6

# The Role of Working Memory in Higher-Level Cognition

*Domain-Specific versus Domain-General Perspectives*

David Z. Hambrick, Michael J. Kane, and Randall W. Engle

## INTRODUCTION

The idea that short-term memory is an important component of intelligence is not new. For example, over a century ago James (1890) wrote, "All the intellectual value for us of a state of mind depends on our after memory of it. Only then is it combined in a system and knowingly made to contribute to a result. Only then does it *count* for us." Around the same time, Binet (1905) included a test of short-term memory in a test battery designed to identify learning disabled children in the Paris school system. And more recently, short-term memory has been conceptualized as a fundamental component of human cognition. For example, Miller (1956) famously proposed that the capacity of short-term memory is limited to $7 \pm 2$ bits of information. Later, Atkinson and Shiffrin (1968) incorporated this idea of a central bottleneck in information processing into their "modal" model of memory.

Nevertheless, the extent to which short-term memory plays an important role in higher-level cognition – intelligence manifested in complex cognitive activities like reasoning and learning – has been a topic of considerable debate in cognitive psychology. Consider, for example, the results of a series of experiments by Baddeley and Hitch (1974). The surprising finding in these experiments was that a secondary task designed to tax short-term memory had little or no effect on a variety of reasoning, comprehension, and memory primary tasks. In one such experiment, subjects performed a task in which the goal was to verify sentences purporting to describe the relationship between two letters (e.g., A precedes B – BA) while maintaining a memory load. The secondary task had little effect on subjects' success in the task – a finding logically inconsistent with the assumption of short-term memory as a central bottleneck in information processing.

Baddeley and Hitch (1974) therefore proposed that short-term memory – the passive storage of information – is but one part of a memory system in which a limited capacity "workspace" can be divided between processing and storage functions. This concept provided a tidy explanation for their findings. Subjects were able to divide this limited capacity workspace between the primary task and the secondary task, as long as the latter did not overtax the system. Following this initial work, Baddeley and his colleagues proposed a working memory model consisting of three major components: two "slave" systems – the *phonological loop* and *visuospatial sketchpad* – devoted to temporary storage and maintenance of information and a *central executive* responsible for planning and control processes involved in higher-level cognition (e.g., Baddeley, 1986). Understanding the nature of this latter component of the system and its involvement in higher-level cognition has since been a major focus of research in cognitive psychology.

## AN INDIVIDUAL-DIFFERENCES PERSPECTIVE ON WORKING MEMORY

In the early 1980s, research on individual differences in working memory (WM) took off with the development of a procedure for measuring the construct – the Daneman and Carpenter (1980) reading span task. Consistent with Baddeley and Hitch's (1974) conception of the central executive, Daneman and Carpenter designed this task to include both a processing component – reading sentences – and a storage component – remembering the final word of each sentence for later recall. For example, given the sentences *When at last his eyes opened, there was no gleam of triumph, no shade of anger* and *The taxi turned up Michigan Avenue where they had a clear view of the lake*, the task would be to report *anger* and *lake*. Daneman and Carpenter discovered that *reading span* – the number of sentences a subject could read while maintaining perfect recall of the sentence-final words – correlated with global measures of language comprehension (e.g., verbal SAT score) as well as with specific measures (e.g., resolving pronominal ambiguity). Moreover, reading span was a better predictor of comprehension than was a measure of short-term memory (word span).

A variety of WM tasks modeled after reading span have been introduced since Daneman and Carpenter's (1980) study. Like reading span, each of these tasks is a dual task in the sense that it involves alternating between interleaved processing and storage subtasks. To illustrate, in *operation span* (Turner & Engle, 1989), the goal is to solve a series of simple math problems while remembering a word following each problem, whereas in *counting span*, the goal is to count the number of target objects in a series of displays (e.g., light blue circles among dark blue circles and light blue triangles) while remembering the count from each display. Nonverbal WM tasks have been developed as well. For example, Shah and Miyake (1996) introduced a task called *spatial span* in which subjects decide whether each of a series of rotated letters is normal or mirror-imaged while remembering the orientation of each letter.

Two observations can be made from the hundreds of independent studies in which WM tasks have been administered. The first observation is that WM tasks are reliable; that is, these tasks measure accurately *whatever it is that they measure*. For example, with approximately two months between test intervals, Klein and Fiss (1999) reported a test–retest reliability coefficient of .88 for the operation span task. Moreover, internal consistency estimates (e.g., coefficient alphas) for WM tasks are typically in the range from .70 to .90. This evidence can be understood in terms of classical test theory (e.g., Novick, 1966; Spearmen, 1927). The basic assumption of classical test theory is that a single test score consists of a *true score* – which reflects stable characteristics of the attribute one is trying to measure – and *error*. Within this framework, reliability is interpreted as an index of the proportion of variance in test scores (*total variance*) that is caused by variability in true scores (*true-score variance*). Because reliability coefficients of WM tasks are seldom lower than .70, and are often much higher, it therefore appears that scores on these tasks are more attributable to stable characteristics of subjects – to true scores – than to error.

