

# Working memory and language comprehension: A meta-analysis

MEREDYTH DANEMAN

University of Toronto, Mississauga, Ontario, Canada

and

PHILIP M. MERIKLE

University of Waterloo, Waterloo, Ontario, Canada

This paper presents a meta-analysis of the data from 6,179 participants in 77 studies that investigated the association between working-memory capacity and language comprehension ability. A primary goal of the meta-analysis was to compare the predictive power of the measures of working memory developed by Daneman and Carpenter (1980) with the predictive power of other measures of working memory. The results of the meta-analysis support Daneman and Carpenter's (1980) claim that measures that tap the combined processing and storage capacity of working memory (e.g., reading span, listening span) are better predictors of comprehension than are measures that tap only the storage capacity (e.g., word span, digit span). The meta-analysis also showed that math process plus storage measures of working memory are good predictors of comprehension. Thus, the superior predictive power of the process plus storage measures is not limited to measures that involve the manipulation of words and sentences.

In this paper we present a meta-analysis of the research investigating the association between working-memory capacity and language comprehension ability. Much of the research was stimulated by a 1980 paper in which Daneman and Carpenter claimed to have developed a measure of working-memory capacity that was an excellent predictor of language comprehension ability. First we will describe Daneman and Carpenter's (1980) claims, and then we will present our meta-analysis of the research findings.

## THE BACKGROUND

In their 1980 paper, Daneman and Carpenter claimed to have resolved a paradox in the literature on individual differences in comprehension ability. Many theorists had suggested that short-term memory capacity plays a crucial role in reading and listening comprehension, and that short-term memory should therefore be an important source of individual differences in comprehension ability (cf. Just & Carpenter, 1980; Kintsch & van Dijk, 1978; Perfetti & Lesgold, 1977). After all, reading and listening involve much more than comprehending a stream of isolated words. A major component of skilled comprehension is the ability to compute the semantic and syntactic relations among successive words, phrases, and sen-

tences, thereby constructing a coherent and meaningful representation of the discourse. Integrating newly encountered information with the previously processed information means that readers and listeners must have access to the results of earlier processes. Otherwise, how could they compute the referent for *he* in the following sentence?—*Although he spoke softly, yesterday's speaker could hear the little boy's question* (Just & Carpenter, 1980). Or how could they make the crucial inference in the following passage?—*Jake bought his new girlfriend a turtle for Christmas. He decided to take it back to the shop when he discovered she liked soup* (Mitchell, 1982). Or how could they resolve the apparent inconsistency in the following passage?—*The dentist used the drill. He found it useful for mastering the Latin and Greek verbs* (Daneman & Carpenter, 1983). If temporary storage is important for successful comprehension, individuals with small temporary storage capacities should be less able to have earlier relevant information still active in memory, and therefore they should be less likely to compute the referent for a pronoun, make the necessary inference, and resolve the inconsistency; in short, they should have deficits in the processes that integrate successive ideas in a written or spoken discourse relative to individuals who have larger short-term memory capacities. So, according to *theory*, short-term memory capacity should be related to comprehension ability. Paradoxically however, the *empirical evidence* suggested that this was not so. Researchers had tried to find a correlation between performance on traditional measures of short-term memory, such as digit span, letter span, and word span, and performance on standardized tests of comprehension ability; however, unless their samples included very young children, or se-

---

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada to M. Daneman and P. M. Merikle. We thank Monica Davidson for her help with data search and categorization, Lee Sechrest for statistical consultation, and Akira Miyake for his useful comments and suggestions. Correspondence concerning this article should be addressed to M. Daneman, Department of Psychology, University of Toronto, Mississauga, ON, Canada L5L 1C6 (e-mail: daneman@psych.utoronto.ca).

verely disabled readers, the correlations obtained were at best very weak (e.g., Farnham-Diggory & Gregg, 1975; Perfetti & Goldman, 1976; Rizzo, 1939).

Daneman and Carpenter (1980) argued that this paradox resided in the theory behind the measures of short-term memory capacity. The theory was inadequate. Traditional digit span, letter span, and word span tests treat short-term memory as a passive storage buffer. The tests measure the maximum number of items that an individual can store in a task that simply requires that the individual attend to some input, encode it, store it, and retrieve it. Daneman and Carpenter (1980) argued that the source of individual differences in memory capacity does not reside in the passive storage capacity or number of "slots." Rather, individuals differ in functional capacity; that is, they differ in the processes that they have for maximally utilizing their limited capacities. To differentiate this more active view from the more classical "slot" conception of short-term memory, Daneman and Carpenter adopted Baddeley and Hitch's (1974) term, *working memory* (see also Baddeley, 1983, 1986; Miller, Galanter, & Pribram, 1960). The working-memory system is assumed to have processing as well as storage capabilities. There are limited resources that must be shared between the work and the memory, between the processing and storage demands of the task to which the working-memory system is being applied. Moreover, individuals differ in the ability to coordinate the processing and storage functions. In particular, individuals with inefficient processes have a functionally smaller temporary storage capacity, because they must allocate more of the available resources to the processes themselves. Daneman and Carpenter (1980) argued that a functionally smaller storage capacity would lead to deficits in comprehension, particularly in the processes that integrate successively encountered words, phrases, and sentences into a coherent representation. Moreover, they argued that if one is to measure this functional capacity, one needs a measure to assess the combined processing and storage resources of working memory—rather than simply the storage resources, as on traditional span tests. And so Daneman and Carpenter (1980) developed a measure to do just that. In their working-memory span test, participants read or listened to a set of unrelated sentences (e.g., *I turned my memories over at random like pictures in a photograph album; He had an odd elongated skull which sat on his shoulder like a pear on a dish; I will not shock my readers with the cold-blooded butchery that followed*), and then, at the end of the set, they attempted to recall the final word of each sentence in the set (e.g., *album; dish; followed*). Participants were presented with increasingly longer sets of sentences, and working memory was defined as the maximum number of sentences the participants could read or listen to while maintaining perfect recall of the final words. Thus, the test involved the usual demands of sentence comprehension, from the lower level processes that encode the visual patterns of individual words and access their meanings, to the higher level processes that compute the semantic,

syntactic, and referential relations among the successive words; and the test imposed the additional and simultaneous component of maintaining the final words. When administered to college students, working-memory span varied from two to five final words. The theory behind the test was that poor comprehenders allocate so much capacity to comprehending the sentences that they can only store two or three final words.

Unlike traditional digit span, letter span, and word span tests, which do not correlate well with reading and listening ability, the reading and listening versions of Daneman and Carpenter's (1980) working-memory-span test were excellent predictors of comprehension performance. Daneman and Carpenter (1980) found that individuals with working-memory spans of only two or three sentence-final words did more poorly on a global test of language comprehension (the Verbal Scholastic Aptitude Test, or VSAT) and particularly poorly on specific tests of the integration of successive ideas in a text, such as computing the antecedent referent for a pronoun. The correlations in the 1980 study were large. They ranged between .42 and .90, and, with an average correlation of .66, they were well above the .30 barrier that typically plagues research on individual differences (Hunt, 1980).

