These comparisons control for effects of sentence position, which have been found to affect reading times in previous work (Balogh, Zurif, Prather, Swinney, & Finkel, 1998). The third comparison directly compared the processing load at the main verbs of the sentences but confounds serial order and complexity because the verbs differ with respect to their serial position in the sentence.

Offline sentence comprehension. The three measures of sentence level comprehension were accuracy scores from the plausibility judgments of CO, CS, and SO sentences in the AMW task. Accuracy scores for OS sentences from the AMW task were not included in the models because they showed only weak correlations with the other sentence types included. The measure of sentence comprehension for each sentence type was the percentage correct across all exemplars of that sentence type.

Text comprehension. Participants completed the reading comprehension subtest of the Nelson-Denny Reading Comprehension Test—Form A (Nelson & Denny, 1960) as a global measure of language comprehension. The Nelson-Denny Reading Comprehension Test is a timed test in which participants are required to read paragraphs and answer comprehension questions about them. Measures of both reading rate and comprehension level can be derived from this test. This test has been widely used as a measure of language comprehension in studies that investigate the relationship between vWM capacity and language comprehension (see Daneman & Merikle, 1996, for a review).

Results

Results for each set of models are presented separately because each set contained a different group of participants.

vWM and Online Sentence Processing

Table 3 presents the mean and standard deviation of the performance on the background measures of the older adults. None of the participants consistently performed outside of the normal range, indicating an absence of dementia in the older adults. Performance on the background measures was similar for the groups of older adults used in the three sets of analyses. Table 1 (under "Models of vWM and text comprehension") summarizes the mean age and education levels of the groups and presents the descriptive statistics of the manifest variables for this set of models for the younger and older adults separately as well as for the group as a whole. Multiple R^2 was used as a measure of reliability for all of the measures. Multiple R^2 was calculated by regressing each variable onto the other variables in the data set, with the assumption that if the task is reliable and covaries to a degree with other tasks that the multiple R^2 will be high. All of the multiple R^2 values for the manifest variables associated with this set of analyses were greater than .4, indicating that the measures were moderately intercorrelated. The groups did not significantly differ with respect to years of education. The younger adults

[2] Difference scores have been argued to have questionable reliability and validity (Lord, 1956). However, others have argued that difference scores are reliable when there is appreciable difference in true change for individual scores (Rogosa & Willett, 1983). More recently, Zimmerman and Williams (1998) demonstrated that gain scores can be reliable and that their reliability coefficients are intermediate between those of scores on which they are based. In the present study, use of difference scores was critical because the difference in listening times between the segments of interest in syntactically complex and simple sentences is an index of effects of processing load on participants' performance.

[3] With respect to part of speech, we have been unable to find any studies of the effect of part of speech on spoken word recognition. According to L. K. Tyler (personal communication, September 2003), there is no evidence for single word-processing differences between noun and verb stems in normal healthy people (from priming and other tasks). Such studies as we could find were for written presentation and came from functional neuroimaging studies. Perani, Cappa, and Schnur (1999) reported no differences in reading times for matched nouns and verbs. Tyler, Russell, and Fadili (2001) found faster lexical decision and semantic categorization times for nouns than for verbs. The magnitude of these effects, which might reflect different time courses of accessing semantic or syntactic features of verbs and nouns, were 31 ms in lexical decision and 85 ms in semantic categorization. Differences of this magnitude are unlikely to be the sole source of differences between the critical items in the sentences we presented, which averaged several centiseconds (see, e.g., Waters & Caplan, 2001). Also, any noun advantage in lexical-access time is likely to be reduced in high-span and young subjects, because they have faster lexical access, and decreases in lexical-access time are likely to be proportional to lexical-access speed. Accordingly, category effects on speed of activation of lexical information, if they exist for spoken language, are likely to reinforce the conclusions of the article. We return to these points in the Discussion section.

Table 3
Background Characteristics of Older Adults for Three Sets of Analyses

	Models		
	Online processing	Offline sentence comprehension	Text comprehension
Measure	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
MMSE	28.1 (1.7)	28.1 (1.7)	28.0 (1.8)
BNT	53.9 (4.8)	53.9 (4.8)	53.8 (4.5)
LMI	62.0 (30.2)	62.0 (30.2)	64.2 (29.6)
LMII	62.7 (29.7)	62.7 (29.7)	64.7 (28.9)

Note. MMSE = Mini-Mental State Examination (scored out of 30 points); BNT = Boston Naming Test (scored out of 60 points); LMI and II = Logical Memory I and II (percentile scores).

had significantly higher spans than did the older adults on all three measures of vWM. Younger and older adults' response times differed significantly on one measure of online syntactic processing (COV − CSN2). These results demonstrate that there are age differences among the criterion variables, although the differences are observed more consistently in the measures of vWM than in the measures of online sentence processing.

Similar results were found in correlational analyses. Table 4 presents intercorrelations between all of the measures for younger and older adults. Age was significantly correlated with the three measures of vWM and with one measure of online sentence processing (COV − CSN2). The three measures of vWM were moderately intercorrelated (range = .41 to .49) and also showed some weak negative correlations with measures of online sentence processing (range = −.17 to .01). The measures of online sentence processing were also moderately intercorrelated (range = .29 to .53).

Model fitting was completed using the least squares method of EQS (Bentler, 1989) to estimate the standardized parameters of the model. Model fit was evaluated on the basis of significant parameter estimates, examination of standardized residuals, and goodness-of-fit indices. Standardized residuals in EQS represent the difference between the observed and predicted correlation matrix. These were examined to ensure that they were symmetrical and centered around zero and that few to none exceeded .20 in absolute value (Hatcher, 1994).

Chi-square assesses the extent to which the model fits the data. When the difference between the sample covariance matrix and the estimated population matrix is small, chi-square will be nonsignificant (Hatcher, 1994). However, chi-square is extremely sensitive to sample size and may be significant because of trivial differences between the sample and estimated populations (Ullman, 1996). For this reason, other model fit indices that are less sensitive to sample size have been developed. The standardized root-mean-square residual (SRMR) was included to assess model fit based on the difference between sample variances and covariances and estimated population variances and covariances. Values less than .08 are indicators of good model fit (Hu & Bentler, 1999). The root-mean-square error of approximation (RMSEA) measures the discrepancy of model fit with respect to the degrees of freedom and is less sensitive than chi-square to sample size (Loehlin, 1998). RMSEA values of less than .05 are considered to be an indication of excellent model fit, whereas values less than .10 are indicators of acceptable model fit (Loehlin, 1998). The comparative fit index (CFI) and non-normed fit index (NNFI) are measures that assess the model's ability to account for the observed data. Values for both range from 0 to 1.0, and values greater than .90 are indicative of a good model fit.

