Current Anthropology Volume 51, Supplement 1, June 2010

Role of Working-Memory Capacity in Cognitive Control

by Randall W. Engle

This paper discusses the psychometric properties important to the measurement of working memory. Domain-specific aspects of working memory, such as the phonological loop and store and the visual and spatial stores, are important to the performance of many real-world tasks and were probably important to the evolution of the modern mind. However, there is little work demonstrating the critical psychometric properties of reliability and validity of measurement of these domain-specific stores. The domain-general aspect of working memory—attention control—on the other hand, has established reliability and validity of measurement. Individual differences in domain-general working-memory capacity have been shown to be important to a wide range of both speech-based and visual/spatial-based tasks. Working-memory capacity appears to be both a trait and a state variable.

What is working memory? What does it mean when we use the term "working-memory capacity" (WMC)? How do we measure it? Do people differ in WMC in any meaningful way? Is that a trait or a state variable? Is it important in feral cognition? How might it have been important in the evolution of the human mind? What brain and genetic mechanisms might have led to that development? These are all questions I will try to deal with in the following pages with varying levels of confidence in the answers and, in some cases, mere speculation about the answers.

Working memory is a system of domain-specific stores or formats for temporarily representing information along with a domain-general supervisory or executive attention mechanism. The Baddeley and Hitch (1974) model proposed two formats for temporary storage: one based on speech or articulation and the other for representing and maintaining visual/spatial information. More recently, Baddeley (2000) added a multidimensional store allowing binding of information across dimensions. More recent work has shown the need to fractionate these structures into more specialized structures such as one for visual information and another for spatial information (Logie 2003).

However, I have argued that there are as many domain-specific stores as there are different ways of thinking—probably on the order of a few dozen such formats (Engle and Kane 2004). Like Cowan (1995), I conceive of the contents of these "stores" as temporarily activated representations in long-term memory, as links or pointers, as it were, to existing representations in long-term or secondary memory. These links will decline in strength over time below a threshold of consciousness, but attention to a link will lead to reactivation that can maintain activation above that threshold. Rehearsal is one way in which that is accomplished.

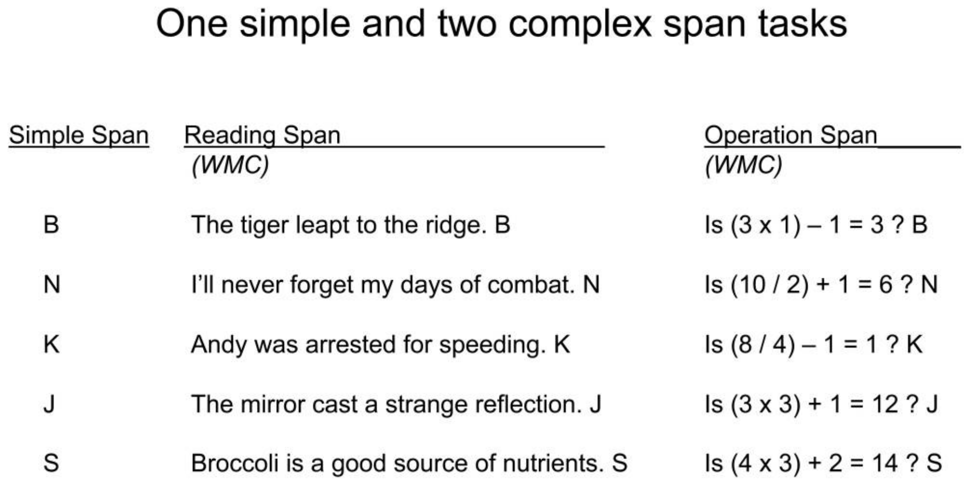
I am confident that each of those domain-specific stores is important in some way to modern cognition and has probably played an important role in the development of the modern mind. The formats proposed in the Baddeley model have been studied extensively, and the phenomena that serve as their basis are large and reliable across experiments, even if they have not demonstrated reliability within the individual. I am also confident that there are individual differences in the practice-developed skill for the various coding formats, and it is also likely that there are differences in the biological mechanisms necessary to perform those functions, both of which would seem to be necessary for a role in evolution. Cognitive psychologists suspect, for instance, that the phonological loop is important, possibly even necessary, the development of language and reading in the individual (Gathercole and Baddeley 1993). However, there has been either far too little work demonstrating the importance of the domain-specific stores in a wide range of real-world cognitive functions or far too little evidence of reliable individual differences in those functions. In fact, while Logie et al. (1996) found strong group effects of two markers of the phonological loop, they found little evidence of reliability of measurement for the effects at the individual level. Thus, we can only speculate about their role in the evolution of the modern mind.

Both visual and spatial working memory were probably important to the development of tool use. Kunde, Müsseler, and Heuer (2007), for example, showed that using a lever

---

Randall W. Engle is Professor at the School of Psychology, Georgia Institute of Technology (654 Cherry Street, Atlanta, Georgia 30332, U.S.A. [randall.engle@gatech.edu]).