The impact of the Daneman and Carpenter (1980) study has been wide ranging. Daneman and Carpenter's reading and listening working-memory-span tests have been adopted by many researchers who have used them or variants of them to replicate and extend the original findings. Not only have the measures been applied to investigating working memory's role in accounting for the kinds of individual differences in language comprehension that one would encounter in a typical college classroom, but they have been applied to many different population groups and to many different cognitive activities. Researchers have used variants of the reading-span and listening-span measures to investigate the young and the old, the learning disabled and the reading disabled, the hearing impaired and the visually impaired, and the victims of stroke and other forms of brain damage (e.g., Daneman, 1988a; Daneman, Nemeth, Stainton, & Huelsmann, in press; Frisk & Milner, 1990; Light & Anderson, 1985; Salthouse & Babcock, 1991; Salthouse, Mitchell, Skovronek, & Babcock, 1989; Stine & Wingfield, 1987; Swanson, 1993; Tompkins, Bloise, Timko, & Baumgaertner, 1994). The cognitive activities investigated have been equally numerous. On the language side, researchers have used Daneman and Carpenter's (1980) measures to investigate working memory's contribution to speaking, writing, speech reading, and learning to read (e.g., Bock & Miller, 1991; Byrd, 1993; Daneman, 1991; Daneman & Green, 1986; Hupert, Nef, & Maroy, 1992; Lyxell & Rönnberg, 1993; Rönnberg, Arlinger, Lyxell, & Kinnerfors, 1989; see also Gathercole & Baddeley, 1993, for an excellent review). On the nonlanguage side, researchers have also used Daneman and Carpenter's (1980) measures to investigate working memory's contribution to cognitive activities as diverse as solving arith-

metic problems and reasoning problems, learning to program a computer, playing bridge, remembering the spatial location of objects, controlling attention, playing complex computer games, and so on (e.g., Cherry & Park, 1993; Clarkson-Smith & Hartley, 1990; Das-Smaal, De Jong, & Koopmans, 1993; Kyllonen & Christal, 1990; Lehrer, Guckenberg, & Lee, 1988; Lehrer & Littlefield, 1993; Logie, Baddeley, Mané, Donchin, & Sheptak, 1989; Salthouse & Kersten, 1993; Shute, 1991; Staver & Jacks, 1988). Many of these attempts to link working memory span with individual differences in performance on everyday cognitive activities have been successful, and consequently they provide an important validation of the working-memory approach. After all, working-memory theory was proposed as an alternative to short-term memory theory because of concerns with the ecological relevance of the short-term memory construct (Baddeley, 1983; Daneman, 1987); consequently, it is encouraging to see that working memory is living up to its promise of doing a better job at accounting for everyday cognitive performance than did the short-term memory theory that it replaced.

Of course, the original Daneman and Carpenter (1980) study was not without its critics. The research with reading span and listening span raised the question of whether comprehension is limited by the capacity of a general working-memory system or by one specialized for the language processes. A legitimate concern about the reading span and listening span tests is that they are too much like comprehension itself. Indeed Daneman and Carpenter (1980; see also Daneman, 1982) argued that their span measures may have been such successful predictors of comprehension precisely because they captured many of the processing requirements of sentence comprehension and consequently had an excellent probability of tapping the aspects of working memory that are important to comprehension. But by the same token, critics (e.g., Baddeley, Logie, Nimmo-Smith, & Brereton, 1985) have pointed out that the complexity of the Daneman and Carpenter (1980) span measures makes interpretation of the correlation difficult. And they have argued that the language-specific nature of the measures may leave us with the rather trivial conclusion that sentence comprehension (reading span/listening span) correlates with paragraph comprehension (the criterion comprehension tests).

To go beyond this trivial interpretation and further explore the issue of whether comprehension is limited by the capacity of a general working-memory system or by one specialized for language, several researchers have pitted reading span against nonverbal working-memory measures (e.g., Baddeley et al., 1985; Daneman & Tardif, 1987; Jurden, 1995; Shah & Miyake, 1996; Turner & Engle, 1989). Turner and Engle, for example, developed an analogue of reading span that taps mathematical/arithmetic processes rather than verbal processes. In the Turner and Engle math-span analogue, participants are given a set of simple arithmetic equations, such as  $(9/3) + 4 = 7$ ,  $(2 \times 3) - 2 = 4$ , and  $(6/3) + 2 = 8$ ; for each equation, their task is to verify whether the stated solution is

correct or incorrect, and then at the end of the set, they have to recall the stated solutions from each equation in the set (here, 7, 4, and 8). The question of theoretical interest is whether math span is as good a predictor of language comprehension as is reading span, a question that this meta-analysis addresses.

It is now 16 years since the publication of the Daneman and Carpenter (1980) study. Given the number of follow-up studies and the number of variants and mutants of the original Daneman and Carpenter measures of working-memory capacity that are being used, we thought it would be timely to take stock—to do a meta-analysis of the literature investigating the association between working-memory capacity and comprehension, and to see how the original Daneman and Carpenter claims are faring. It is difficult to assess the validity of Daneman and Carpenter's claims by sampling only a few of the published studies. Because many of the individual studies have small sample sizes, they often lack power and reliability; consequently, it is not surprising that the findings across studies are not always consistent. Some studies have replicated Daneman and Carpenter's (1980) finding that working-memory span is a good predictor of comprehension (e.g., Dixon, LeFevre, & Twilley, 1988; Masson & Miller, 1983), whereas other studies have not (e.g., Light & Anderson, 1985; Morrow, Leirer, & Altieri, 1992). Some studies have replicated Daneman and Carpenter's (1980) finding that working-memory span is a better predictor of comprehension than traditional word span and digit span are (e.g., Dixon et al., 1988; Masson & Miller, 1983; Turner & Engle, 1989), whereas other studies have not (e.g., Calvo, Ramos, & Estevez, 1992; La Pointe & Engle, 1990). Some studies have supported the superiority of verbal working-memory-span measures over math working-memory-span measures at predicting language comprehension (e.g., Baddeley et al., 1985; Daneman & Tardif, 1987), whereas other studies have not (e.g., Norman, Kemper, & Kynette, 1992; Turner & Engle, 1989). The advantage of using a meta-analysis to assess Daneman and Carpenter's (1980) claims is that it provides a method for combining and quantifying the results of individual studies so that we may see the general trends across all relevant studies and obtain more reliable estimates of population characteristics.

## THE META-ANALYSIS

### The Studies

The meta-analysis includes studies published through the end of April 1995. Most of the studies postdate the 1980 Daneman and Carpenter study, but we have included the earlier studies that investigated the relation between traditional span measures and comprehension, with a 1939 study by Rizzo being the earliest. Studies were located by manual and computer-assisted searches of *PsycLIT*, *Psychological Abstracts*, and the *Social Sciences Citation Index*, as well as the reference lists of published papers. To be included in the meta-analysis, a study had to meet the following three criteria: (1) The study had to have

**Table 1**  
**Summary of Predictor Tasks**

Process and storage (Daneman & Carpenter analogues)	verbal (e.g., reading span, listening span) math (e.g., math span, counting span, backward digit span)
Storage alone (traditional span)	verbal (e.g., word span, letter span) math (e.g., digit span, probe digit span)

examined the relation between one or more measures of working-memory capacity or short-term memory capacity and one or more measures of reading or listening comprehension accuracy.<sup>1</sup> (2) The study had to have quantified the association between working memory and comprehension by either a product-moment correlation coefficient ( $r$ ) or another statistic (such as a  $t$  or  $F$  ratio) that could be converted to a correlation coefficient. (3) The study had to have excluded extremely unskilled readers (i.e., beginning readers, severely disabled/dyslexic readers) from its sample; Daneman and Carpenter (1980) tried to show that their working-memory measures captured the kinds of variance in comprehension that one encounters in a typical college or school classroom, and so in the meta-analysis, we did not want to potentially inflate the estimates of the working-memory/comprehension correlation by including extreme groups. In all, the meta-analysis included data from 6,179 participants in 77 studies (independent samples).