Because the sample included extreme groups of older and younger adults, the first step was to demonstrate that the factor structure to be tested was invariant across age groups. The results of these analyses are presented in the Appendix. The first step was accomplished by first fitting measurement models to each age group separately. Next, invariance of factor loadings and invariance of structural parameters (factor covariances) were separately tested. This series of analyses indicated that the factor structures of online sentence processing and vWM were equivalent for the younger and older age groups, and therefore we proceeded to model the data for the two age groups together.

A measurement model was fit to the vWM and online processing variables in order to ensure that the manifest variables provided a good measure of the constructs of interest. This model is presented in Figure 1. The standardized residuals ranged from .00 to .20, with no value exceeding .20. All factor loadings were significant at $p < .05$ and ranged from .58 to .83 for the vWM construct and from .50 to .86 for the online sentence-processing construct. The covariance between the latent constructs was not significant ($z = -1.3$). On the basis of these results and the fit indices, the model was considered to provide an acceptable fit to the data.

The next step was to test causal relationships among the factors and the variable age. Figure 2 presents the relationships among the factors and age that were tested in the present set of structural models. These models tested whether there is a direct effect of vWM on online sentence processing and whether there are direct and mediated effects of age on online processing. Goodness-of-fit measures for the models tested in this set of analyses are summarized in Table 5 (under "Models of vWM and online sentence processing").

Table 4
Correlations Among Manifest Variables in Online Models

Variable	1	2	3	4	5	6	7
1. Age	—						
2. Alphabet span	−.43**	—					
3. Subtract 2 span	−.36**	.49**	—				
4. Sentence span	−.35**	.41**	.41**	—			
5. COV − CSN2	.32**	−.14*	−.14*	−.17*	—		
6. SOV1 − OSN2	.11	−.06	.01	−.15*	.53**	—	
7. SOV2 − OSV1	−.06	.01	.01	−.07	.29**	.43**	—

Note. COV − CSN2 = Listening times for verb (V) of the cleft object (CO) minus the 2nd noun phrase (N2) of the cleft subject (CS) sentences; SOV1 − OSN2 = the embedded verb (V1) of the subject object (SO) minus the 2nd noun phrase (N2) of the object subject (OS) sentences; SOV2 − OSV1 = the main verb (V2) of the SO minus the main verb (V1) of OS sentences.
* $p < .05$. ** $p < .001$.

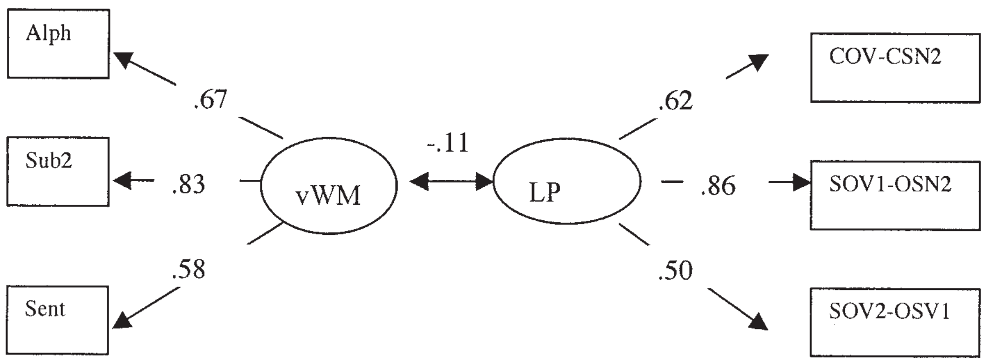

Figure 1. Measurement model of verbal working memory (vWM) and online sentence processing. Alph = alphabet span, Sub2 = subtract 2 span, Sent = sentence span, LP = online syntactic processing, COV-CSN2 = verb (V) of the cleft object (CO) minus the 2nd noun (N2) phrase of the cleft subject (CS) sentences, SOV1-OSN2 = the embedded verb (V1) of the subject object (SO) minus the 2nd noun phrase of object subject (OS) sentences, SOV2-OSV1 = the main verb of SO (V2) minus the main verb (V1) of OS sentences.

Model 1 postulated a direct influence of age on vWM and online sentence processing, as well as a direct effect of vWM on online sentence processing. Examination of the goodness-of-fit indices in Table 5 (under "Models of vWM and online sentence processing") reveals an adequate model fit, and the standardized residuals ranged from .00 to .19 with an average off-diagonal value of .05. Age was significantly related to vWM ($z = -6.7$). The parameter representing direct effects of vWM on online processing was not significant ($z = -0.5$), and the parameter representing direct effects of age on online language processing only approached