0011-3204/2010/510S1-0004$10.00. DOI: 10.1086/650572



One simple and two complex span tasks

Simple Span	Reading Span *(WMC)*	Operation Span *(WMC)*
B	The tiger leapt to the ridge. B	Is (3 x 1) – 1 = 3 ? B
N	I'll never forget my days of combat. N	Is (10 / 2) + 1 = 6 ? N
K	Andy was arrested for speeding. K	Is (8 / 4) – 1 = 1 ? K
J	The mirror cast a strange reflection. J	Is (3 x 3) + 1 = 12 ? J
S	Broccoli is a good source of nutrients. S	Is (4 x 3) + 2 = 14 ? S

Figure 1. Example of a simple span task, a reading span task, and an operation span task. *WMC* = working-memory capacity.

requires that the individual be able to manage the image of the hand moving in one direction *and* the other point of the lever moving in an opposite direction at the same time. This development is quite likely independent of language development. It is also likely that both speech-based and visually and spatially based coding formats will require the use of limited-capacity attention control under complex situations involving interference and distraction.

The vast majority of the work on individual differences in working memory has been done in the context of WMC using complex span tasks. I will discuss at some length the measurement of WMC, the observation that many different measures reflect a common construct, psychometric characteristics of reliability and validity of those measures, and the nature of the construct underlying those measures.

Measurement of WMC

The issue of measurement of WMC became important because of psychometric problems in the measurement of traditional ideas of short-term memory (Crowder 1982). Psychological measures need to satisfy standards of reliability and validity to be of value to either theory or application. Reliability, in psychometric terms, means that there is a consistency of measurement across time and administrations of a task within the same individual. Validity means that the measure consistently predicts performance on some cognitive measure of interest such as comprehension or reasoning. Simple span tasks such as digit span, the historical and prototypical measure of short-term memory, are insufficiently reliable (Dempster and Corkill 1999) and so inconsistently valid that Crowder (1982) signaled "the demise of short-term memory" (291) based partly on the failure of span measures of short-term memory to consistently correlate with important real-world cognition.

Complex span measures modeled after Daneman and Carpenter's (1980) reading span task have shown themselves to be at least moderately reliable and consistently valid (Engle and Kane 2004) in predicting a huge array of higher-level and real-world cognitive tasks. Figure 1 shows one simple and two complex span tasks using letters as to-be-remembered items.

In the simple span tasks, the subject is presented with one letter at a time and is then asked to recall the letters in correct order. The number of letters will generally vary from two to eight or so, but at least in the way my lab does these tasks, the order of the list length varies randomly so the subject cannot know ahead of time how many letters will be presented until the list presentation has ended. In our version of the reading span task, the subject is given a brief period to read the sentence and to judge whether the sentence makes sense; the letter is then presented for 800 milliseconds before the next sentence is presented. At the end of three to seven such items, the subject is cued to recall the letters by clicking the mouse on letters in a matrix on the screen. In the operation span task, the subject is shown an operation such as "$(3 \times 1) - 1 =$" for a period of time; the next screen shows a digit such as 6, and the subject is to click on the "yes" or "no" box on the screen to indicate whether the digit is the correct answer to the arithmetic operation. The subject is then shown a letter for 800 milliseconds. After three to seven such items, the subject performs recall by clicking with the mouse on a matrix of letters on the screen, as with the reading span task (Unsworth et al. 2005).

Notice that these two tasks have some similarities and some differences. They both involve an easy yet nontrivial verbal task (reading a sentence for meaning in one case and solving a simple arithmetic expression in the other), and that task is iteratively interleaved with a verbal item (in this case a letter) for later recall. However, reading and arithmetic are different skills, and it would not be surprising if the scores on these two tasks differentially correlated with a higher-level task such as reading comprehension. But that is not what we find. The two tasks account for pretty much the same variance in a wide array of higher-order cognitive tasks, even tasks with a high level of spatial skill such as the Raven's Matrices. The Raven's Progressive Matrices task is arguably the gold standard of general fluid intelligence. It requires the test taker to pick an object or pattern that would best fit in a matrix of other objects or figures. In fact, even complex tasks that have no discernible verbal component seem to measure the same variance in higher-order cognition. Figure 2 shows three different spatial tasks we have used, and they account for quite similar variance to the verbal tasks described above. Each of these three tasks requires the subject to perform some spatial task, such as mentally rotating a letter, deciding whether a figure is symmetrical around a vertical axis, or mentally traveling around a large block letter and deciding whether each subsequent corner was internal or external. That operation was followed by something that the subject was instructed to recall, such as an arrow of a variable length and direction, a highlighted cell in a matrix, or the distance and direction traveled by a ball. Recall was accomplished by drawing the to-be-recalled objects on paper.