The studies included in the meta-analysis were categorized according to the kinds of working-memory measures and comprehension measures that were used. On the *predictor task* side, we included studies that used analogues of the Daneman and Carpenter working-memory-span measures—that is, measures that assessed the simultaneous processing and storage resources of working memory—and we included studies that used the traditional span tests that tapped predominantly storage resources, such as forward digit span, letter span, and word span. As can be seen in Table 1, each of these categories was subdivided further into verbal and math tasks.<sup>2</sup>

Tables 2 and 3 summarize some of the more commonly used verbal and math process plus storage measures. In the case of the verbal process plus storage measures (see Table 2), the most common measure was Daneman and Carpenter's (1980) reading span test, in which participants read aloud sets of unrelated sentences and then recalled the final words. Other commonly used variants added a comprehension check to the process component of the task; after they read or listened to each sentence, participants were required to make a true/false judgment or a sensibility judgment about the sentence (e.g., Baddeley et al., 1985; Daneman & Carpenter, 1980). In most variants, participants were required to recall the final word of each sentence in the set; however, there were a few variants that cued participants to recall a different word in the sentence, such as the person noun or the object noun (e.g., Baddeley et al., 1985; see Table 2). In the case of the math process plus storage measures, the process component always tapped computations involving digits rather than computations involving letters, words, or sentences (see Table 3 for examples).

On the *criterion task* side, the comprehension measures included global or standardized tests of comprehension and vocabulary knowledge, the most common being the VSAT and the Nelson-Denny Reading Test, as well as more specific tests of integration—tests that measured people's ability to compute the referent for a pronoun, make inferences, monitor and revise inconsistencies, acquire new word meanings from contextual cues, abstract the main theme, and so on. Table 4 provides a summary of the more common criterion measures.

**Table 2**  
**Commonly Used Verbal Process Plus Storage Measures**

The rain and howling wind kept beating against the rattling window panes. He covered his heart with both hands to keep anyone from hearing the noise it made. Process: read aloud Recall: final word (e.g., panes, made) (e.g., Daneman & Carpenter, 1980)
Mammals are vertebrates that give birth to live young. March is the first month in the year that has thirty-one days. Process: true/false judgment Recall: final word (e.g., young, days) (e.g., Daneman & Carpenter, 1980)
The policeman ate the apple. The girl sang the water. Process: sensibility judgment Recall: person or object (e.g., policeman & girl or apple & water) (e.g., Baddeley et al., 1985)

**Table 3**  
**Commonly Used Math Process Plus Storage Measures**

---

$(2 \times 3) - 2 = 4$
$(6/3) + 2 = 8$
Process: verify the stated solution
Recall: stated solution (e.g., 4, 8)
(e.g., Turner & Engle, 1989)
$(2 \times 3) - 2 = 4$ tree
$(6/3) + 2 = 8$ door
Process: verify the stated solution
Recall: word (e.g., tree, door)
(e.g., Turner & Engle, 1989)
$3 + 2 = ?$
$7 - 4 = ?$
Process: perform arithmetic computation
Recall: the highlighted second digit (e.g., 2, 4)
(e.g., Salthouse et al., 1989)

---

With the preceding classification of tasks, we used the meta-analysis to address the following questions: (1) What are the magnitudes of the different working-memory/comprehension correlations? (2) Do measures that tax the processing as well as the storage resources of working memory predict comprehension better than do measures that tax only the storage resources? (3) How specific or language related do the processes have to be? Can a measure that taps math processes predict comprehension, or do the original Daneman and Carpenter (1980) findings boil down to a demonstration that comprehension correlates with comprehension? And if math span does predict comprehension, is it as good a predictor as verbal span?

### Data Analysis

With the four categories of predictor tasks (verbal process plus storage, math process plus storage, verbal stor-

age alone, math storage alone) and the two categories of criterion tasks (global and specific), there were eight different working-memory/comprehension effect sizes to be computed for the overall meta-analysis. The first step involved computing the effect size for each working-memory/comprehension association investigated in an individual study. Following Rosenthal (1991), we used the product-moment correlation coefficient ( $r$ ) as the effect size estimate. In many studies, the association between a working-memory measure and a comprehension measure was reported as a correlation coefficient, so the effect size was taken directly from the written report. In other instances, the correlation coefficients were computed from reported  $t$  ratios or  $F$  ratios. The number of effect sizes extracted from a single study varied from 1 to 4, depending on how many of the four categories of working-memory measures and two categories of comprehension measures were included in the study. For example, if a study included only one category of working-memory measure (e.g., a verbal process plus storage measure such as reading span) and one category of comprehension measure (e.g., a global measure such as the VSAT), then only one effect size was computed for that study (verbal process plus storage/global comprehension). If a study included two categories of working-memory measures (e.g., a verbal process plus storage measure, and a math process plus storage measure) and two categories of comprehension measures (a global measure and a specific measure), four effect sizes were computed for that study (verbal process plus storage/global comprehension; verbal process plus storage/specific comprehension; math process plus storage/global comprehension; math process plus storage/specific comprehension). In some cases, a study produced several different estimates of the same working-memory/comprehension association; this occurred if the study included more than one working-memory measure and/or comprehension measure from the same category. For example, a study may have used two

**Table 4**  
**Summary of Criterion Tasks**

---

Global (Standardized Tests of Comprehension and Vocabulary Knowledge)
Verbal Scholastic Aptitude Test
Nelson-Denny Reading Test
Mill Hill Vocabulary
General Aptitude Test Battery: Vocabulary
WAIS Vocabulary
Metropolitan Achievement Test
Armed Services Vocational Aptitude Battery: Word Knowledge
Armed Services Vocational Aptitude Battery: Paragraph Comprehension
Specific (Nonstandardized)
Assigning pronominal reference
Making inferences
Detecting ambiguity
Monitoring and revising inconsistencies
Abstracting the main theme
Following verbal directions
Recalling propositional content

---

verbal process plus storage measures of working memory such as reading span and listening span, and two measures of global comprehension such as the VSATs and the Nelson–Denny, a design that would produce four correlations: reading span/VSAT; reading span/Nelson–Denny; listening span/VSAT; listening span/Nelson–Denny. According to our eight-way classification system, all four correlations were estimates of the same working-memory/comprehension association—the verbal process plus storage/global comprehension association. In cases such as these, the correlations were averaged to yield a single effect size estimate for that particular working-memory/comprehension association for that study. The Appendix shows the studies that contributed to each of the eight working-memory/comprehension cells in the design, as well as the sample size and the effect size for each study.

Once we had computed the effect size for each working-memory/comprehension association in each study, the next step was to compute average effect sizes across studies by weighting individual studies according to their sample size. These average weighted effect sizes were computed by (1) transforming each  $r$  into a Fisher's  $Zr$ ; (2) multiplying each  $Zr$  by the number of participants; (3) summing the weighted  $Zr$ s and dividing each sum by the total number of participants; and (4) converting the resulting  $Zr$  back to an  $r$ .

The final step was to compute the 95% confidence interval for each average weighted effect size (see Hedges & Olkin, 1985, for procedure). The confidence interval provides the range within which an observed effect size is likely to fall 95 times out of 100, and it can be used to compare effect sizes. For example, if one effect size is .41 with a 95% confidence interval of .38 to .44, and another effect size is .28 with a 95% confidence interval of .23 to .33, the nonoverlapping confidence intervals provide support for the conclusion that the .41 effect size is larger than the .28 effect size.<sup>3</sup> Table 5 shows the number of independent samples and the number of participants for each working-memory/comprehension association. Table 6 shows the average weighted effect size and the confidence interval for each working-memory/comprehension association.