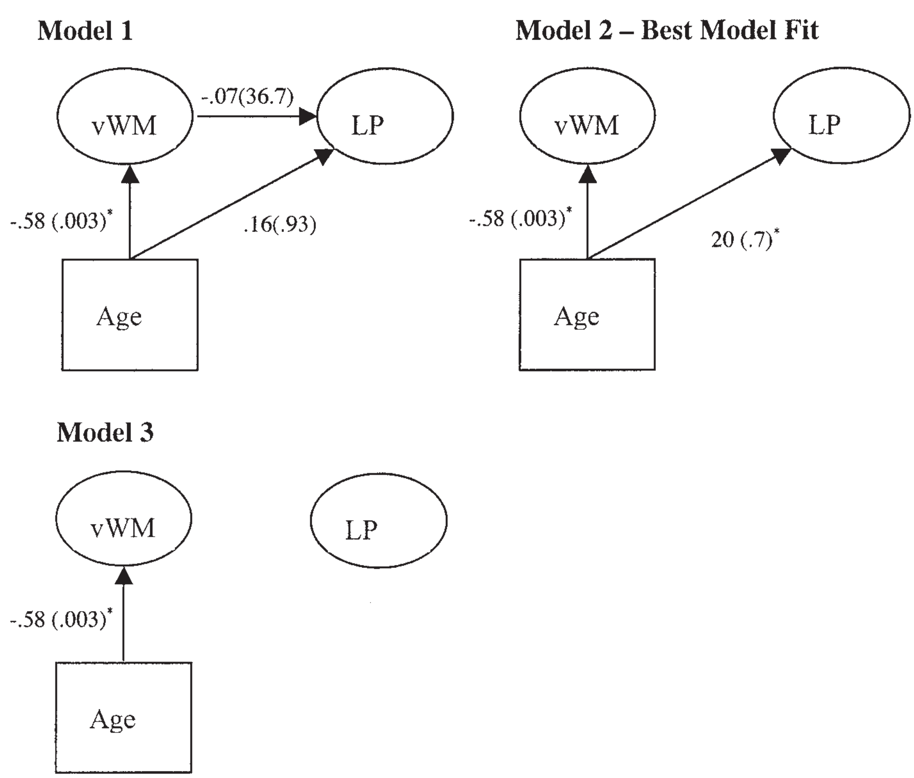

Figure 2. Models of age, verbal working memory (vWM), and online sentence processing. vWM includes the manifest variables of alphabet span, subtract 2 span, and sentence span. LP = online syntactic processing, with the manifest variables of (a) COV-CSN2 = verb (V) of the cleft object (CO) minus the 2nd noun phrase (N2) of the cleft subject (CS) sentences; (b) SOV1-OSN2 = the embedded verb (V1) of the subject object (SO) minus the 2nd noun phrase (N2) of the object subject (OS) sentences; (c) SOV2-OSV1 = the main verb (V2) of SO minus the main verb (V1) of OS sentences. Values given are betas (standard error). *$p < .05$.

Table 5
Model Fit Indices for All Models

Model	χ^2	*df*	*p*	CFI	NNFI	RMSEA	SRMR
Models of vWM and online sentence processing							
Measurement	18.39	9	.03	.96	.93	.07	.08
Model 1	30.26	12	.001	.94	.90	.08	.06
Model 2	30.54	13	.001	.94	.91	.08	.06
Model 3	35.78	14	.001	.93	.90	.09	.09
Models of vWM and offline sentence comprehension							
Measurement	16.98	9	.05	.97	.96	.06	.10
Model 4	18.02	12	.12	.99	.97	.05	.04
Model 5	18.03	13	.16	.99	.98	.04	.04
Models of vWM and text comprehension							
Model 6	10.40	4	.03	.98	.95	.08	.03
Model 7	10.97	5	.05	.98	.96	.07	.03

Note. CFI = comparative fit index; NNFI = nonnormed fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; vWM = verbal working memory.

significance ($z = 1.5$). Examination of the effect sizes of these parameters revealed that the coefficient of the former was less than twice its standard error, whereas the coefficient of the latter was greater than twice its standard error.

Two further models were generated. Model 2 was created by eliminating the nonsignificant parameter that linked vWM and online language processing. Because the parameter postulating a direct effect of age on online language processing was borderline significant and had a reasonable effect size, it was retained in this analysis. Model 2 was associated with adequate goodness-of-fit indices, and the standardized residuals ranged from .00 to .19 with an average off-diagonal value of .05. All parameter estimates were significant at $p < .05$. Because Model 2 was nested in Model 1, the models were compared using the chi-square difference test, $\chi^2(1) = 0.28$, $p > .05$, suggesting that dropping the nonsignificant parameter did not significantly reduce the model fit (Ullman, 1996). Given that the relationship between vWM and online sentence processing was nonsignificant and that retaining it did not significantly improve the model fit, Model 2 is preferred over Model 1.

Model 3 was generated to determine the effects of dropping both the parameter linking vWM to online sentence processing and the parameter linking age to online sentence processing. This model had acceptable goodness-of-fit indices and standardized residuals (range = 0.00 to 0.30, with only one value greater than .20; average off-diagonal value = .06). When Model 3 was compared to Model 2, the difference was significant, $\chi^2(1) = 5.24$, $p > .05$, suggesting that eliminating the relationship between age and online sentence processing significantly reduced model fit. Model 2 therefore best captures the relationship between vWM, age, and online language processing.

vWM and Offline Sentence Comprehension

None of the older participants consistently performed below the normal range for their age group on the background measures (see Table 3). Table 1 (under "Models of vWM and offline sentence comprehension") presents the mean age and education for both groups, as well as descriptive statistics of the manifest variables entered into this set of models for the younger and older participants separately as well as for the group as a whole. The two groups did not differ significantly in their years of education.

Age effects on vWM and offline sentence comprehension were apparent in both *t* tests and correlational analyses. Younger adults had significantly higher spans than the older adults on the three vWM measures, and they made more accurate plausibility judgments about sentences. Age was significantly correlated with the three measures of vWM and with the three measures of offline sentence comprehension (Table 6). The three measures of vWM were also moderately intercorrelated (range = .41 to .55), as were the measures of offline sentence comprehension (range = .45 to .67). Measures of vWM and offline sentence comprehension showed weak to moderate positive intercorrelations (range = .14 to .34).

Model estimation and fitting procedures were identical to those described above. Model fitting was first conducted separately for

Table 6
Correlations Among Manifest Variables in Offline Sentence Models

Variable	1	2	3	4	5	6	7
1. Age	—						
2. Alphabet span	−.43**	—					
3. Subtract 2 span	−.36**	.50**	—				
4. Sentence span	−.35**	.41**	.41**	—			
5. Acc CO	−.23*	.34**	.31**	.22**	—		
6. Acc CS	−.31**	.30**	.14*	.20*	.53**	—	
7. Acc SO	−.17*	.31**	.20*	.18*	.67**	.45**	—

Note. Accuracy (Acc) cleft object (CO), cleft subject (CS), and subject object (SO) refer to the accuracy of plausibility judgments on the offline sentence comprehension measure.
* $p < .05$. ** $p < .001$.