The second observation is that individual differences in WM span correlate with measures of many aspects of higher-level cognition, including reading comprehension (e.g., Daneman & Carpenter, 1980), abstract reasoning (e.g., Kyllonen & Christal, 1990), problem solving (e.g., Welsh, Satterlee-Cartmell, & Stine, 1999), and complex learning (e.g., Kyllonen & Stephens, 1990). Nevertheless, on the basis of the available evidence, it remains unclear *what* various measures of WM reflect and *why* they correlate with higher-level cognition. In other words, although it is evident that WM tasks accurately measure *some* capability that seems to be important for higher-level cognition, what is the nature of this capability? At least two major hypotheses concerning this question have been advanced.

The premise of the first hypothesis is that WM tasks capture factors that are applicable to only a particular task or class of tasks. For example, according to this domain-specific hypothesis, reading span correlates with reading comprehension simply because reading span itself involves reading comprehension. In line with this hypothesis, Daneman and Carpenter (1980) proposed that by virtue of their greater efficiency in the processing component of the reading span task – reading sentences – the high-span individuals in their study had more residual capacity to devote to memorization of the sentence-final words than did the low-spans. Similarly, MacDonald and Christiansen (2002) claimed that "the distinction commonly drawn between language-processing tasks and linguistic WM tasks is an artificial one, and ... all of these tasks are simply different measures of language processing skill" (p. 36) (see also Ericsson & Kintsch, 1995).

In contrast, the second hypothesis proposes that, in addition to any domain-specific factors, WM tasks capture factors that are involved in a wide range of cognitive tasks. In particular, this domain-general hypothesis assumes that there is nothing special about a particular WM task like reading span or operation span. Rather, all WM tasks, regardless of their specific requirements, tap domain-general factors that play a role in many different cognitive tasks. For example, consistent with this hypothesis, we have argued that one domain-general factor captured by WM tasks is the capability for attention control, which we believe underlies the ability to maintain goals and other task-relevant information in a highly activated and accessible state, particularly under conditions of interference or distraction (Engle, Kane, & Tuholski, 1999a; Engle, Tuholski et al., 1999b). As another example, Hasher and Zacks (1988) proposed that individual differences in WM span arise from the efficiency and effectiveness of a number of inhibitory processes that regulate the contents of conscious thought. Although the theoretical mechanisms of these theories differ – ours emphasizes maintenance of task-relevant information whereas theirs emphasizes inhibition of task-irrelevant information – the theories are similar in that both assume that domain-general factors underlie individual differences in WM and its involvement in higher-level cognition.

## Which Hypothesis Is Correct?

Domain-specific factors almost certainly account for some of the true-score variance in WM tasks because, as Spearman (1927) observed, we must assume that performance on any test of mental ability is influenced by factors unique to that test, in addition to any factors that operate across different tests. Stated differently, no task is "process-pure" in the sense that it captures only the task-independent construct of interest. For example, skill in math may contribute to the total variance in operation span, whereas skill in reading may contribute to the total variance in reading span. In fact, dozens of factors may contribute to the total variance in WM span as measured by a particular task. At the same time, evidence suggests that a sizeable proportion of the true-score variance in WM tasks is accounted for by domain-general factors, above and beyond the contribution of any domain-specific factors. For example, in a study by Engle et al. (1999b), subjects completed a battery of WM tasks that included reading span, operation span, and counting span. Even though the requirements of these tasks were quite different, the average inter-task correlation was .43, indicating that an average of 18% of the variance in one task was accounted for by factors operating in the other tasks (i.e., $.43^2 = .184$). Of course, another way to interpret this observation is that 82% of the variance in these tasks was accounted for by factors *not* operating in the other tasks. However, the central claim of the domain-general hypothesis is not that the total variance in WM span is accounted for entirely by domain-general factors – or even mostly – but rather that these factors explain the correlation of WM span with higher-level cognition, with little or no contribution from domain-specific factors. In other words, if the true-score variance in WM span can be decomposed into two types – domain-specific and domain-general – then the prediction is that the latter drives correlations of WM span with higher-level cognition. Evidence from studies that have followed two quite different research approaches supports this conclusion.

### Microanalytic Research

The first approach is *microanalytic* because the goal is to investigate how WM span relates to performance in what might be considered "elementary" attention tasks; that is, tasks designed to capture basic information processes underlying higher-level cognition. This research has revealed that individual differences in these elementary tasks are strongly related to individual differences in a variety of WM tasks, suggesting that the capability for attention control may lie at the heart of individual differences in WM span.