How does the original Daneman and Carpenter (1980) pattern of results based on 41 participants hold up over 77 studies and 6,179 participants?

The first point to note is that the results of the meta-analysis support Daneman and Carpenter's (1980) finding that verbal process plus storage measures of working-memory capacity are good predictors of comprehension. As Table 6 shows, the verbal process plus storage span measures correlate .41 and .52 with global and specific tests of comprehension ability, respectively. Although smaller in magnitude than those in the original Daneman and Carpenter (1980) study with only 41 participants,<sup>4</sup> these correlations are certainly respectable in an area that rarely sees correlations above .30 (Hunt, 1980). Moreover, given the large sample size, we can be reasonably sure that these correlations provide good estimates of the

**Table 5**  
Number of Independent Samples ( $N$ ) and  
Number of Participants ( $n$ ) for  
Each Working-Memory/Comprehension Association

	Global Comprehension		Specific Comprehension	
	$N$	$n$	$N$	$n$
Process and storage				
Verbal	38	2,961	36	2,745
Math	15	1,183	8	856
Storage alone				
Verbal	25	1,431	10	810
Math	10	2,344	8	1,487

magnitude of the working-memory/comprehension correlation in the population.

Not only does the meta-analysis support Daneman and Carpenter's (1980) finding that verbal process plus storage measures are good predictors of comprehension, it also supports Daneman and Carpenter's (1980) finding that working-memory capacity is a better predictor of performance on specific tests of integration (.52) than on global tests of comprehension and vocabulary knowledge (.41). This finding likely reflects the fact that working memory plays a particularly important role in the processes that integrate successively encountered ideas in a written or spoken discourse.

Do measures that tax the combined processing and storage resources of working memory predict comprehension better than do measures that tax only the storage resources? The results of the meta-analysis suggest that the answer to this question is "yes." As Table 6 shows, the meta-analysis supports Daneman and Carpenter's (1980) claim that measures that tap the combined processing and storage resources of working memory are better predictors of comprehension than are measures that tap only the storage resources. Whereas the verbal process plus storage measures (the reading spans and listening spans) correlated .41 and .52 with the global and specific measures of comprehension, respectively, verbal storage measures (word span and letter span) correlated only .28 and .40 with the global and specific measures of comprehension, respectively. Thus, for both the global and specific measures of comprehension, the average effect size was greater for the verbal process plus storage measures than for the verbal storage alone measures. This conclu-

**Table 6**  
Average Weighted Effect Size ( $r$ ) and  
95% Confidence Interval (CI) for  
Each Working-Memory/Comprehension Association

	Global Comprehension		Specific Comprehension	
	$r$	CI	$r$	CI
Process and storage				
Verbal	.41	.38–.44	.52	.49–.55
Math	.30	.25–.35	.48	.43–.53
Storage alone				
Verbal	.28	.23–.33	.40	.34–.46
Math	.14	.10–.18	.30	.25–.35

sion is supported by the fact that the 95% confidence intervals are nonoverlapping (see Table 6).<sup>5</sup>

The superior predictive power of the verbal process plus storage measures over the simple verbal storage measures does not appear to be attributable to the fact that the process plus storage measures are more reliable measures than the simple span measures. Because the reliability of a measure sets an upper limit on the extent to which that measure can correlate with other measures (Spearman, 1904), and because some researchers (e.g., Dempster, 1985) have argued that the storage-alone measures may have only moderate reliability, it is possible that the greater predictive power of the verbal process plus storage measures has more to do with its greater reliability than with the type of working-memory capacity it taps. We were able to evaluate this possibility by conducting a small meta-analysis of the reliability data reported in a subset of the studies included in the overall meta-analysis. Although most of the studies included in the meta-analysis did not report reliability coefficients for their working-memory measures, there is a subset of studies by Engle and colleagues (Engle, Cantor, & Carullo, 1992; Engle, Carullo, & Collins, 1991; La Pointe & Engle, 1990; Turner & Engle, 1989) that not only reported reliability estimates for their working-memory measures, but used the same procedure for estimating reliability across studies (Cronbach's alpha, a measure of internal consistency). Using the same meta-analytical procedures as used in the overall analysis, we computed the average weighted reliability estimates across these studies for both types of working-memory measures. The results of this analysis showed very similar reliability estimates for the two types of working-memory measures. The average weighted reliability estimate for the verbal process plus storage measures was .80 ( $n = 473$ ); the average weighted reliability estimate for the verbal storage measures was .79 ( $n = 523$ ). These results argue against differences in reliability being responsible for the greater predictive power of the verbal process plus storage measures of working-memory capacity.

The superior predictive power of the verbal process plus storage measures over the simple verbal storage measures also does not appear to be attributable to the fact that a number of the studies in the meta-analysis included participants whose ages varied quite widely. To make sure that our correlations between working-memory capacity and comprehension were not unduly influenced by age effects, we conducted a smaller meta-analysis in which we only included studies that had homogeneous ages in their sample (e.g., all college students, all elderly participants, all elementary school students) or that provided separate correlations for each age group in the study. The pattern of correlations for this meta-analysis (which included 64 independent samples and 4,982 participants) was very similar to the overall meta-analysis presented in Table 6. In this analysis, the verbal process plus storage measures correlated .40 ( $\pm .03$ ) and .55 ( $\pm .03$ ) with the global and specific tests of comprehen-

sion, respectively, a pattern that was very close to the pattern in the overall meta-analysis in Table 6.<sup>6</sup> And in this analysis, the verbal-storage-alone measures correlated .28 ( $\pm .05$ ) with global comprehension and only .18 ( $\pm .05$ ) with specific comprehension, showing again that the verbal process plus storage measures are better predictors of comprehension than are the verbal-storage-alone measures.

The meta-analysis also shows that the superior predictive power of measures that tap the combined processing and storage resources of working memory is not limited to measures that involve the processing and comprehension of words or sentences. Table 6 shows that the math process plus storage measures are better predictors of comprehension than are their math-storage-alone counterparts. Whereas the math process plus storage measures correlated .30 and .48 with the global and specific measures of comprehension, respectively, the math-storage-alone measures (traditional digit span) correlated only .14 and .30 with the global and specific measures of comprehension, respectively. Thus, for both the global and specific measures of comprehension, the average effect size was greater for the math process plus storage measures than for the math-storage-alone measures, a finding supported by the fact that the confidence intervals are nonoverlapping (see Table 6).

How do the math process plus storage measures compare with the verbal process plus storage measures in predicting comprehension? The first point to note is that math process plus storage measures are themselves good predictors of comprehension. By showing that the predictive power of the process plus storage measure does not depend on a process component that taps sentence comprehension processes, we believe that the results of the meta-analysis argue against the trivial part/whole account of the working-memory/comprehension correlation, the account that relegates the original Daneman and Carpenter (1980) finding to a simple demonstration that sentence comprehension (i.e., reading span) is correlated with paragraph comprehension (i.e., the criterion comprehension tests). Of course, one could argue that the math process plus storage tasks predict comprehension simply because they also tap verbal processes to some extent; when one executes numerical computations (such as mental arithmetic), one typically operates on verbally coded numbers (Logie, Gilhooly, & Wynn, 1994). However, even though the math process plus storage measures tap some verbal processes, they do not tap the entire range of processes involved in sentence comprehension that are tapped by the verbal process plus storage measures (e.g., reading span). Consequently, we believe that the finding of a substantial correlation between math process plus storage measures and comprehension goes well beyond a simple demonstration that sentence comprehension correlates with paragraph comprehension. The finding suggests that it is an individual's efficiency at executing a variety of symbolic manipulations and computations that is related to comprehension ability.