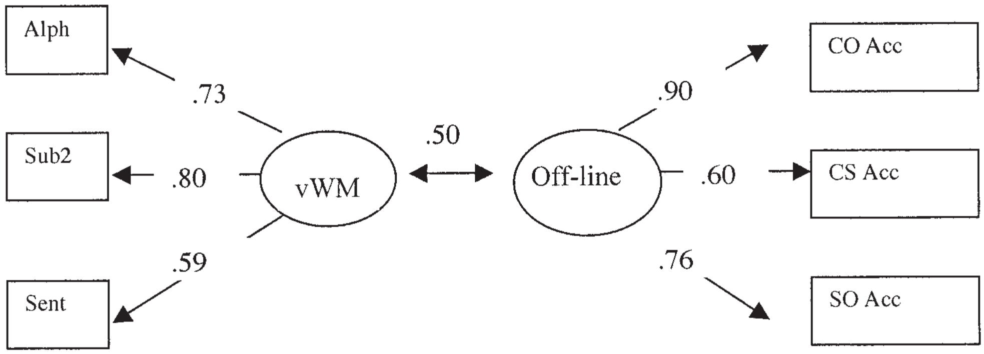

Figure 3. Measurement model of verbal working memory (vWM) and offline sentence comprehension. Alph = alphabet span, Sub2 = subtract 2 span, Sent = sentence span. Off-line = offline sentence comprehension. CO (cleft object) Acc, CS (cleft subject) Acc, and SO (subject object) Acc refer to the accuracy of plausibility judgments on the offline sentence comprehension measure.

the older and younger participants to verify that the assumption of factor structure invariance was met. The results of these analyses are presented in the Appendix. Although the assumptions of factor invariance were met, the relationship between vWM and offline sentence comprehension was noted to be stronger in the older than in the younger adults. A measurement model was then fit to the data to ensure that the observed variables were good predictors of the underlying latent variables (Figure 3). The goodness-of-fit indices of the measurement model are presented in Table 5 (under "Models of vWM and offline sentence comprehension"). The indices indicate that the model provides an acceptable fit to the data. In addition, the absolute standardized residuals ranged from .01 to .25 (with only one value >.20), with an average off-diagonal value of .07. All factor loadings were significant at $p < .05$. Factor loadings for the vWM construct ranged from .59 to .80, and factor loadings for the offline sentence comprehension construct ranged from .60 to .90.

The models generated to test the hypothesized relationships between age, vWM, and offline sentence comprehension are presented in Figure 4, and the goodness-of-fit indices associated with each model are presented in Table 5 (under "Models of vWM and offline sentence comprehension"). Model 4 hypothesized that age has both a direct effect on offline sentence comprehension and an indirect effect that is mediated through vWM. Examination of the goodness-of-fit indices for Model 4 indicates that it provides a good fit to the data. Evidence of model fit is also observed in the small absolute standardized residuals (range = .00 to .14; average offdiagonal value = .03). All parameter estimates were significant with the exception of the parameter directly linking age and offline sentence comprehension ($z = 0.12$).

Model 5 was generated by dropping the nonsignificant parameter to test its relative contribution to the model fit. Examination of the model fit indices and standardized residuals (range = .00 to .14; average off-diagonal value = .03) indicates that this model also provides a good fit to the data. Comparison of Models 4 and 5 using the chi-square difference method indicated that Model 5 provided the best fit to the data, $\chi^2(1) = 0.01$, $p < .05$ (Ullman, 1996).

vWM and Text Comprehension

Table 3 presents the mean and standard deviation of the performance on the background measures of the older adults whose

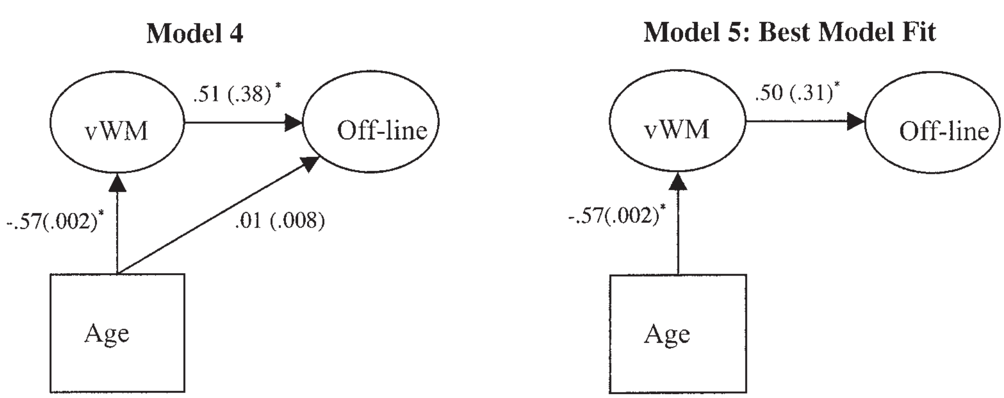

Figure 4. Models of age, verbal working memory (vWM), and offline sentence comprehension. vWM includes the manifest variables of alphabet span, subtract 2 span, and sentence span. Off-line = offline sentence comprehension, with the manifest variables of cleft subject accuracy, cleft object accuracy, and subject object accuracy. Values given are beta (standard error). *$p < .05$.

results are included in these models. None of the participants showed a consistent pattern of performing below the normal range for their age group. Table 1 (under "Models of vWM and text comprehension") presents the mean age and education for the groups and the descriptive statistics of the variables included in this set of models, for the younger and older participants separately and for the group as a whole. As with the previous sets of analyses, younger adults had significantly higher spans on all measures of vWM than did the older adults. Younger adults also achieved higher scores on the Nelson-Denny Reading Comprehension Test (Nelson & Denny, 1960) than the older adults. Table 7 presents the intercorrelations among the measures for both participant groups. All of the measures were significantly correlated with age (range = −.21 to −.48) and with one another (range = .21 to .51).

Model estimation and fitting proceeded as has been described above. Models were fit to the younger and older adults' data separately, confirming that the assumption of factor invariance was met. These data are presented in the Appendix. However, because there was only one observed variable for the text comprehension factor and the strength of the vWM construct was established in the preceding analyses, model fitting began with the testing of theoretical causal predictions. Figure 5 presents the different causal relationships considered among the factors and the variable age, and Table 5 (under "Models of vWM and text comprehension") presents the values of the associated fit indices.

Model 6 postulated a causal relationship between vWM and text comprehension, as well as direct effects of age on vWM and on text comprehension. This model provided an adequate fit to the data as evidenced by the fit indices and examination of the standardized residuals (range = .00 to .06; average off-diagonal value = .03). All parameters were significant at $p < .05$, with the exception of the direct link between age and text comprehension.