Consider the results of a study by Kane et al. (2001). Subjects classified as either low or high in WM span (*low-span* or *high-span*) performed a version of the so-called antisaccade task. The procedure was simple: In the *prosaccade* condition, a flashing cue appeared in the same location on the screen as an upcoming stimulus – the letter B, P, or R – and the task was to press a key corresponding to the stimulus. By contrast, in the *antisaccade* condition, the target always appeared in the location opposite to that of the cue. The results were straightforward: the advantage of high-spans over low-spans in both reaction time and accuracy was larger in the antisaccade condition than in the prosaccade condition. Moreover, in a follow-up experiment, Kane et al. monitored eye movements and found that this was because low-spans made more reflexive eye movements toward the flashing cue in the antisaccade condition than did high-spans. Similarly, in a study by Schrock and Engle (in preparation), in which the subject simply had to look at a box on the opposite side of the screen from a flashing cue, low-spans were much more likely than high-spans to make their first saccade an erroneous movement to the flashing cue. In fact, even when low-spans were correct in their first saccade, they were slower than high-spans to begin the eye movement.

We believe that the results of these studies provide especially strong support for a domain-general hypothesis of WM because there are no apparent domain-specific factors to which span-related differences in the antisaccade task can be attributed. Results of other studies from our labs are consistent with this hypothesis as well. For example, Kane and Engle (2000) used a three-trial serial recall task in which subjects were presented with three 10-word lists, each of which was followed by a 30-second rehearsal preventative task before recall. As predicted, there was greater buildup of proactive interference in low-spans than in high-spans. One interpretation of this finding is that, after the first trial, high-spans were better able to maintain the words in an activated state than were low-spans and were hence less likely to confuse these words with those from the previous trial or trials. To test this hypothesis, in a second experiment, subjects performed the task as before or while performing a continuous, attention-demanding secondary task. If attention control was responsible for the span-related difference in proactive interference observed in the first experiment, then the secondary task should have produced more of an increase in proactive interference for high-spans than for low-spans. This is what happened; indeed, in the divided-attention condition, the performance of low-spans and high-spans was indistinguishable.

In another microanalytic study, Kane and Engle (2003) used the Stroop task to investigate the possibility that WM span is related to a phenomenon Duncan (1990) termed "goal neglect." The basic idea of goal neglect is that attention failures occur when goal-relevant information is lost from the active portion of memory because the environment lacks external cues for appropriate action. In a series of experiments, Kane and Engle set up this type of situation by manipulating percentages of congruent and incongruent trials in the Stroop task. In the 0% congruent conditions, almost all of the trials were incongruent (e.g., BLUE displayed in red), and so the task context reinforced the goal, to ignore the word, on virtually every trial. By contrast, in the 75% congruent conditions, subjects could neglect the task goal on a majority of trials with no negative consequences. However, accurate responding on the rare incongruent trials here required that subjects maintain access to the *ignore-the-word* goal. Taken together, the results revealed that low-spans were much more error-prone than high-spans in the 75% conditions but not in the 0% conditions. Thus, low-spans were at a disadvantage when the task placed a premium on actively maintaining the goal of ignoring words in a task environment that lacked external prompts to action.

As a final example, Conway, Cowan, and Bunting (2001) found that WM span is related to a phenomenon first reported by Moray (1959). In a series of experiments by Cherry (1953), subjects were instructed to repeat a message presented in one ear and to ignore a message presented in the other ear. Subjects had little difficulty performing this task, and thus theorists such as Broadbent (1958) proposed that attention acts as an all-or-none filter, letting relevant information into short-term memory but blocking out irrelevant information. Nevertheless, Moray demonstrated that content from an unattended message is not rejected completely. In particular, a substantial number of subjects (33%) heard their name when it was presented in the unattended message. By contrast, very few participants could recall a word that was repeated 35 times in the unattended ear. Why, though, did 100% of Moray's subjects not hear their own names? Conway et al. reasoned that if what WM tasks capture is related to the ability to control attention – to direct it toward relevant information and away from irrelevant information – then high-spans would be *less* likely to notice their names in an unattended message than low-spans. Thus, Conway et al. replicated Moray's experiment, but with low-span and high-span subjects. The results were striking: 65% of low-spans heard their names in the unattended message, whereas only 20% of high-spans did so.

*The Role of Strategies?*

We believe that the evidence considered thus far supports the hypothesis that individual differences in various span tasks reflect differences in the capability for attention control, and elsewhere we have argued that this individual-difference characteristic is a relatively stable aspect of cognition (e.g., Engle et al., 1999a). Nevertheless, an alternative hypothesis – and one that is particularly appealing because it implies that deficits in WM can be ameliorated through instruction – posits that these differences stem not from differences in any fixed information processing capacity, but rather from differences in the *strategies* that low-spans and high-spans use to perform the tasks.