At the other extreme, one could challenge our conclusion not on the basis of the math process plus storage measures' sharing too many processes in common with the verbal process plus storage measures, but on the basis that they are totally unrelated measures. In other words, the correlation between math process plus storage measures and comprehension could result from factors totally unrelated to those responsible for the correlation between verbal process plus storage measures and comprehension; if this were the case, the math process plus storage/comprehension correlation would tell us nothing about the nature of the verbal process plus storage/comprehension correlation. To rule out this possibility, one would need to demonstrate that the math process plus storage measures and verbal process plus storage measures are correlated with one another, and that their shared variance predicts comprehension. Of course, meta-analytic techniques do not allow us to compute partial correlations. However, individual studies have shown math and verbal working memory measures to be correlated with one another (e.g., Daneman & Tardif, 1987; Kyllonen, 1993; Turner & Engle, 1989), and the correlation between math span and comprehension or between verbal span and comprehension is substantially lower when the effects of the other span measure are statistically removed (Daneman & Tardif, 1987). Consequently, we feel confident in concluding that the math and verbal measures draw on the same limited-capacity working-memory system.

Even though a working-memory measure does not require a process component that taps verbal processes to predict comprehension, the meta-analysis provides some evidence to suggest that the domain of the process component does make a difference. As seen in Table 6, the meta-analysis shows that the verbal process plus storage measures of working-memory capacity are better predictors of global comprehension (.41) than are the math process plus storage measures (.30). The verbal process plus storage measures of working-memory capacity tend to be better predictors of performance on specific tests of comprehension (.52) than are the math process plus storage measures (.48), although here there is an overlap in confidence intervals. The superior predictive power of the verbal process plus storage measures over the math process plus storage measures is unlikely to be attributable to differences in the reliabilities of the two kinds of measures. Our small meta-analysis of the reliability data reported by Engle and colleagues (Engle et al., 1992; Engle et al., 1991; La Pointe & Engle, 1990; Turner & Engle, 1989) showed very similar reliability estimates for the verbal and math working-memory measures. The average weighted reliability estimate for the verbal process plus storage measures was .80 ( $n = 473$ ); the average weighted reliability estimate for the math process plus storage measures was .76 ( $n = 393$ ). These results argue against differences in reliability being responsible for the greater predictive power of the verbal process plus storage measures of working-memory capacity. Thus, whereas a more general measure of an individual's

efficiency at symbol manipulating processes in working memory can be used to predict comprehension performance, a measure of the individual's efficiency at specific linguistic processes will be a more sensitive predictor.<sup>7</sup>

So far we have focused heavily on the process side of the working-memory equation, as if the domain of the storage component were not relevant. However, this does not appear to be the case if one compares the relative predictive powers of the verbal- versus math-storage-alone tasks. As Table 6 shows, the verbal storage tasks (the word spans and letter spans) are better predictors of verbal comprehension (.28 and .40) than are the math storage tasks (digit span, .14 and .30), a finding that suggests that the domain of the storage component is relevant too.

Further support for the relevance of the domain of the storage component comes from a subsidiary analysis that we performed on a subset of studies involving the math process plus storage measures. This subsidiary analysis suggests that the predictive power of the math process plus storage measures is better if these measures include a *verbal* storage component rather than a math storage component. Although most of the math process plus storage tasks in the overall meta-analysis involved the processing and storage of digits, one task involved the processing of digits but the storage of words. As Table 3 shows, there were two versions of Turner and Engle's (1989) math span tasks; in one version, participants verified an arithmetic equation [e.g.,  $(2 \times 3) - 2 = 4$ ] and then recalled the product of that operation (i.e., the stated solution, 4); in the second version, they verified the same arithmetic operation, but instead of recalling the stated solution, they recalled a word that accompanied the equation (e.g., *tree*). In the overall meta-analysis presented in Table 6, both versions of Turner and Engle's (1989) math span tasks were categorized as math process plus storage measures because the categorization system was dictated by the domain of the process component. In order to determine whether the domain of the storage component affects the predictive power of the math process plus storage measures, we computed separate average effect sizes for the global comprehension data, depending on whether the storage component of the task required word recall or digit recall.<sup>8</sup> Although we should treat these results as tentative because the number of independent samples is small relative to those of the overall meta-analysis, the pattern of results suggests that the predictive power of the math process plus storage measures is influenced by the domain of the storage component. If the math process measure is accompanied by a verbal storage component, the predictive power is .37 ( $N = 5$ ,  $n = 490$ ,  $CI = .29-.45$ ), a predictive power that approximates that of the verbal process plus storage measures (.41). In contrast, if the math process measure is accompanied by a math storage component, its predictive power is only .25 ( $N = 11$ ,  $n = 936$ ,  $CI = .19-.31$ ), a predictive power no better than that of the traditional word span or letter span tasks (.28). These results, together with the



results of the overall meta-analysis (Table 6), illustrate that one gets the best predictive power for measures that tap verbal processes *and* verbal storage (e.g., reading span), and the poorest predictive power for measures that have no process component and a math storage component (e.g., digit span).

## CONCLUSIONS

The meta-analysis supports Daneman and Carpenter's (1980) contention that working memory plays an important role in language comprehension (see also Carpenter & Just, 1989; Gathercole & Baddeley, 1993; Just & Carpenter, 1992). It also supports Daneman and Carpenter's (1980) approach for assessing an individual's working-memory capacity during language comprehension by showing that measures that tax the combined processing and storage resources of working memory are better predictors of language comprehension performance than are measures that tax only storage. The meta-analysis also rules out the "trivial" account of the working-memory/comprehension correlation. The predictive power of the working-memory measure does not depend on a process component that taps *language* comprehension processes; as long as the processes involve the manipulation of symbolic information (i.e., words or digits), the process plus storage span measures are better than the traditional storage span measures at predicting comprehension. However, to achieve the best predictive validity, the meta-analysis indicates that the working-memory measure should include a verbal process component *and* a verbal storage component.