Model 7 was generated from Model 6 by dropping the nonsignificant link between age and text comprehension. Model 7 also provided an excellent fit to the data, as seen in the fit indices and the standardized residuals (range = 0 to .05; average off-diagonal value = .04). Models 6 and 7 were compared by using the chi-square difference method (Ullman, 1996). The results indicated that dropping the nonsignificant parameter did not significantly change the model fit, $\chi^2(1) = 0.06$, suggesting that the more parsimonious model (Model 7) should be retained.

Discussion

This series of structural equation models demonstrate that the relationship between age, vWM, and language differs, depending on what aspect of language processing is measured. Although the preferred models of online processing included a direct effect of age, they did not incorporate direct or mediated effects of vWM. In contrast, for both offline sentence comprehension and text level processing, effects of age were mediated by vWM.

Although the models did not show that online processing was mediated by vWM, two measures of online processing (COV − CSN2 and SOV1 − OSN2) showed significant negative correlations with age and vWM. The correlations between the online measures and vWM were weak (less than .20) and may have been significant because of the large sample size. For the sentences with two full clauses (SO and OS), the correlations were inconsistent. However, COV − CSN2 was negatively correlated with all of the vWM measures. Waters and Caplan (in press) argued that the verb of CO sentences is a locus of increased difficulty for constructing discourse structure. In CO sentences, the focus of discourse is the theme of the verb. This violates the typical agent focus of English sentences (Kintsch, 1998), and the verb is the first opportunity for listeners to recognize this violation. Hence, the correlations observed between COV − CSN2 and vWM measures may in part reflect these additional discourse demands. Nonetheless, the magnitude of the correlations and the failure to find a significant relationship between vWM and online sentence processing in the structural models argue against the presence of significant effects of vWM on online sentence processing.

Before concluding, we consider several issues that have risen about the materials and participants included in our study. The first issue pertains to the materials. As discussed in Footnote 1, the critical words might have been recognized earlier in complex sentences than in simple sentences, leading to an underestimation of the effect of syntax on online processing. One question is whether there could be individual differences in lexical access that might affect the interpretation of our results. Although there are few data on this subject, it is reasonable to assume (and the single resource theory maintains) that high-span individuals recognize words more efficiently than do low-span individuals. If this assumption is true, then any underestimation of the effect of syntax because of early word recognition would have been greater in high-span than in low-span individuals. This would reduce an inverse correlation between our measure of syntactic processing and vWM and would have biased the materials against providing support for the separate resource theory.

If there were systematic age differences in uniqueness points (or in the effect of uniqueness points on word recognition), then the underestimation of the effect of syntax might influence the sensitivity of the Age × Complexity interaction. To our knowledge, there are no data that address the stability of uniqueness points over the life span. However, it is unlikely that age effects of syntactic processing are being underestimated because of age-related differences in uniqueness points. The logic is as follows. As discussed above, the earlier the uniqueness point, the more corrected listening times underestimate the syntactic processing time. Therefore, if older adults had earlier uniqueness points, then we might be underestimating the effect of complexity in them relative to younger adults. However, it is unlikely that uniqueness points are earlier in older adults. Older adults do not process words faster; if anything, they are likely to process them more slowly than younger adults (Salthouse, 1996). Also, older adults tend to have larger vocabularies (Bayles & Kaszniak, 1987), and there-

Table 7
Correlations Among Manifest Variables in Text Comprehension Models

Variable	1	2	3	4	5
1. Age	—				
2. Nelson-Denny	−.21**	—			
3. Alphabet span	−.48**	.36**	—		
4. Subtract 2 span	−.34**	.21**	.51**	—	
5. Sentence span	−.42**	.29**	.44**	.44**	—

Nelson-Denny = Nelson-Denny Reading Comprehension Test.
** $p < .001$.

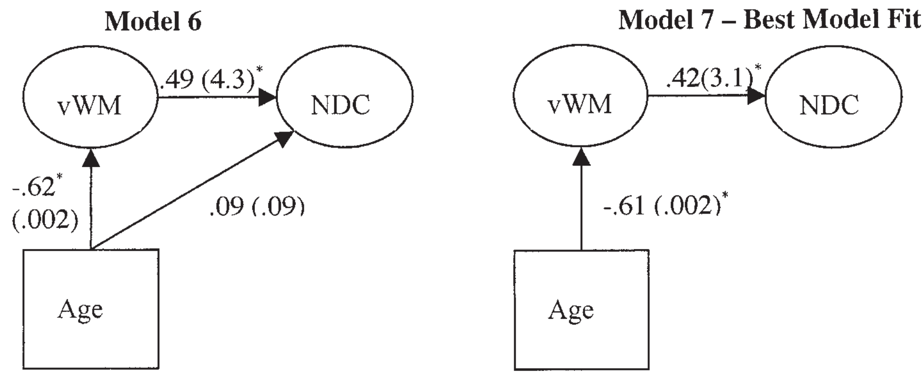

Figure 5. Models of age, verbal working memory (vWM), and text processing. vWM includes the manifest variables of alphabet span, subtract 2 span, and sentence span. NDC = Nelson-Denny Reading Comprehension test. Values given are beta (standard error). $^*p < .05$.

fore, if their uniqueness points differ from younger adults, then the uniqueness points would be later in the words. This is likely a very small effect: Vocabulary does not increase so much as to change the uniqueness point for words such as those used in the present study.

The second issue pertains to the participants included in these studies. The effects of age observed in the present study may have been inflated by the use of extreme age groups. In fact, the magnitude of the age effect on vWM in the present study (range = −.57 to −.62) is greater than the average age by WM correlation of −.27 found by Verhaeghen and Salthouse (1997) in a meta-analysis of 75 studies. This finding is particularly important in the analyses of offline sentence comprehension, because the observed relationship between vWM and age might be inflated because of the extreme groups design. Separate group analyses (presented in the Appendix) revealed a weaker relationship between vWM and offline sentence comprehension in the younger compared with older adults. Similarly, direct effects of age on online sentence processing may appear stronger in the present study than is true of the general population. The use of an extreme groups design also may have inflated the magnitude of age-related vWM effects on online sentence processing, creating a possible bias toward finding results in support of the single resource theory. That such effects were not found lends further support to the separate resource theory.