Using the reading span task, McNamara and Scott (2001) investigated this possibility by training subjects in the use of a mnemonic technique called "chaining" that involves memorizing words by generating sentences to connect them. McNamara and Scott found that training improved reading span performance by 41% and 53% in two experiments. Moreover, these improvements did not come at the expense of poorer performance in the comprehension component of the reading span task, as comprehension actually improved from pretest to post-test in both experiments. McNamara and Scott concluded that strategy training enhanced subjects' efficiency in performing the reading span task, thereby freeing up resources for use in the comprehension component of the task.

The McNamara and Scott (2001) study convincingly suggests that strategies can influence performance in WM tasks; in addition, this study is important because it highlights the importance of taking into account the possibility of strategy use when assessing WM. Nevertheless, McNamara and Scott's finding is not surprising because many studies have demonstrated beneficial effects of strategy instruction on cognitive performance. For example, a number of researchers have reported that strategy training enhances performance on a task that is regarded as a relatively pure indicator of general intelligence – Raven's Progressive Matrices (e.g., Blieszner, Willis, & Baltes, 1981; Klauer, Willmes, & Phye, 2002; Denney & Heidrich, 1990). There simply is no reason to expect that strategy training would not also enhance WM span. Furthermore, McNamara and Scott did not address the important question of whether differential strategy use by low-spans and high-spans accounts for the correlation of WM span with higher-level cognition.

To answer this question, Turley-Ames and Whitfield (2003) conducted an impressive, large-scale study ($N = 360$) to investigate effects of different types of strategies on the correlation between operation span and reading comprehension. After taking a pretest of operation span, subjects were assigned to a control condition or to a condition in which they were instructed in use of a strategy for the operation span task involving rote rehearsal, visual imagery, or forming semantic associations. Subjects then completed another version of operation span. Consistent with McNamara and Scott's (2001) finding, strategy training enhanced WM performance. However, strategy training did not reduce – much less eliminate – the correlation between operation span and reading comprehension. In fact, at post-test, operation span correlated *more* positively with reading comprehension in each strategy condition – rehearsal ($r = .56$), imagery ($r = .32$), and semantic ($r = .47$) – than in the control condition ($r = .30$).

Therefore, the results of the Turley-Ames and Whitfield (2003) study suggest that differential strategy use by low-spans and high-spans may *suppress* rather than account for the relationship between WM span and higher-level cognition. Results of an earlier study by Engle, Cantor, and Carullo (1992) provide additional support for this conclusion. In this study, using a "moving-window" technique in which elements of either operation span or reading span were presented sequentially rather than simultaneously, Engle et al. measured the amount of time subjects spent on the processing component of the task. They then interpreted this measure as an estimate of the extent to which subjects strategically traded off time on the processing component for time on the storage component. In agreement with Turley-Ames and Whitfield's (2003) finding, for both operation span and reading span, there was no evidence for a decrease in the correlation between WM span and reading comprehension after controlling for this estimate; that is, the correlation increased slightly for operation span (.34 → .40) and was unchanged for reading span (.40 → .40).

To sum up, based on the available evidence, it appears that the main effect of strategy use may be on the total variance in WM performance. That is, as both McNamara and Scott (2001) and Turley-Ames and Whitfield (2003) demonstrated, it seems clear that strategy use can influence scores in WM tasks. At the same time, the available evidence does not support the hypothesis that differential strategy use by low-spans and high-spans accounts for the relationship between WM span and higher-level cognition. To the contrary, if anything, differential strategy use appears to suppress the true magnitude of this relationship. Additional research like that by Turley-Ames and Whitfield will be critical to understanding why this is so.

## Macroanalytic Research

An advantage of microanalytic research on WM is that it is potentially informative about the precise nature of basic information processing mechanisms underlying individual differences in WM. That is, if WM span correlates with individual differences in some experimental task, then the implication is that a common mechanism is operating in both tasks. To reiterate, based on results of the microanalytic studies just reviewed, we argue that WM reflects the capacity for attention control, which is critical for tasks that demand maintenance of task-relevant information. However, a potential disadvantage of this approach is a consequence of a basic psychometric principle alluded to earlier: no single task can be expected to provide a process-pure measure of the construct it is hypothesized to measure. For this reason, although a factor like attention control may indeed play an important role in the experimental tasks we have investigated in our research, we must assume that a number of other factors contribute to true scores in the tasks. Furthermore, on the basis of evidence from microanalytic studies alone, the possibility that these factors contribute to the correlation of scores in the task with WM span cannot be unequivocally rejected.