## REFERENCES

- BADDELEY, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society*, **302**, 311-324.
- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- BADDELEY, A. D., & HITCH, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- BADDELEY, A. D., LOGIE, R. H., NIMMO-SMITH, I., & BRERETON, N. (1985). Components of fluent reading. *Journal of Memory & Language*, **24**, 119-131.
- BAKER, L. (1985). Working memory and comprehension: A replication. *Bulletin of the Psychonomic Society*, **23**, 28-30.
- BOCK, K., & MILLER, C. A. (1991). Broken agreement. *Cognitive Psychology*, **23**, 45-93.
- BYRD, M. (1993). Adult age differences in the ability to write prose passages. *Educational Gerontology*, **19**, 375-393.
- CALVO, M. G., RAMOS, P. M., & ESTEVEZ, A. (1992). Test anxiety and comprehension efficiency: The role of prior knowledge and working memory deficits. *Anxiety, Stress, & Coping*, **5**, 125-138.
- CANTOR, J., ENGLE, R. W., & HAMILTON, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence*, **15**, 229-246.
- CARIGLIA-BULL, T., & PRESSLEY, M. (1990). Short-term memory differences between children predict imagery effects when sentences are read. *Journal of Experimental Child Psychology*, **49**, 384-398.
- CARPENTER, P. A., & JUST, M. A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Handbook of psycholinguistics* (pp. 1075-1122). San Diego, CA: Academic Press.
- CHEN, H.-C. (1986). Effects of reading span and textual coherence on rapid-sequential reading. *Memory & Cognition*, **14**, 202-208.
- CHERRY, K. E., & PARK, D. C. (1993). Individual difference and contextual variables influence spatial memory in younger and older adults. *Psychology & Aging*, **8**, 517-526.
- CLARKSON-SMITH, L., & HARTLEY, A. A. (1990). The game of bridge as an exercise in working memory and reasoning. *Journals of Gerontology*, **45**, P233-P238.
- DANEMAN, M. (1982). The measurement of reading comprehension: How not to trade construct validity for predictive power. *Intelligence*, **6**, 331-345.
- DANEMAN, M. (1987). Reading and working memory. In J. Beech & A. Colley (Eds.), *Cognitive approaches to reading* (pp. 57-86). Chichester, U.K.: Wiley.
- DANEMAN, M. (1988a). How reading braille is both like and unlike reading print. *Memory & Cognition*, **16**, 497-504.
- DANEMAN, M. (1988b). Word knowledge and reading skill. In M. Daneman, G. E. MacKinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 6, pp. 145-175). San Diego: Academic Press.
- DANEMAN, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, **20**, 445-464.
- DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.
- DANEMAN, M., & CARPENTER, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 561-584.
- DANEMAN, M., & GREEN, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory & Language*, **25**, 1-18.
- DANEMAN, M., NEMETH, S., STANTON, M., & HUELSMANN, K. (in press). Working memory as a predictor of reading achievement in orally educated hearing-impaired children. *Volta Review*.
- DANEMAN, M., & TARDIF, T. (1987). Working memory and reading skill re-examined. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 491-508). London: Erlbaum.
- DAS-SMAAL, E. A., DE JONG, P. F., & KOOPMANS, J. R. (1993). Working memory, attentional regulation, and the star counting test. *Personality & Individual Differences*, **14**, 815-824.
- DEMPSTER, F. N. (1985). Short-term memory development in childhood and adolescence. In C. J. Brainerd & M. Pressley (Eds.), *Basic processes in memory development: Progress in cognitive development research* (pp. 209-248). New York: Springer-Verlag.
- DIXON, P., LEFEVRE, J., & TWILLEY, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, **80**, 465-472.
- DREVENSTEDT, J., & BELLEZZA, F. S. (1993). Memory for self-generated narration in the elderly. *Psychology & Aging*, **8**, 187-196.
- ENGLE, R. W., CANTOR, J., & CARULLO, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 972-992.
- ENGLE, R. W., CARULLO, J. J., & COLLINS, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, **84**, 253-262.
- ENGLE, R. W., NATIONS, J. K., & CANTOR, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology*, **82**, 799-804.
- FARNHAM-DIGGORY, S., & GREGG, L. W. (1975). Short-term memory function in young readers. *Journal of Experimental Child Psychology*, **19**, 279-298.
- FRISK, V., & MILNER, B. (1990). The relationship of working memory to the immediate recall of stories following unilateral temporal or frontal lobectomy. *Neuropsychologia*, **28**, 121-135.
- GATHERCOLE, S. E., & BADDELEY, A. D. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.
- GAULIN, C. A., & CAMPBELL, T. F. (1994). Procedure for assessing verbal working memory in normal school-age children: Some preliminary data. *Perceptual & Motor Skills*, **79**, 55-64.

- HAENGGLI, D., & PERFETTI, C. A. (1994). Processing components of college-level reading comprehension. *Discourse Processes*, 17, 83-104.
- HARTLEY, J. (1986). Reader and text variables as determinants of discourse memory in adulthood. *Psychology & Aging*, 1, 150-158.
- HARTLEY, J. (1988). Aging and individual differences in memory for written discourse. In L. L. Light & D. M. Burke (Eds.), *Language, memory, and aging* (pp. 36-57). Cambridge: Cambridge University Press.
- HEDGES, L. V., & OLKIN, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- HESS, T. M., & TATE, C. S. (1992). Direct and indirect assessments of memory for script-based narratives in young and older adults. *Cognitive Development*, 7, 467-484.
- HOLLAND, C. A., & RABBITT, P. M. A. (1990). Autobiographical and text recall in the elderly: An investigation of a processing resource deficit. *Quarterly Journal of Experimental Psychology*, 42A, 441-470.
- HUNT, E. (1980). Intelligence as an information-processing concept. *British Journal of Psychology*, 71, 449-474.
- HUPERT, M., NEF, F., & MAROY, M. (1992). Étude comparative du langage spontané d'adultes jeunes et âgés. *L'Année Psychologique*, 92, 511-526.
- JACKSON, M. D., & MCCLELLAND, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General*, 108, 151-181.
- JURDEN, F. H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology*, 87, 93-102.
- JURDEN, F. H., & REESE, H. W. (1992). Educational context differences in prose recall in adulthood. *Journal of Genetic Psychology*, 153, 275-291.
- JUST, M. A., & CARPENTER, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 4, 329-354.
- JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- KING, J., & JUST, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory & Language*, 30, 580-602.
- KINTSCH, W., & VAN DIJK, T. A. (1978). Toward a model of text comprehension. *Psychological Review*, 85, 363-394.
- KYLLONEN, P. C. (1993). Aptitude testing inspired by information processing: A test of the four-sources model. *Journal of General Psychology*, 120, 375-405.
- KYLLONEN, P. C., & CHRISTAL, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence*, 14, 389-433.
- LA POINTE, L. B., & ENGLE, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 1118-1133.
- LEATHER, C. V., & HENRY, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. *Journal of Experimental Child Psychology*, 58, 88-111.
- LEE-SAMMONS, W. H., & WHITNEY, P. (1991). Reading perspectives and memory for text: An individual differences analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17, 1074-1081.
- LEHRER, R., GUCKENBERG, T., & LEE, O. (1988). Comparative study of cognitive consequences of inquiry-based logo instruction. *Journal of Educational Psychology*, 80, 543-553.
- LEHRER, R., & LITTLEFIELD, J. (1993). Relationships among cognitive components in logo learning and transfer. *Journal of Educational Psychology*, 85, 317-330.
- LIGHT, L. L., & ANDERSON, P. A. (1985). Working-memory capacity, age, and memory for discourse. *Journal of Gerontology*, 40, 737-747.
- LOGIE, R. [H.], BADDELEY, A. D., MANÉ, A. M., DONCHIN, E., & SHEP-TAK, R. (1989). Working memory in the acquisition of complex cognitive skills. *Acta Psychologica*, 71, 53-87.
- LOGIE, R. H., GILHOOLY, K. J., & WYNN, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, 22, 395-410.
- LYXELL, B., & RÖNNBERG, J. (1993). The effects of background noise and working memory capacity on speechreading performance. *Scandinavian Audiology*, 22, 67-70.
- MASSON, M. E. J., & MILLER, J. A. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology*, 75, 314-318.
- MILLER, G. A., GALANTER, E., & PRIBRAM, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.
- MITCHELL, D. C. (1982). *The process of reading: A cognitive analysis of fluent reading and learning to read*. New York: Wiley.
- MORROW, D. G., LEIRER, V. O., & ALTIERI, P. A. (1992). Aging, expertise, and narrative processing. *Psychology & Aging*, 7, 376-388.
- NORMAN, S., KEMPER, S., & KYNETTE, D. (1992). Adults' reading comprehension: Effects of syntactic complexity and working memory. *Journals of Gerontology*, 47, P258-P265.
- PERFETTI, C. A., & GOLDMAN, S. R. (1976). Discourse memory and reading comprehension skill. *Journal of Verbal Learning & Verbal Behavior*, 15, 33-42.
- PERFETTI, C. A., & LESGOLD, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141-183). Hillsdale, NJ: Erlbaum.
- PRATT, M. W., BOYES, C., ROBINS, S., & MANCHESTER, J. (1989). Telling tales: Aging, working memory, and the narrative cohesion of story retellings. *Developmental Psychology*, 25, 628-635.
- RADVANSKY, G. A., GERARD, L. D., ZACKS, R. T., & HASHER, L. (1990). Younger and older adults' use of mental models as representations for text materials. *Psychology & Aging*, 5, 209-214.
- RANKIN, J. L. (1993). Information-processing differences of college-age readers differing in reading comprehension and speed. *Journal of Reading Behavior*, 25, 261-278.
- RIZZO, N. D. (1939). Studies in visual and auditory memory span with special reference to reading disability. *Journal of Experimental Education*, 8, 208-244.
- RÖNNBERG, J., ARLINGER, S., LYXELL, B., & KINNEFORS, C. (1989). Visual evoked potentials: Relative to adult speechreading and cognitive functions. *Journal of Speech & Hearing Research*, 32, 725-735.
- ROSENTHAL, R. (1991). *Meta-analytic procedures for social research* (rev. ed.). Newbury Park, CA: Sage.
- SALTHOUSE, T. A., & BABCOCK, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 763-776.
- SALTHOUSE, T. A., & KERSTEN, A. W. (1993). Decomposing adult age differences in symbol arithmetic. *Memory & Cognition*, 21, 699-710.
- SALTHOUSE, T. A., MITCHELL, D. R. D., SKOVRONEK, E., & BABCOCK, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 507-516.
- SHAH, P., & MIYAKE, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125, 4-27.
- SHUTE, V. J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research*, 7, 1-24.
- SINGER, M., ANDRUSIAK, P., REISDORF, P., & BLACK, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition*, 20, 539-548.
- SPEARMAN, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72-101.
- STAYER, J. R., & JACKS, T. (1988). The influence of cognitive reasoning level, cognitive restructuring ability, disembedding ability, working memory capacity, and prior knowledge on students' performance on balancing equations by inspection. *Journal of Research in Science Teaching*, 25, 763-775.
- STINE, E. A. L., LACHMAN, M. E., & WINGFIELD, A. (1993). The roles of perceived and actual control in memory for spoken language. *Educational Gerontology*, 19, 331-349.
- STINE, E. [A.] L., & WINGFIELD, A. (1987). Process and strategy in memory for speech among younger and older adults. *Psychology & Aging*, 2, 272-279.
- STINE, E. A. L., WINGFIELD, A., & MYERS, S. D. (1990). Age differences