The use of extreme age groups also prevents us from directly measuring the linearity of relationships among measures of vWM, language processing, and age. Van der Linden et al. (1999) demonstrated that these relationships were linear for their data, which included measures of vWM, sentence recall, and text comprehension. Such data lend some support to the claim that there is a linear relationship between age, vWM, and sentence and text level processing. Such data are not available for the online measures. For this reason, the results must be interpreted with the caveat that nonlinear relationships among the variables could underlie the data and be influencing the observed pattern of results.

The present study does not directly compare relationships among the three language-processing measures. The reason is that there were not enough participants who had completed all of the language comprehension tasks to complete a larger analysis involving all of the variables. This prevents us from making strong statements about the relative strength of the relationships between age, vWM, and the measures of language comprehension. Nonetheless, the results do suggest that these relationships are not the same for all measures of language comprehension and that vWM does not have a direct effect on online syntactic processing. The results also suggest that studies that conflate different types of measures of language processing do not adequately examine the complex relationship between language comprehension and other cognitive processes. All of the comprehension measures in the present study tapped immediate comprehension, with minimal emphasis on recall of sentences or text, but the effects of vWM differed among these measures. Measures of comprehension that differ to greater extents are even more likely to bear different relationships to vWM. The results thus support claims such as the one by Van der Linden et al. (1999) that differences in measures of language processing in their study and that of Kwong See and Ryan's (1995) study may account for the different results between those studies.

Our study was restricted to the relationship between vWM and sentence processing and comprehension, and we did not examine the effects of differences and age-related decline in other cognitive functions, such as processing speed and inhibition, on these abilities. We therefore cannot say whether such variables might affect online sentence processing. However, the caution about using unitary measures of sentence processing and comprehension that we have voiced about the effect of vWM on these processes applies to the effects of these other cognitive variables on online sentence processing and different aspects of offline comprehension as well. For instance, as noted in the introductory section of this article, Kwong See and Ryan (1995) found that WM no longer predicted age-related variance in a composite measure language processing when processing speed and inhibition were taken into account. However, the finding in the present study that the effects of age and of age-related changes in WM differ in online sentence processing, offline sentence comprehension, and offline text comprehension suggests that these different aspects of the comprehension process may be mediated in different ways by cognitive variables. Specifically, one possibility is that processing speed and inhibition may mediate performance on online sentence processing, whereas offline sentence and text comprehension may be primarily mediated through vWM. Other results that point to potentially important differences in the role of general cognitive variables in mediating different comprehension tasks come from the differences in the variables that mediate performance as a function of modality. For example, Hultsch, Hertzog, and Dixon (1990) found that both speed and WM contributed to written text recall, whereas Frieske and Park (1999) found that measures of acuity and speed best predicted auditory text recall. The relations between general cognitive functions and different aspects of comprehension require studies that include measures of different aspects of language processing in a model with measures of WM, processing speed, inhibition, and sensory function.

The results of this study support the separate resource theory, which claims that variation in age and vWM is predictive of performance on offline measures of sentence processing, such as sentence or text comprehension but not online sentence comprehension (Caplan & Waters, 1999a, 1999b) over the single resource theory, which claims that variation in age or vWM is predictive of performance on both online and offline processing tasks (Just & Carpenter, 1992; King & Just, 1991). This raises the question of what distinctions need to be made in the complex set of processes that are grouped under the terms "comprehension" and "sentence processing." Caplan and Waters (1999b; Waters & Caplan, 2003a) have claimed that the separate processing resource for language is used for both assignment of syntactic structure and generation of the preferred, literal meaning of an utterance. Clearly, this general characterization needs to be made much more specific, both in theoretical terms and in terms of what tasks measure operations that are related to the two hypothesized vWM systems. Our results provide some empirical data relevant to the second of these questions, because they suggest that measures of comprehension taken immediately after completion of the sentence tap the postinterpretative processing resource that relies on the same system as that tapped by standard tests of vWM.

WM measures such as those used in our study (e.g., sentence span) have been postulated to capture a domain general capacity (Engle & Kane, 2004; Engle, Kane, & Tuholski, 1999). On this view, domain-specific effects are reflective of the short-term memory components of the WM system. In this study, we have used the term vWM because all of our WM measures were verbal. However, it is possible that these measures tap the domain-general capacity postulated by Engle and colleagues (Engle & Kane, 2004; Engle, Kane, & Tuholski, 1999). If so, the results of our study suggest that offline sentence and text comprehension measures are part of the domain general capacity, whereas online sentence processing taps a separate WM resource. One prediction that follows from these claims is that measures of sentence and text comprehension should correlate with measures of general intelligence (e.g., Raven's Progressive Matrices), whereas online sentence processing may not.

To summarize, the results of the present study suggest that age-related differences in vWM are not associated with online syntactic processing, but they may be related to differences in sentence and text comprehension among younger and older adults. These results provide evidence for the theory that online syntactic processing taps a separate vWM resource than that measured by standard vWM measures such as reading span.

References

Baddeley, A. D. (1986). *Working memory*. Oxford, United Kingdom: Clarendon Press.

Balogh, J., Zurif, E. B., Prather, P., Swinney, D., & Finkel, L. (1998). Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language, 61,* 169–182.

Baum, S. (1991). Sensitivity to syntactic violations across the age span: Evidence from a word-monitoring task. *Journal of Clinical Linguistics and Phonetics, 5,* 317–328.

Bayles, K. A., & Kaszniak, A. W. (1987). *Communication and cognition in normal aging and dementia.* Austin, TX: Pro-Ed.

Bentler, P. (1989). EQS: Structural Equations Program (Version 5.6) [Computer software]. Encino, CA: Multivariate Software.

Caplan, D., Hildebrandt, N., & Waters, G. S. (1994): Effects of verb selectional restrictions on the construction of syntactic form. *Language and Cognitive Processes, 9,* 549–585.

Caplan, D., & Waters, G. S. (1990). Short-term memory and language comprehension: A critical review of the literature. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments in short-term memory* (pp. 337–389). Cambridge, United Kingdom: Cambridge University Press.

Caplan, D., & Waters, G. S. (1999a). Issues regarding general and domain-specific resources. *Behavioral and Brain Sciences, 22,* 77–94.