With this in mind, a second approach that we have used in research on the nature of individual differences in WM is *macroanalytic* in that the goal is to investigate the relationship between WM and individual differences in broad, psychometrically established constructs. In particular, this research has focused on the link between WM and the aspect of cognition that Cattell (1943) first termed *fluid intelligence* ($g_f$) – the ability to solve novel problems and adapt to new situations. Summarized, evidence from macroanalytic research suggests that WM may be an important component of $g_f$. For example, at the latent-variable level, Kyllonen and Christal (1990) found a strong positive correlation (.90) between WM and $g_f$. Furthermore, Kyllonen (1996) also reported high positive correlations between $g_f$ and latent variables representing WM in three content areas: verbal (.94), spatial (.96), and numerical (.95). Kyllonen summarized his research as follows:

> We have observed in study after study, under a variety of operationalizations, using a diverse set of criteria, that working memory capacity is more highly related to performance on other cognitive tests, and is more highly related to learning, both short-term and long-term, than is any other cognitive factor. This finding of the centrality of the working memory capacity factor leads to the conclusion that working memory capacity may indeed be essentially Spearman's *g*. (p. 73)

Engle et al. (1999b) further investigated the relationship between WM and $g_f$. WM was measured with span tasks similar to those described earlier, while short-term memory (STM) was measured with simple memory span tasks (e.g., word span); $g_f$ was measured with two nonverbal tests of abstract reasoning ability. Engle et al. predicted that latent variables representing WM and STM would correlate, given that some of the same domain-specific skills and procedures were captured by both. For example, skill in encoding information into long-term memory could contribute to performance in both the reading span and word span tasks. However, they also predicted that once this correlation was taken into account, the WM residual variance would reflect individual differences in attention control and would correlate positively with $g_f$, whereas the STM residual would not. The data were consistent with this prediction: the WM residual correlated significantly with $g_f$ (.49) whereas the STM residual did not (.12).

### *Verbal versus Spatial* WM?

Recently, we have focused more directly on the question of whether WM is domain-specific or domain-general. Given that verbal WM tasks predict both $g_f$ and low-level attention control, it is quite likely that WM tasks measure a general cognitive capability. However, other work suggests that verbal and spatial WM tasks may measure different constructs. For example, Shah and Miyake (1996) observed the correlations between scores in verbal and spatial WM tasks (reading span and spatial span) and independent estimates of verbal ability and spatial ability. The major finding of this study was that spatial span correlated with spatial ability (.66), but not with verbal ability (.07), whereas the reading span measure correlated with verbal ability (.45), but not with spatial ability (.12). In addition, the correlation between the two WM tasks was weak (.23). Shah and Miyake (1999) therefore concluded that "the predictive powers of the two complex memory span tasks seem to be domain specific" (p. 11).

Nevertheless, a limitation of the Shah and Miyake (1996) study is that the subjects were college students from two relatively selective universities. Therefore, as Shah and Miyake themselves acknowledged, it is possible that variability in the span scores due to a domain-general WM factor was restricted compared to what might be expected within a more heterogeneous sample. With this in mind, we recently conducted a study in which over 200 subjects, recruited from university subject pools and from the general population, completed both verbal and spatial WM and STM tasks; in addition, subjects completed tests of verbal reasoning and spatial reasoning, as well as "decontextualized" reasoning (e.g., Raven's Progressive Matrices). As described before, each WM task included a processing component and a storage component, while each STM task included only a storage component.

As expected, there were moderate positive correlations among all of the memory tasks. However, the patterns of intercorrelations differed for STM and WM. The mean correlation among domain-matching WM measures was .64, compared to a mean of .56 among domain-mismatching measures. By contrast, the mean correlation among domain-matching STM measures

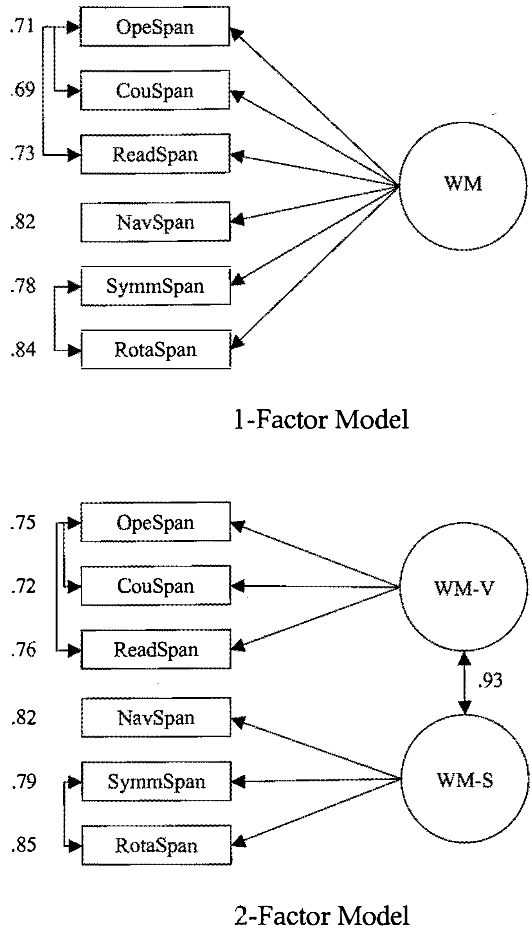


FIGURE 1. Domain-general model (top panel); domain-specific model (bottom panel).

was .68, compared to a mean of .47 among domain-mismatching measures. Thus, the domain-matching versus domain-mismatching difference was greater for the STM measures (.21) than for the WM measures (.08). In line with other research (e.g., Park et al., 2002; Swanson & Howell, 2001), these results suggest that the verbal and spatial STM span tasks measured more distinct constructs than did the verbal and spatial WM span tasks.