In a structural equation modeling study of the tasks, scores from the complex spatial and complex verbal tasks all loaded on a common construct. As can be seen in figure 3 (*left*), the three verbal and three spatial tasks all had similar loadings on the WMC construct. Further, this construct was strongly associated with a construct for fluid intelligence. Even if the verbal and spatial tasks were forced into a two-construct model, as in the model in figure 3 (*right*), the two were highly correlated at 0.93 (Kane et al. 2004). This and other studies provide support for the idea that the great variety of complex span tasks reflects a unitary construct common to verbal and spatial processing and that the construct is important to a wide range of higher-level tasks, including reading and listening comprehension, complex learning, and reasoning (Engle and Kane 2004; Turner and Engle 1989).

As I mentioned earlier, the two criteria important to any psychological measure are reliability and validity. One way reliability is assessed is to do a split-half correlation of the test. This assesses internal consistency within a session. All of the complex span tasks have split-half reliabilities of 0.7–0.9,

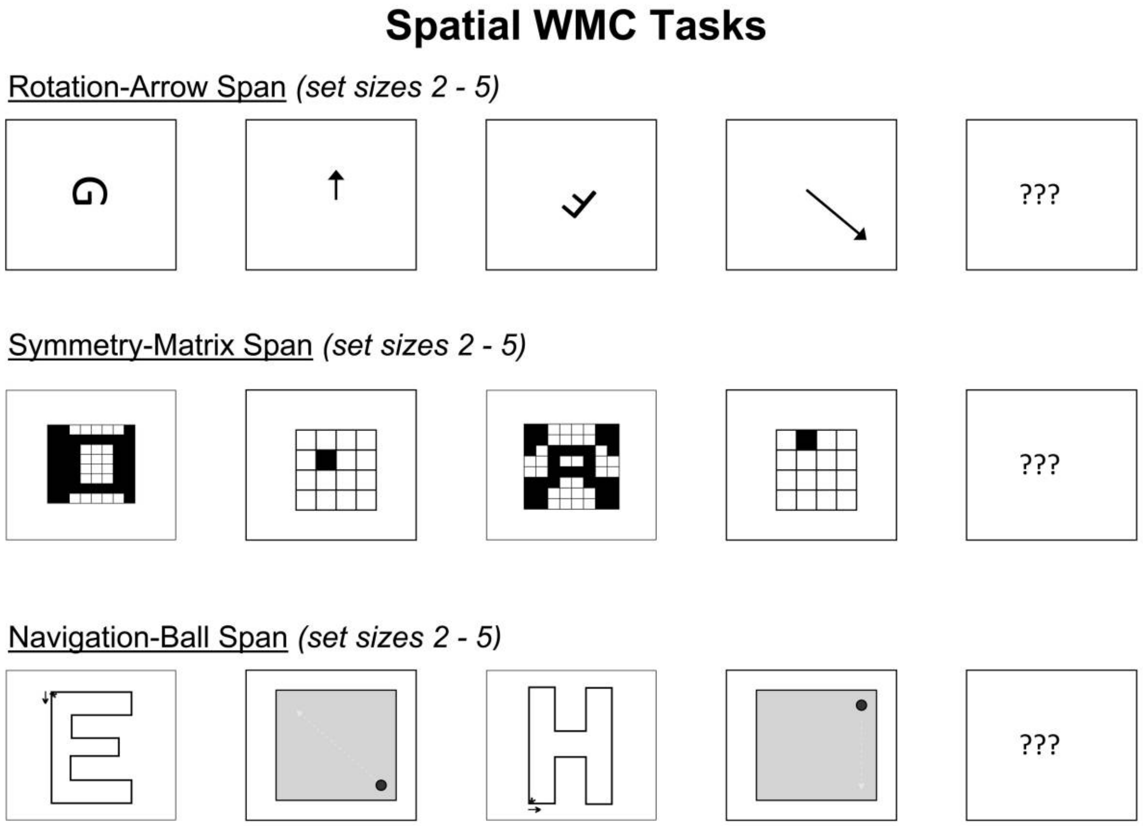


Figure 2. Three different spatial tasks used by Kane et al. (2004). *WMC* = working-memory capacity.