Caplan, D., & Waters, G. S. (1999b). Verbal working memory capacity and language comprehension. *Behavioral and Brain Science, 22,* 114–126.

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension: Evidence from individual difference, aphasia and aging. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1075–1122). New York: Academic Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research: Methods, Instruments and Computers, 25,* 257–271.

Cohen, L., & Mehler, J. (1996). Click monitoring revisited: An on-line study of sentence comprehension. *Memory & Cognition, 24,* 94–102.

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

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

Davis, G. A., & Ball, H. E. (1989). Effects of age on comprehension of complex sentences in adulthood. *Journal of Speech and Hearing Research, 32,* 143–150.

Delis, D. C., Freeland, J. H., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test—Research Edition*. New York: Psychological Corporation.

Dunn, J. (1994). SoundEdit (Version 2) [Computer software]. San Francisco, CA: Macromedia.

Emery, O. (1985). Language and aging. *Experimental Aging Research, 11,* 3–60.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). New York: Elsevier.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

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

Feier, C. D., & Gerstman, L. J. (1980). Sentence comprehension abilities throughout the adult life span. *Journal of Gerontology, 35,* 722–728.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence from the auditory moving window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 324–335.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Ford, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 203–218.

Francis, W. N., & Kucera, H. (1982). *The frequency analysis of English usage*. Boston: Houghton Mifflin.

Frauenfelder, U., Segui, J., & Mehler, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior, 19,* 328–337.

Frieske, D. A., & Park, D. C. (1999). Memory for news in young and old adults. *Psychology and Aging, 14,* 90–98.

Gibson, E. (1998). Syntactic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1–76.

Hartley, J. T. (1986). Reader and text variables as determinants of discourse memory in adulthood. *Psychology and Aging, 1,* 150–158.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6(1),* 1–55.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1984). Text recall in adulthood: The role of intellectual abilities. *Developmental Psychology, 20,* 1193–1209.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5,* 356–368.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122–149.

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review, 103,* 773–780.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Kemper, S. (1986). Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics, 7,* 277–287.

Kemper, S. (1988). Geriatric psycholinguistics: Syntactic limitations of oral and written language. In L. Light & D. Burke (Eds.), *Language, memory, and aging* (pp. 58–76). Cambridge, United Kingdom: Cambridge University Press.

Kemper, S., & Kemtes, K. (1999). Limitations on syntactic processing. In S. Kemper & R. Kliegl (Eds.), *Constraints on language: Aging, grammar, and memory* (pp. 79–105). Boston: Kluwer Academic.

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in older and younger adults. *Psychology and Aging, 16,* 312–322.

Kemtes, K. (1999). *Decomposing adults' sentence comprehension: The role of age, working memory, inhibitory functioning, and perceptual speed*. Unpublished doctoral dissertation, University of Kansas, Lawrence.

Kemtes, K., & Kemper, S. (1997). Younger and older adults' on-line processing of syntactically ambiguous sentence. *Psychology and Aging, 12,* 362–371.

King, J. W., & Just, M. A. (1991). Individual difference in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580–602.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, United Kingdom: Cambridge University Press.

Kwong See, S. T., & Ryan, E. B. (1995). Cognitive mediation of adult age differences in language performance. *Psychology and Aging, 10,* 458–468.

Light, L. (1990). Interactions between memory and language in old age. In J. E. Birren & K. W. Schaie (Ed.), *Handbook of the psychology of aging* (pp. 275–290). San Diego: Academic.

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis*. Mahwah, NJ: Erlbaum.

Lord, F. M. (1956). Elementary models for measuring change. *Educational and Psychological Measurement, 16,* 421–437.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56–98.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676–703.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10,* 29–63.

Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple components view. *Neuropsychology, 8,* 506–523.

Nelson, M. J., & Denny, E. C. (1960). *The Nelson-Denny Reading Test*. New York: Houghton Mifflin.

Norman, S., Kemper, S., Kynette, D., Cheung, H. T., & Anagnopoulos, C. (1991). Syntactic complexity and adults' running memory span. *Journals of Gerontology, Series B*: *Psychological Sciences and Social Sciences, 46,* P346–P351.

Obler, L. K., Fein, D., Nicholas, M., & Albert, M. L. (1991). Auditory comprehension and aging: Decline in syntactic processing. *Applied Psycholinguistics, 12,* 433–452.

Perani, D., Cappa, S. F., & Schnur, T. (1999). The neural correlates of verb and noun processing: A PET study. *Brain, 122(12),* 2337–2344.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20,* 335–343.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428.

Stine, E. (1990). On-line processing of written text by younger and older adults. *Psychology and Aging, 5,* 67–78.

Stine, E., & Wingfield, A. (1990). How much do working memory deficits contribute to age differences in discourse memory? *European Journal of Cognitive Psychology, 2,* 289–304.

Stine-Morrow, E., Loveless, M. K, & Soederberg, L. M. (1996) Resource allocation in on-line reading by younger and older adults. *Psychology and Aging, 11,* 475–486

Stine-Morrow, E., Ryan, S., & Leonard, J. S. (2000). Age differences in on-line syntactic processing. *Experimental Aging Research, 26,* 315–322.

Stine-Morrow, E., Soederberg-Miller, L., & Leno, R. (2001). Patterns of on-line resource allocation to narrative text by younger and older adults. *Aging, Neuropsychology, and Cognition, 8,* 36–53.

Tun, P. A., Wingfield, A., & Stine, E. (1991). Speech-processing capacity in young and older adults: A dual-task study. *Psychology and Aging, 6,* 3–9.

Tyler, L. K., Russell, R., & Fadili, J. (2001). The neural representation of nouns and verbs: PET studies. *Brain, 124(8)*, 1619–1634.

Ullman. J. B. (1996) Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (3rd ed., pp. 709–819). New York: HarperCollins.

Van der Linden, M., Hupet, M., Feyereisen, P., Schelstraete, M., Bestgen, Y., Bruyer, M. R., et al. (1999). Cognitive mediators of age-related differences in language comprehension and verbal memory performance. *Aging, Neuropsychology, and Cognition, 6,* 11–24.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122,* 231–249.

Waldstein, R., & Baum, S. (1992). The influence of syntactic and semantic context on word monitoring latencies in normal aging. *Journal of Speech Language Pathology and Audiology, 6,* 217–222.