To further investigate the possibility that WM is domain-general, we conducted both exploratory and confirmatory factor analyses on the WM measures. In an exploratory factor analysis, the first factor accounted for a large proportion of the variance (65.9%), and it was the only factor that met the criterion for extraction (i.e., eigenvalue greater than one), suggesting that the WM measures tapped a single construct. To perform a more rigorous test, we also conducted confirmatory factor analyses to compare 1-factor and 2-factor models, with the latter model consisting of separate but correlated verbal and spatial factors (WM-V and WM-S). The results, illustrated in Figure 1, were as follows: In the 1-factor model, each of the WM measures had a strong positive loading on the common factor. In addition, while the 2-factor model provided a slightly better fit to the data than did the 1-factor model, the improvement was not statistically significant, and the verbal and spatial factors correlated near one (.93). The data clearly do not lead us to reject the parsimonious view that WM capacity reflects a domain-general construct.

We conducted two additional analyses to examine the involvement of a domain-general WM in $g_f$. In the first analysis, the "predictor-side" model was a 1-factor WM model, whereas the "criterion-side" model was a 3-factor reasoning model with a $g_f$ factor, onto which all of the reasoning measures loaded, plus domain-specific verbal and spatial factors, onto which the verbal and spatial measures loaded (REA-V and REA-S). This "nested" model of the reasoning tasks allowed us to isolate the variance shared among all the reasoning tasks ($g_f$), as well as the residual variance that was uniquely shared among the verbal tasks and among the spatial tasks. As shown in Figure 2, WM predicted about 35% of the variance in $g_f$, a value consistent with estimates from prior studies (Conway et al., 2002; Engle et al., 1999b). In addition, WM had weaker, but still significant, effects on domain-specific aspects of both verbal and spatial reasoning (.27 and .30, respectively). Thus, the variance shared by verbal and spatial WM tasks, reflecting domain-general WM, predicted both general and specific reasoning abilities.

In the second analysis, we added the STM measures to the structural equation model shown in Figure 2. In this model, all of the memory measures loaded onto a factor that we hypothesized to represent the central factor underlying individual differences in WM: executive attention (EA). In addition, the six verbal memory measures simultaneously loaded onto a verbal factor, whereas the six spatial memory measures loaded onto a spatial factor. We interpreted these domain-specific factors (STORAGE-V and STORAGE-S) as reflecting storage or coding processes specific to verbal or spatial stimuli and independent of domain-general executive attention. The logic guiding specification of this model was that no WM or STM task is purely domain-general or domain-specific. Instead, WM measures capture a domain-general factor primarily but also domain-specific factors, whereas STM tasks capture domain-specific factors primarily but also a domain-general factor. Therefore, from each measure, we extracted domain-general variance and domain-specific variance.

Consistent with the model in Figure 2, EA had a strong effect on $g_f$ (.57) and weaker effects on REA-V (.26) and REA-S (.33). These correlations