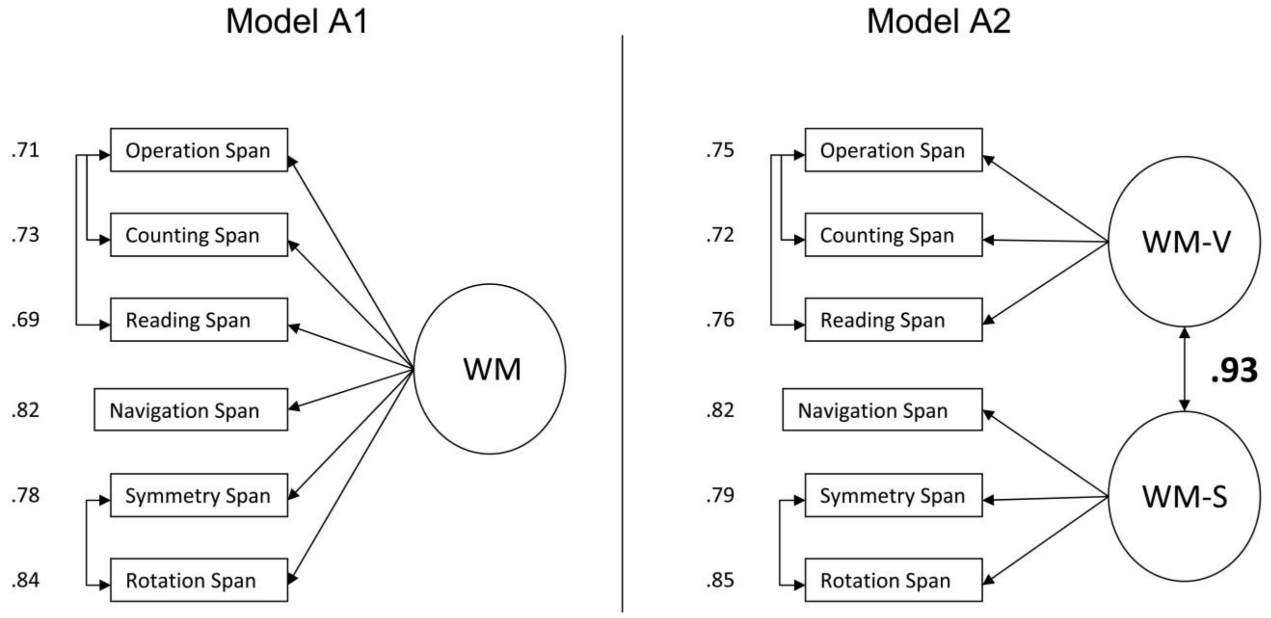


Figure 3. Structural equation model of verbal complex span tasks (operation span, reading span, and counting span) and spatial complex span tasks (navigation span, symmetry span, and rotation span). Adapted from Kane et al. (2004). *WM* = working memory; *WM-V* = verbal working memory; *WM-S* = spatial working memory.

which is quite reasonable. Another way to assess reliability is to administer the same or equivalent forms of the test to the same individuals at two or more different times. Klein and Fiss (1999) tested a group of individuals three times on the operation span task over 9–10 weeks. Scores on the operation span task increased over the administrations because of practice, but the overall corrected reliability was 0.88, suggesting that performance increased evenly across ability levels. Another study of the reliability of the operation span task was conducted by Unsworth et al. (2005). They tested 78 subjects on the automated operation span task and retested them with a mean delay of 13 days. The automated operation span showed a reliability coefficient of 0.83, which is quite high for a lab-based task. Similar studies have not been done with other complex span tasks, but the studies with operation span suggest a high level of reliability over lengthy periods of time and with repeated administrations.

There are multiple ways one can assess validity. Construct validity means that measures thought to reflect the underlying unobservable hypothetical construct should correlate with performance in other tasks for which that construct is deemed important. By this thinking, any of the WMC tasks we have described here, as well as dozens of others that have been used in published reports, should predict performance on any other task in which the hypothesized construct is a crucial mediator of performance. Performance on WMC span tasks has been shown to correlate with a wide range of higher-order cognitive tasks, such as reading and listening comprehension (Daneman and Carpenter 1983; Daneman and Merikle 1996), language comprehension (King and Just 1991), following oral and spatial directions (Engle, Carullo, and Collins 1991), vocabulary learning from context (Daneman and Green 1986), note taking in class (Kiewra and Benton 1988), writing (Benton et al. 1984), reasoning (Barrouillet 1996; Kyllonen and Christal 1990), hypothesis generation (Dougherty and Hunter 2003), bridge playing (Clarkson-Smith and Hartley 1990), and complex-task learning, such as learning to write programs in a computer language (Kyllonen and Stephens 1990). Low-WMC individuals are less good at blocking intrusive thought than are high-WMC individuals (Brewin and Beaton 2002), which would seem to interact with many different psychopathologies to impair cognition proportionally to the extent that troubling thoughts need to be suppressed.

Many of our studies have used extreme-groups designs to test whether a dependent variable is sensitive to WMC differences. In these studies, we test a broad range of individuals on measures of WMC and use the upper quartile as the high-WMC group and the lower quartile as the low-WMC group. It should be pointed out that the subjects in our broad sample are recruited from college campuses and from large urban centers. Some labs use college students only from large comprehensive universities. Thus, the low-WMC individuals are in no way pathological or nonnormal.