- in processing information from television news: The effects of bisensory augmentation. *Journals of Gerontology*, **45**, P1-P8.
- SWANSON, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, **84**, 473-488.
- SWANSON, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, **56**, 87-114.
- SWANSON, H. L., COCHRAN, K. F., & EWERS, C. A. (1989). Working memory in skilled and less skilled readers. *Journal of Abnormal Child Psychology*, **17**, 145-156.
- TAYLOR, J. L., YESAVAGE, J. A., MORROW, D. G., DOLHERT, N., BROOKS, J. O., & POON, L. W. (1994). The effects of information load and speech rate on younger and older aircraft pilots' ability to execute simulated air-traffic controller instructions. *Journals of Gerontology*, **49**, P191-P200.
- TIRRE, W. C., & PENA, C. M. (1992). Investigation of functional working memory in the reading span test. *Journal of Educational Psychology*, **84**, 462-472.
- TOMPKINS, C. A., BLOISE, C. G. R., TIMKO, M. L., & BAUMGAERTNER, A. (1994). Working memory and inference revision in brain-damaged and normally aging adults. *Journal of Speech & Hearing Research*, **37**, 896-912.
- TUN, P. A., WINGFIELD, A., & STINE, E. A. L. (1991). Speech-processing capacity in young and older adults: A dual-task study. *Psychology & Aging*, **6**, 3-9.
- TURNER, M. L., & ENGLE, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, **28**, 127-154.
- WINGFIELD, A., STINE, E. A. L., LAHAR, C. J., & ABERDEEN, J. S. (1988). Does the capacity of working memory change with age? *Experimental Aging Research*, **14**, 103-107.

### NOTES

1. A few studies have examined the association between working-memory capacity and comprehension *speed*; however, because there have not been a sufficient number of these studies, they were not included in the meta-analysis.
2. A few studies have investigated the association between spatial working-memory capacity and language comprehension; however, because there are so few of these studies, they were not included in the formal meta-analysis.
3. In fact, whenever confidence intervals do not overlap, a *t* test comparing the two correlation coefficients will reveal a statistically significant difference.
4. The comparable average weighted effect sizes in Daneman and Carpenter's (1980) Experiments 1 and 2 were .56 and .76, for global and specific comprehension, respectively.
5. It is worth noting that although the results of the meta-analysis support Daneman and Carpenter's (1980) claim that measures of the combined processing and storage resources of working memory are better predictors of comprehension than are measures of storage alone, they do not support Daneman and Carpenter's (1980) finding that verbal-storage-alone measures are not significant predictors of language comprehension. Indeed, the results of the meta-analysis show a rather respectable correlation between verbal-storage-alone measures and specific tests of comprehension.
6. The number in parenthesis following an effect size indicates the confidence limits for the 95% confidence interval.
7. Further support for the relevance of the domain of the process component comes from a few studies of spatial working memory that were not included in the formal meta-analysis (e.g., Daneman & Tardif, 1987; Kyllonen, 1993; Shah & Miyake, 1996). The results of these studies suggest that verbal process plus storage measures are better predictors of comprehension than are spatial process plus storage measures.
8. A comparable analysis could not be done for the specific comprehension data because the studies that used the math process-verbal storage version of Turner and Engle's (1989) task used only global/standardized tests of language comprehension as criterion tasks.