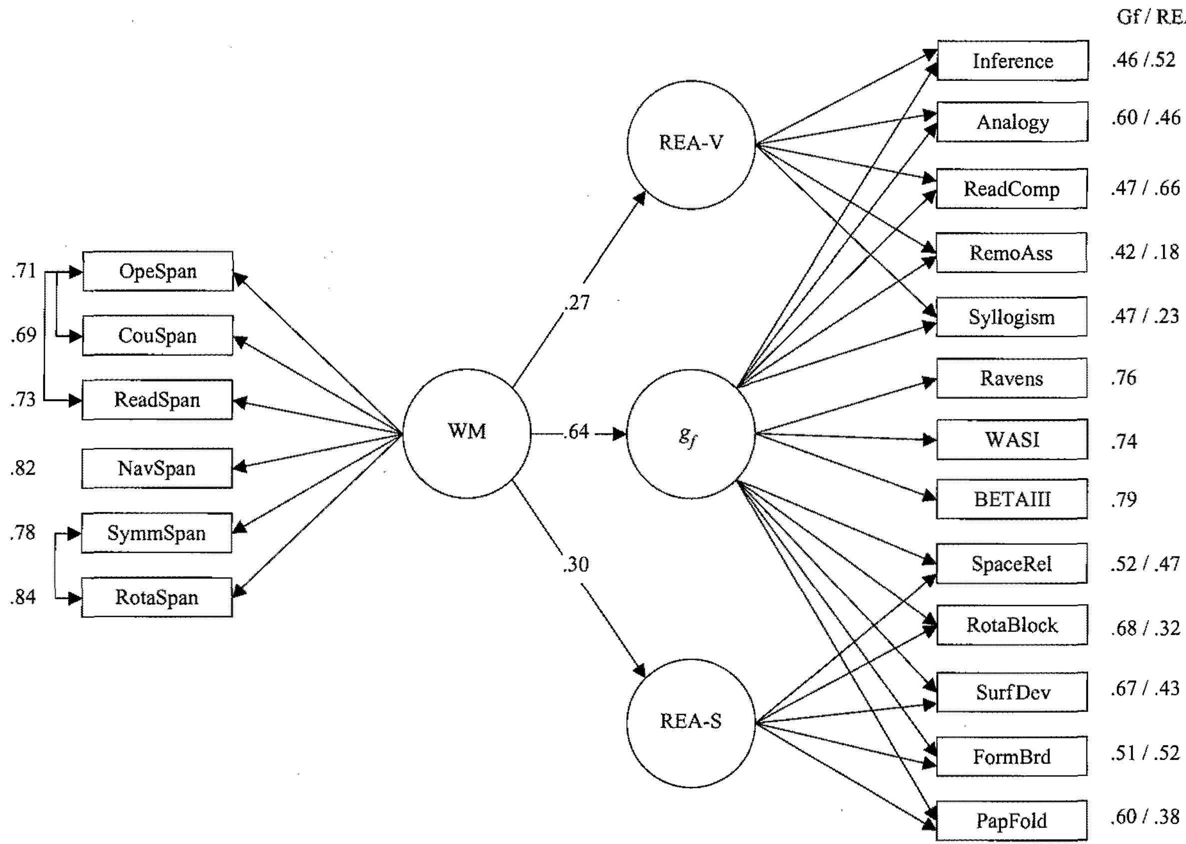


FIGURE 2. Structural equation modeling relating WM and $g_f$.

almost perfectly matched the ones we found when we modeled "executive attention" with only the WM span tasks (cf. Fig. 2). In addition, STORAGE-V had a positive effect on REA-V (.42) but a nonsignificant effect on $g_f$. As prior studies have found (e.g., Engle et al., 1999b), verbal storage and rehearsal processes account for unique variance in verbal ability over and beyond that accounted for by WM. However, the same is not true for $g_f$, where only WM accounts for unique variance. Lastly, STORAGE-S showed a quite different pattern of relations to reasoning, with strong effects on both REA-S (.39) and $g_f$ (.51). Thus, not only did spatial-storage processes account for aspects of spatial reasoning that are independent of $g_f$, but they also accounted for a sizeable proportion of $g_f$ variance that is not shared with EA. Consistent with other research (e.g., Miyake et al., 2001; Oberauer, 1993), this finding suggests that spatial storage may be more closely tied to executive functioning than is verbal storage. A possible interpretation of this finding is that verbal storage is supported by well-learned coding and storage processes (e.g., rehearsal), whereas spatial storage, due to its novelty, must rely more on attention control ability. This is an intriguing hypothesis as it suggests that executive attention can be measured in span tasks without dual-task requirements, but it must await further investigation.

## Toward a Broader Perspective on the Role of WM in Higher-Level Cognition

To sum up, evidence from two types of research is consistent with a domain-general hypothesis of individual differences in WM. First, microanalytic research suggests that an important factor underlying individual differences in WM is the capacity for attention control. That is, WM span correlates with performance in elementary attention tasks like antisaccade. Once again, we believe that this evidence is compelling because there are no apparent domain-specific factors to which span-related differences can be attributed in these tasks. Second, macroanalytic research suggests that WM plays an important role in the broad aspect of cognition referred to as $g_f$. That is, WM span predicts $g_f$ even after the contribution of domain-specific factors has been taken into account.

But how important is WM for real-world tasks in which many other factors might be expected to play a role? For example, does WM contribute to success in tasks like choosing a move in a chess game, or even in more mundane tasks like financial planning? We have begun to explore this sort of question. The general approach in this research is to create a laboratory simulation of some real-world task and then to determine whether, and to what extent, WM contributes to performance above and beyond the influence of other possible predictors. For example, in a recent study by Hambrick and Engle (2002), subjects performed a task that involved listening to, and then answering questions about, simulated radio broadcasts of baseball games. The subjects were 181 adults with wide ranges of WM and knowledge about baseball, and the radio broadcasts were recorded by a baseball announcer for a local radio station.