A measure shows discriminant validity to the extent that it does not correlate with measures that the construct should *not* be important to. I will argue below that the construct at the core of individual differences in measures of WMC is the ability to control attention to keep representations most relevant to the task at hand in active memory or most easily retrievable from inactive memory and that this comes into play most directly under conditions of interference from competing representations. To the extent that this idea has credibility, individuals who are different in WMC should not show differences when cognitive control is minimally required because the task is minimally affected by interference. Kane and Engle (2002) showed that high- and low-WMC individuals showed large differences in the recall of 10-item lists when the lists had been preceded by at least four other lists for recall; however, on the very first list, with no proactive interference to affect recall, high- and low-WMC individuals did not differ in level of recall. Similarly, Unsworth, Schrock, and Engle (2004) showed that in the high-interference antisaccade task (described in more detail below), low-WMC individuals made many more errors than did high-WMC individuals. However, in the low-interference prosaccade condition, the two groups did not differ in errors.

Another interesting demonstration of the validity of WMC is that individuals measured to be low in WMC appear to have their mind wander when engaged in a task more so than do high-WMC individuals. Michael Kane and his colleagues (Kane et al. 2007*a*) tested a large number of individuals on a battery of complex span tasks that included the operation span, reading span, and symmetry span. They then provided the subjects with a Palm PDA that beeped a signal randomly eight times each day for 7 days. When the signal occurred, the subjects were to immediately consider whether their mind had wandered from the task they were supposed to be performing at that time. They then entered that information into the PDA and answered a series of other questions. Kane et al. found that low-WMC individuals were more likely than high spans (all students at a large state university) to have their mind wander as the task became more challenging or required more effort or if the subjects were trying hard to concentrate on the task. This finding has great implications for the role of WMC in planning events over the course of the day and in implementing those plans. It would be particularly important to more complex and more difficult plans requiring greater concentration.

There are other measures of WMC than the ones described here, but none of them have the extensive assessment of reliability and validity of the complex span tasks. The N-back task, for example, presents subjects with a series of verbal or pictorial objects and asks them to press a key when an item is identical to the one presented three back in a three-back task or two back in a two-back task. The N-back task has been used almost exclusively in studies using the functional magnetic resonance imaging (fMRI) technology to study brain circuits associated with working memory because it conveniently requires only a two-choice button press to perform the memory task. A recent paper by Kane, Conway, Miura, and Colflesh (Kane et al. 2007*b*) showed that performance on a three-back task did correlate with the Raven's score but that scores from a two-back task did not. Kane et al. (2007*b*) found that the N-back, even three-back, seemed to have little or no overlap with the operation span because they accounted for nearly unique variance in Raven performance. Thus, N-back and the complex span tasks appear to reflect quite different constructs, but there has been little work to determine their similarities and differences. This would appear to be an important field for future work. The vast majority of work on the role of WMC in real-world cognition has been done using complex span tasks, and the vast majority of work on the cognitive neuroscience of working memory has been done using N-back tasks. The fact that these two tasks appear to have little or no overlap in variance seems to present a large conceptual problem to the field.

Working Memory as a State/Trait Variable

I have presented the work on individual differences in WMC as if it were a trait variable, an abiding characteristic of the individual that remains relatively immutable over time. Of course, it is the trait aspect that is most germane to discussions of the evolution of working memory. But, while there is good evidence to support that approach, it needs to be pointed out here that we need to think of WMC as a state variable as well. In much the same way that psychologists talk about anxiety as reflecting both a trait of the individual and a state that depends on the context, we should think about WMC as both a trait and a state variable. Let me provide some examples.

"Stereotype threat" refers to the fact that individuals perform poorly on a test if a relevant stereotype is associated with performance on that task. For example, women score worse on a mathematics test if they are told before hand that women typically score worse than men on math tests. Schmader and Johns (2003) found that the effect of stereotype threat was mediated by a reduction in WMC. Similarly, Richeson and Shelton (2003) found that it is especially taxing of attention control for a racially biased white individual to interact with a black individual and that the white individual will perform worse on a subsequent task also demanding attention control.

Sleep deprivation and fatigue are also associated with reductions in WMC. A recent study with 10 highly experienced U.S. Air Force pilots measured performance on a sophisticated flight simulator along with a battery of other measures including the operation span every 2 hours while the pilots were kept awake for 35 hours (N. Lopez, F. H. Previc, J. Fischer, C. M. DaLuz, A. J. Workman, W. R. Ercoline, R. H. Evans, N. A. Dillon, R. W. Engle, and R. P. Heitz, unpublished data). Even though the pilots were quite experienced and often flew with little sleep, they made errors in the last half of the 35-hour period, and the errors were highly predicted by performance on the operation span at 0.65. Thus, as fatigue increased, WMC declined, and that was associated with increased errors on the simulator. WMC has also been shown to be affected by stress (Beilock and Carr 2005) and alcohol (Finn 2002). These results do not diminish the importance of WMC as a trait variable; they do, however, point out the difficulty in specifying whether the results for a given individual at a given point in time are driven by state or trait.