Waters, G., & Caplan, D. (1996). The capacity theory of sentence comprehension: A reply to Just and Carpenter (1992). *Psychological Review, 103,* 761–772.

Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging,* 16, 128–144.

Waters, G. S., & Caplan, D. (2002). The relationship between age, processing speed, working memory capacity, and language comprehension. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 57,* P289–P311.

Waters, G. S., & Caplan, D. (2003a). Individual differences in working memory capacity and on-line syntactic processing: Evidence from self-paced listening. *Quarterly Journal of Experimental Psychology, 57A,* 129–163.

Waters, G. S., & Caplan, D. (2003b). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, and Computers, 35*(4), 550–564.

Waters, G. S., & Caplan, D. (in press). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory.*

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale—Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1987). *The Wechsler Memory Scale—Revised*. San Antonio, TX: Psychological Corporation.

Wingfield, A., & Littlefield, K. C. (1995). Multiple memory systems in the processing of speech: Evidence from aging. *Experimental Aging Research, 21,* 101–121.

Zimmerman, D. W., & Williams, R. H. (1998). Reliability of gain scores under realistic assumptions about properties of pre-test and post-test scores. *British Journal of Mathematical and Statistical Psychology, 51,* 343–351.

Zurif, E., Swinney, D., Prather, P., Wingfield, A., & Brownell, H. (1995). The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research, 24,* 165–182.

Appendix

Additional Structural Equation Models

Online Sentence Processing

Model fitting for younger and older adults began with specification of a measurement model identical to the one specified in the main analyses of online sentence processing and vWM (e.g., see Figure 1). Goodness-of-fit indices (Table A1, under "Models of vWM and online processing") indicated good model fit for both groups. Standardized residuals for the older adults ranged from .00 to .16 (average standardized residual = .05), whereas younger adults' standardized residuals ranged from .00 to .15 (average standardized residual = .05). For the younger and older groups, all parameter estimates representing factor loadings were significant. The covariance of the two factors was not significant in either measurement model.

Factor invariance of the measurement model was tested to ensure that the relationship between the factors was invariant across age groups. Table A1 (under "Models of vWM and online processing") presents the goodness-of-fit measures for the models postulating invariance of factor loadings and invariance of factor structure. These models were compared using the chi-square difference test. The results indicate that the additional constraint on the factor loadings did not significantly reduce model fit, $\chi^2(1)= 1.60$, $p > .05$. The Lagrange multiplier test provided a test of the assumption that the parameters were equivalent across the two age groups. None of the tests reached statistical significance, indicating that the assumption of structural invariance was met.

Offline Sentence Comprehension

Measurement models identical to the one specified in the main analyses of offline sentence comprehension (see Figure 3) were generated separately for the younger and older adults. Goodness-of-fit indices for the measurement models (Table A1, under "Models of vWM and offline sentence comprehension") indicated acceptable model fit. Standardized residuals for older adults ranged from .00 to .35 (offdiagonal average = .08, with two values greater than .20). Standardized residuals for younger adults ranged from .00 to .21 (off-diagonal average = .07, with one value greater than .20). All parameter estimates representing factor loadings were significant for both groups. Covariance between the factors was significant for the older adults and approached significance for the younger adults ($z = 1.14$). The covariance between vWM and offline sentence comprehension was weaker in the younger ($\beta = .14$) than in the older adults ($\beta = .50$), but both groups showed a positive relationship between vWM and offline sentence comprehension.

Factor invariance was tested as above. Table A1 (under "Models of vWM and offline sentence comprehension") includes the goodness-of-fit measures for models postulating invariance of factor loadings and factor structure. As with the online processing model, invariance of the covariance among the factors was included to determine whether the relationship between vWM and offline sentence comprehension differed significantly across age groups. The results indicate that the additional constraint did not significantly reduce model fit, though it did approach significance, $\chi^2(1) = 3.40$, $p < .10$. Lagrange multiplier tests for releasing constraints revealed that all of the parameter constraints held. These results indicate that the assumption of structural equivalence cannot be rejected.

Text Comprehension

As in the main analyses, measurement models were not separately tested in the separate group analyses. For this reason, structural models corresponding with Model 6 in the main analyses were generated for the younger and older adults separately. Goodness-of-fit indices (Table A1, under "Models of vWM and text comprehension") indicate that these models provided a good fit to the data. Standardized

Table A1
Model Fit Indices for Models of Older and Younger Age Groups Separately

Age group and model	χ^2	*df*	*p*	CFI	NNFI	RMSEA	SRMR
Models of vWM and online processing							
Younger							
Measurement	5.8	9	.75	1.0	1.0	.00	.06
Older							
Measurement	10.5	8	.23	.98	.97	.05	.07
Younger & older: Invariance models							
Factor loading	20.2	21	.51	1.0	1.0	.00	.08
Factor structure	21.8	22	.47	1.0	1.0	.00	.07
Models of vWM and offline sentence comprehension							
Younger							
Measurement	12.2	9	.20	.97	.94	.07	.08
Older							
Measurement	17.0	9	.05	.93	.96	.08	.12
Younger & older: Invariance models							
Factor loading	30.5	22	.10	.97	.97	.04	.10
Factor structure	33.9	24	.09	.97	.96	.05	.12
Models of vWM and text comprehension							
Younger							
Dependent	2.0	4	.72	1.0	1.0	.00	.02
Older							
Dependent	7.5	4	.11	.94	.86	.08	.04
Younger & older: Invariance models							
Factor loading	10.8	10	.37	.99	.99	.01	.04
Factor structure	11.4	11	.40	.99	.99	.01	.04

Note. CFI = comparative fit index; NNFI = non-normed fit index; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual; vWM = verbal working memory.

residuals for older adults ranged from .00 to .13 (off-diagonal average = .04). Standardized residuals for younger adults ranged from .00 to .05 (off-diagonal average = .03). All parameter estimates were significant for both age groups. Factor invariance was tested as described above. A chi-square difference test indicated that constraining factor structures across age groups did not significantly reduce model fit, $\chi^2(1) = 0.60$, $p > .05$. Lagrange multiplier tests were nonsignificant, suggesting that the parameters were equivalent across age groups.

Received April 22, 2004
Revision received June 8, 2004
Accepted June 25, 2004 ■