Not surprisingly, baseball knowledge was a strong predictor of memory for information from the baseball games, including changes in which bases were occupied after each turn at bat and information about the players (e.g., batting averages). In fact, baseball knowledge accounted for over half of the variance. However, there was evidence that WM enhanced the effect of domain knowledge on memory performance. That is, for information that was judged directly relevant to the games (e.g., players' batting averages), the effect of domain knowledge on memory performance was greater for high-spans than for low-spans. Based on this finding, we suggested that WM may serve as a "bottom-up" constraint on knowledge use in cognitive performance. In particular, we suggested that to the extent that integrating new information with preexisting knowledge depends on maintaining that information in an activated state for some period of time, high-spans should benefit more from preexisting knowledge than low-spans.

Additional evidence concerning the interplay between domain knowledge and WM was reported by Wittmann and Süß (1999), who investigated the effects of domain knowledge and WM on performance in work simulations. For example, in one task, the goal was to control the energy output of a coal-fired power plant by manipulating a number of variables (e.g., coal input); another task involved managing the production of a garment manufacturing company. A consistent finding from this research was that task-specific knowledge (i.e., knowledge acquired during the simulations) was a strong predictor of final performance. However, Wittmann and Süß also reported that WM was a significant predictor of performance above and beyond knowledge. Thus, there is reason to believe that effects of WM on higher-level cognition are not limited to simple laboratory tasks. Rather, WM may be an important contributor to success in complex task environments in which many other factors might also be expected to play a role.

## SUMMARY AND CONCLUSIONS

Working memory has now been a topic of intensive research in cognitive psychology for more than 25 years. What has this research revealed about the nature of this construct and its involvement in higher-level cognition? At least two conclusions seem warranted. First, the work from two complementary perspectives – microanalytic and macroanalytic – converges on the conclusion that individual differences in WM span reflect something more general than factors tied to particular domains. For example, WM span correlates with individual differences in elementary attention tasks and in tests of general intelligence. Second, it now seems clear that these domain-general factors may be responsible for the correlation of WM span with higher-level cognition. Important goals for future research are to refine understanding of the nature of these factors and to study their involvement in complex, real-world activities.

## References


Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Advances in the psychology of learning and motivation research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

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

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Broadbent, D. (1958). *Perception and communication*. Oxford: Permagon.

Binet, A. (1905). New methods for the diagnosis of the intellectual level of subnormals. *L'Année Psychologique, 12*, 191–244.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology, 2*, 247–265.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40*, 153–193.

Cherry, C. E. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*, 975–979.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin and Review, 8*, 331–335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D., & Minkoff, S. (2002). A latent variable analysis of working memory capacity, short term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*, 163–183.

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

Denney, N. W., & Heidrich, S. M. (1990). Training effects on Raven's Progressive Matrices in young, middle-aged, and elderly adults. *Psychology and Aging, 5*, 144–145.

Duncan, J. (1990). Goal weighting and the choice of behavior in a complex world. *Ergonomics, 33*, 1265–1279.

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., Kane, M. J., & Tuholski, S. W. (1999a). 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. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999b). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309–331.
Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.
Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology, 44*, 339–384.
Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–225). New York: Academic Press.
James, W. (1890). *Principles of psychology*. New York: Holt.
Kane, M. J., & Engle, R. W. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132*, 47–70.
Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 333–358.
Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*, 169–183.
Klauer, K. J., Willmes, K., & Phye, G. D. (2002). Inducing inductive reasoning: Does it transfer to fluid intelligence? *Contemporary Educational Psychology, 27*, 1–25.
Klein, K., & Fiss, W. H. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments and Computers, 31*, 429–432.
Kyllonen, P. C. (1996). Is working memory capacity Spearman's *g*? In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 49–75). Mahwah, NJ: Erlbaum.
Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*, 389–433.
Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences, 2*, 129–160.
MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.
McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory and Cognition, 29*, 10–17.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.
Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General, 130*, 621–640.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56–60.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology, 3*, 1–18.

Oberauer, K. (1993). Die Koordination kognitiver Operationen. Eine Studie zum Zusammenhang von 'working memory' und Intelligenz [The coordination of cognitive operations. A study of the relationship of working memory and intelligence]. *Zeitschrift für Psychologie, 201*, 57–84.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult lifespan. *Psychology and Aging, 17*, 299–320.

Schrock, C., & Engle, R. W. (in preparation). Working memory capacity and behavioral inhibition.

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.

Shah, P., & Miyake, A. (1999). Models of working memory: An introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 1–27). Cambridge, UK: Cambridge University Press.

Spearman, C. (1927). *The abilities of man*. London: MacMillan.

Swanson, H. L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology, 93*, 720–734.

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. *Journal of Memory and Language, 49*, 446–468.

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

Welsh, M. C., Satterlee-Cartmell, T., & Stine, M. (1999). Towers of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain & Cognition, 41*, 231–242.

Wittmann, W. W., & Süß, H. (1999). Investigating the paths between working memory, intelligence, and complex problem-solving performances via Brunswik Symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences* (pp. 77–110). Washington, DC: American Psychological Association.