WMC as Attention Control

I have described the considerable work on the psychometric properties of measures of WMC and the huge variety of tasks that seems to depend on that capacity. I will now turn to work on the nature of the construct that is reflected by those tasks. I have argued that the complex span tasks reflect individual differences in the ability to control attention to task-critical representations to keep that information either available in active memory or easily and quickly retrievable from inactive memory. The construct reflects the interface between attention and memory. It is important to strengthening the activation of representations critical to the current task but also important to the dampening or inhibition of representations that would interfere with the task. I will discuss several lines of work supporting the idea that individual differences in WMC reflect differences in attention capability even though the complex span tasks measure the number of items recalled.

One paradigm that has been used to study the regulation of attention is the dichotic listening procedure in which two different auditory messages are played simultaneously to the two ears and the subjects are instructed to ignore the message coming to one of the ears and to simultaneously repeat the words in the message coming to the attended ear. Subjects are unable to report anything about the words in the ignored message other than some primitive features of the voice of the speaker. However, Moray (1959) found that if he included the subjects' first name among the random words in the ignored ear, about one-third of the subjects reported hearing their name even though they had no idea what the other words were in the ignored message.

If individual differences in WMC reflect differential ability to attend to events important to the current task and to block distracting events, then we should see an interesting difference between high- and low-WMC individuals in the dichotic listening task. If high-WMC individuals are better at blocking distracting information, they should be less likely to hear their own name in the ignored message than should low-span subjects. Conway, Cowan, and Bunting (2001) tested high- and low-WMC individuals on Moray's task, with the subjects' first name inserted by computer into the ignored ear. While only 20% of high-WMC individuals reported hearing their name, 65% of low-WMC individuals heard their name. Presumably, high spans were better at focusing attention on the attended message, which required blocking the distracting message.

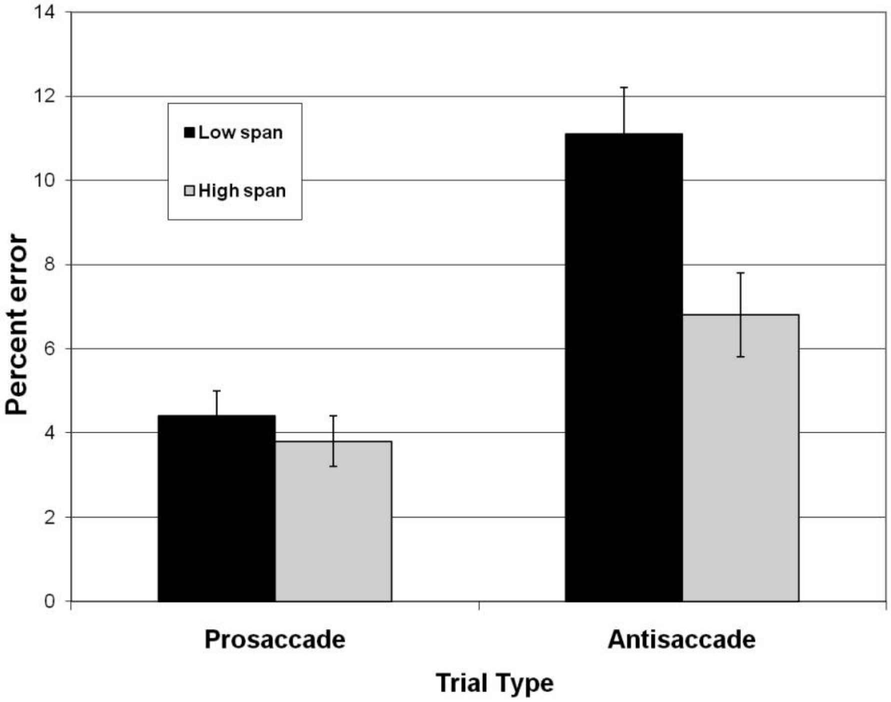


Figure 4. Percent errors for low working-memory capacity (WMC; *gray bars*) and high-WMC (*black bars*) subjects in the prosaccade and antisaccade task. Adapted from Unsworth, Schrock, and Engle (2004).

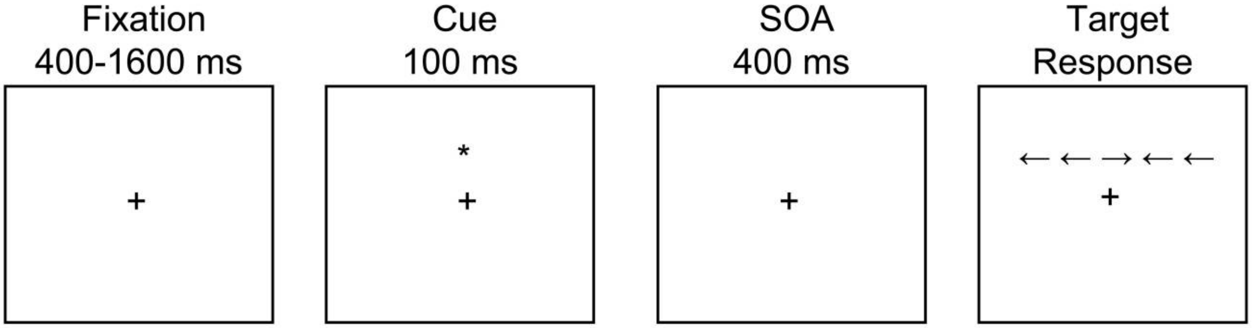


Figure 5. Example of a trial in the attention network task with incompatible flankers surrounding the critical arrow. Adapted from Redick and Engle (2006). *SOA* = stimulus onset asynchrony.