### APPENDIX

#### Effect Size and Number of Participants for Each Sample in Each Cell of the Meta-Analysis

Study	<i>r</i>	<i>n</i>
Verbal Process Plus Storage/Global Comprehension		
Baddeley, Logie, Nimmo-Smith, & Brereton (1985): Experiment 1	.40	51
Baddeley, Logie, Nimmo-Smith, & Brereton (1985): Experiment 2	.31	102
Baker (1985)	.67	28
Calvo, Ramos, & Estevez (1992)	.32	36
Cantor, Engle, & Hamilton (1991)	.42	49
Cherry & Park (1993): young adults	.05	64
Cherry & Park (1993): old—low income	.51	64
Cherry & Park (1993): old—univ ed	.30	64
Daneman & Carpenter (1980): Experiment 1	.59	20
Daneman & Carpenter (1980): Experiment 2	.53	21
Daneman & Carpenter (1983): Experiment 1	.58	26
Daneman & Carpenter (1983): Experiment 2	.46	27
Daneman & Green (1986)	.57	30
Daneman & Tardif (1987)	.56	36
Dixon, LeFevre, & Twilley (1988)	.37	95
Drevenstedt & Bellezza (1993)	.50	67
Engle, Cantor, & Carullo (1992): Experiment 2	.40	70
Engle, Carullo, & Collins (1991): Grade 3	.53	40
Engle, Carullo, & Collins (1991): Grade 6	.44	40
Gaulin & Campbell (1994)	.63	68
Haenggi & Perfetti (1974)	.76	34
Hartley (1986)	.47	72
Holland & Rabbitt (1990)	.21	33
Jurden (1995): Experiment 1	.28	84
Jurden (1995): Experiment 2	.41	52
Kyllonen & Christal (1990): Study 1	.37	723
La Pointe & Engle (1990): Experiment 1	.54	37
Leather & Henry (1994)	.51	71
Masson & Miller (1983)	.53	29
Norman, Kemper, & Kynette (1992)	.30	89
Radvansky, Gerard, Zacks, & Hasher (1990)	.43	32
Rankin (1993)	.39	63
Singer, Andrusiak, Reisdorf, & Black (1992)	.26	135
Swanson (1992)	.58	98
Swanson, Cochran, & Ewers (1989)	.60	50
Turner & Engle (1989): Experiment 1	.33	243
Turner & Engle (1989): Experiment 2	.32	50
Wingfield, Stine, Lahar, & Aberdeen (1988)	.33	68
Verbal Process Plus Storage/Specific Comprehension		
Baker (1985)	.50	28
Calvo, Ramos, & Estevez (1992)	.28	36
Chen (1986)	.46	32
Daneman (1988b)	.66	49
Daneman & Carpenter (1980): Experiment 1	.83	20
Daneman & Carpenter (1980): Experiment 2	.70	21
Daneman & Carpenter (1983): Experiment 1	.71	32
Daneman & Carpenter (1983): Experiment 2	.60	27
Daneman & Green (1986)	.69	30
Dixon, LeFevre, & Twilley (1988)	.28	95
Drevenstedt & Bellezza (1993)	.56	67
Engle, Carullo, & Collins (1991): Grade 3	.27	40
Engle, Carullo, & Collins (1991): Grade 6	.48	40
Haenggi & Perfetti (1994)	.68	34
Hartley (1986)	.03	72
Hartley (1988)	.33	94
Hess & Tate (1992)	.32	65
Holland & Rabbitt (1990)	.23	33
Jurden & Reese (1992)	.50	36
King & Just (1991): Experiment 1	.54	46

## APPENDIX (Continued)

Study	<i>r</i>	<i>n</i>	Study	<i>r</i>	<i>n</i>
King & Just (1991): Experiment 2	.31	48	Engle, Nations, & Cantor (1990)	.55	90
Kyllonen & Christal (1990): Study 1	.61	723	Farnham-Diggory & Gregg (1975)	.11	24
Lee-Sammons & Whitney (1991)	.73	48	Jackson & McClelland (1979)	.37	24
Light & Anderson (1985): Experiment 1	.34	50	La Pointe & Engle (1990): Experiment 1	.37	37
Light & Anderson (1985): Experiment 2	.12	40	La Pointe & Engle (1990): Experiment 2	.49	37
Masson & Miller (1983)	.56	29	Leather & Henry (1994)	.27	71
Morrow, Leirer, & Altieri (1992)	.27	75	Masson & Miller (1983)	.03	29
Norman, Kemper, & Kynette (1992)	.35	89	Rizzo (1939): Grade 3	.26	42
Pratt, Boyes, Robins, & Manchester (1989)	.37	60	Rizzo (1939): Grade 4	.23	40
Salthouse & Babcock (1991): Experiment 1	.62	227	Rizzo (1939): Grade 5	.18	44
Salthouse & Babcock (1991): Experiment 2	.49	233	Rizzo (1939): Grade 6	.57	43
Stine, Lachman, & Wingfield (1993)	.63	32	Rizzo (1939): Grade 8	.28	29
Stine & Wingfield (1987)	.66	48	Rizzo (1939): Grade 10	.16	38
Stine, Wingfield, & Myers (1990)	.62	24	Rizzo (1939): Grade 12	.19	38
Tirre & Pena (1992)	.62	86	Swanson (1992)	.33	98
Tun, Wingfield, & Stine (1991)	.55	36	Turner & Engle (1989)	.08	243
Math Process Plus Storage/Global Comprehension			Wingfield, Stine, Lahar, & Aberdeen (1988)	.18	68
Baddeley, Logie, Nimmo-Smith, & Brereton (1985)			Verbal Storage Alone/Specific Comprehension		
Cherry & Park (1993): young adults	.09	102	Calvo, Ramos, & Estevez (1992)	.22	36
Cherry & Park (1993): old—low income	.51	64	Daneman & Carpenter (1980): Experiment 1	.35	20
Cherry & Park (1993): old—univ ed	.33	64	Dixon, LeFevre, & Twilley (1988)	.09	95
Daneman & Tardif (1987)	.47	36	Engle, Carullo, & Collins (1991): Grade 3	.48	40
Drevenstedt & Bellezza (1993)	.44	67	Engle, Carullo, & Collins (1991): Grade 6	.51	40
Engle, Cantor, & Carullo (1992): Experiment 1	.34	70	Kyllonen & Christal (1990): Study 1	.28	723
Engle, Nations, & Cantor (1990)	.41	90	Light & Anderson (1985): Experiment 1	.25	50
Jurden (1995): Experiment 1	.03	84	Light & Anderson (1985): Experiment 2	-.11	40
Jurden (1995): Experiment 2	.20	52	Masson & Miller (1983)	.04	29
La Pointe & Engle (1990): Experiment 2	.54	37	Salthouse & Babcock (1991): Experiment 1	.56	227
Leather & Henry (1994)	.27	71	Salthouse & Babcock (1991): Experiment 2	.48	233
Norman, Kemper, & Kynette (1992)	.30	89	Math Storage Alone/Global Comprehension		
Turner & Engle (1989): Experiment 1	.28	243	Cantor, Engle, & Hamilton (1991)	.04	49
Turner & Engle (1989): Experiment 2	.20	50	Dixon, LeFevre, & Twilley (1988)	.04	95
Math Process Plus Storage/Specific Comprehension			Gaulin & Campbell (1994)	.35	68
Drevenstedt & Bellezza (1993)	.52	67	Kyllonen & Christal (1990): Study 1	.12	723
Light & Anderson (1985)	.29	50	Kyllonen & Christal (1990): Study 2	.17	399
Light & Anderson (1985)	.05	40	Kyllonen & Christal (1990): Study 4	.12	562
Norman, Kemper, & Kynette (1992)	.39	89	Norman, Kemper, & Kynette (1992)	.23	89
Salthouse & Babcock (1991): Experiment 1	.57	227	Perfetti & Goldman (1976)	.15	48
Salthouse & Babcock (1991): Experiment 2	.50	233	Turner & Engle (1989): Experiment 1	.11	243
Salthouse, Mitchell, Skovronek, & Babcock (1989)	.48	120	Wingfield, Stine, Lahar, & Aberdeen (1988)	.16	68
Taylor et al. (1994)	.56	30	Math Storage Alone/Specific Comprehension		
Verbal Storage Alone/Global Comprehension			Dixon, LeFevre, & Twilley (1988)	.02	95
Calvo, Ramos, & Estevez (1992)	.25	36	Light & Anderson (1985): Experiment 1	.26	50
Cariglia-Bull & Pressley (1990): imagery group	.34	84	Light & Anderson (1985): Experiment 2	.09	40
Cariglia-Bull & Pressley (1990): control group	.46	81	Norman, Kemper, & Kynette (1992)	.21	89
Daneman & Carpenter (1980): Experiment 1	.35	20	Salthouse & Babcock (1991): Experiment 1	.42	227
Dixon, LeFevre, & Twilley (1988)	.08	95	Salthouse & Babcock (1991): Experiment 2	.39	233
Engle, Cantor, & Carullo (1992): Experiment 3	.49	40	Taylor et al. (1994)	.47	30
Engle, Carullo, & Collins (1991): Grade 3	.36	40			
Engle, Carullo, & Collins (1991): Grade 6	.41	40			

(Manuscript received September 1, 1995;  
revision accepted for publication June 6, 1996.)