Another relatively primitive attention task that has been used to study attention differences in WMC is the eye movement task, or the antisaccade task, referred to earlier. In our version of this task (Unsworth, Schrock, and Engle 2004), the subject is seated in front of a computer monitor and must closely monitor a fixation point on the screen. There are two boxes printed on the screen, each approximately 11° from the fixation point. At some point, the fixation figure changes, and soon thereafter one of the boxes flickers. An eye tracker is used to monitor fixation and saccades. There are two conditions in the experiment. In the prosaccade condition, as soon as the box flickers, the subject is to move his or her gaze to the flickering box. In the antisaccade condition, the subject is to immediately move his or her gaze to the nonflickering box. This is an elegantly simple task but is devilishly difficult. Millions of years of evolution have prepared us to immediately move our eyes toward a flickering stimulus. Flicker affords movement, and things that move possibly can eat you—or you can eat them. Either way, detection of movement is vitally important. The experimenter is asking the subject to go against this behavior that is strongly predisposed in virtually all animals and to make the first eye movement in the opposite direction. Even a slight glimpse toward the flickering box is considered an error. However, as can be seen in figure 4, even though high- and low-WMC subjects did not differ on the prosaccade task, low-WMC individuals made many more erroneous saccades on the antisaccade version of the task, meaning that they were much more likely to have their attention captured by the flickering box than were the high-WMC individuals. Notice how dissimilar this task is from the complex span tasks we use to measure WMC. Subjects were chosen on the basis of a task in which they recalled sets of letters or words interleaved with calculating a simple arithmetic string or reading a simple sentence. However, the WMC tasks predicted the number of errors in the antisaccade task involving no verbal or complex spatial processing.

Another approach to studying differences in attention is the attention network task (ANT) developed by Michael Posner and his colleagues (Fan et al. 2002). This task requires subjects to look at an arrow on a computer screen and press one of two buttons to indicate the direction the arrow is pointing. The task manipulates the types of cues the subject receives before the critical arrow occurs and the nature of the arrows surrounding the critical arrow, and it tests for three different aspects of attention—alerting, orienting, and executive attention—reflecting the ability to resist attention capture by the environment and attend to a task-critical event. An example from an incompatible trial for the executive attention network is shown in figure 5. The subject on this trial is to press a key indicating the direction the center arrow in the fourth panel is pointing. To the extent that attention is captured by the arrows flanking the critical arrow, performance will be slowed, and more errors will occur. Individuals who can effectively block the attention response to the flankers will do better on this task, and that is the prediction of our theory about WMC and attention control. Redick and Engle (2006) tested high- and low-WMC individuals on the ANT. We predicted high/low differences on the executive attention task but were unclear whether to expect differences on the other two aspects of attention. We found that the two groups did *not* differ on alerting or orienting. However, as we expected, low-WMC individuals performed much more poorly on the test of executive attention than did high-WMC individuals. Also, notice that the results of this experiment would not be predicted by any view of capacity limits based on a finite number of objects, such as 7 ± 2 (Miller 1956) or 4 ± 1 (Cowan 2001); in fact, it is not clear why one would even do the study based on that view. However, if one believes that the construct underlying WMC is the ability to regulate and control attention, the study makes perfect sense.

Psychological Mechanisms Responsible for WMC Differences

I have argued that individual differences are, at base, a result of differences in ability to effectively select representations that are relevant to the task at hand and to deselect, inhibit, or suppress competing representations. Clearly, low-WMC subjects are very different from high-WMC subjects on very low-level attention tasks such as antisaccade and the ANT. However, can we connect the dots between the complex span tasks and attention control? How do differences in attention control lead to differences in the complex span tasks? First, we need to establish whether people who do well on complex span tasks do well because they maintain more information in active memory or because they are better at constantly moving information from inactive memory back into active memory.

Nash Unsworth and I (Unsworth and Engle 2007*a*, 2007*b*) have proposed that complex span tasks work because they require a constant updating of active memory. Think about the operation span task in which the subject is presented with an operation string and then a letter to store for later recall, then another operation string, and then another letter to be stored for later recall. When attention is required to solve the second operation, the subject must allow the representation of the first letter (or all the earlier letters in the emerging list) to decay below the threshold of consciousness. Then, after the operation is solved, the emerging list must be retrieved. If this is the very first list in the task, then the retrieval is pretty simple because there is no interference from previously presented letters. However, if the list is not the first list, then retrieval requires searching among a set of recently active but now inactive (i.e., below-threshold) representations. We have argued that low-WMC subjects are more vulnerable to the effects of interference; thus, during the retrieval phase, they would have to search among a larger set of items—the items from the emerging list *and* the items from the previous list. High-WMC individuals are better able to inhibit or block the activation of the items from the previous lists; thus, they must search among a smaller set of items and are more likely to find and retrieve the new items from the emerging list. One advantage of our executive attention theory is that it is flexible enough to allow high-WMC individuals to use their capability in a number of different ways, depending on what is called for by successful performance on the task at hand.

Brain and Genetic Mechanisms Responsible for WMC Differences

For working memory and WMC differences to play any role in evolution, they must, of course, have a heritable base. We have argued that individual differences in WMC correspond to the dopaminergic system, specifically, differences in circuitry associated with the prefrontal cortex and the anterior cingulate (Kane and Engle 2002). Studies using fMRI point to these structures as important to individual differences, particularly under the effects of proactive interference (Gray, Chabris, and Braver 2003). Further, the relationship between dopamine and WMC appears to be a nonlinear one. Kimberg, D'Esposito, and Farah (1997) found that administration of a dopamine agonist, such as bromocriptine, led to enhanced working-memory performance for low-WMC individuals but hurt performance for high-WMC individuals.

Several alleles have been proposed to account for individual differences of the type discussed here. Unfortunately, the empirical research has tended to be unfocused in the nature of the working-memory tasks or has focused on tasks, such as the N-back task or the Wisconsin Card Sort Task, where the reliability and validity concerns about the tasks make conclusions difficult. For example, the COMT allele is important in the reuptake of dopamine and would be a reasonable choice as one of a family of alleles important to individual differences in WMC. While some studies have found a relationship between the COMT allele and WMC on some working-memory tasks (e.g., letter-number sequencing), the relationship has not been found with other tasks (e.g., N-back; Bruder et al. 2005).

Another possibility is that WMC, which I have referred to as if it were a unitary and monolithic construct, is actually composed of several correlated but different abilities that are mediated by different genotypes. It is possible that an allele such as COMT or the various dopamine transporters are associated with the skills we think of as updating information in working memory but that one or more different alleles are associated with more general attention control. For example, Reuter et al. (2007) found that COMT was not associated with performance on the ANT, even the executive attention network, but that the tryptophan hydroxylase 2 gene, a rate-limiting enzyme for serotonin, was associated with performance on the executive attention task. Many of these relatively recent gene studies suffer from problems of sample size that are compounded by the fact that they often use cognitive tasks such as N-back with unknown psychometric properties of reliability and validity.

Conclusion

Working memory and WMC are fundamental concepts in modern cognitive psychology and in understanding why people differ on the performance of a wide array of real-world tasks. Those concepts are likely also to be important in understanding how our ancestors were able to perform such tasks as verbal and nonverbal communications, shelter building, social interactions in groups, instructing others about tasks and ideas important to the group, tool making, and artistic representation of ourselves and our world. I am confident that as we know more and more about why we differ today in WMC, both as trait and as state, it will help us better understand how that might have played a role in our own evolution.

Acknowledgment

This paper was written while I was supported by a grant from the Office of Naval Research.

References Cited

Baddeley, A. D. 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4:417–423.

Baddeley, A. D., and G. Hitch. 1974. Working memory. In *The psychology of learning and motivation*, vol. 8. G. H. Bower, ed. Pp. 47–89. New York: Academic Press.

Barrouillet, P. 1996. Transitive inferences from set-inclusion relations and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22:1408–1422.

Beilock, S. L., and T. H. Carr. 2005. When high-powered people fail: working memory and "choking under pressure" in math. *Psychological Science* 16:101–105.

Benton, S. L., R. G. Kraft, J. A. Glover, and B. S. Plake. 1984. Cognitive capacity differences among writers. *Journal of Educational Psychology* 76:820–834.

Brewin, C. R., and A. Beaton. 2002. Thought suppression, intelligence, and working memory capacity. *Behaviour Research and Therapy* 40:923–930.

Bruder, G. E., J. G. Keilp, H. Xu, M. Shikhman, E. Schori, J. M. Gorman, and T. C. Gilliam. 2005. Catechol-O-methyltransferase (COMT) genotypes and working memory: associations with differing operations. *Biological Psychiatry* 58:901–907.

Clarkson-Smith, L., and A. A. Hartley. 1990. The game of bridge as an exercise in working memory and reasoning. *Journal of Gerontology* 45:233–238.

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

Cowan, N. 1995. *Attention and memory: an integrated framework.* New York: Oxford University Press.

———. 2001. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:97–185.

Crowder, R. G. 1982. The demise of short-term memory. *Acta Psychologica* 50:291–323.

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

———. 1983. Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9:561–584.

Daneman, M., and I. Green. 1986. Individual differences in comprehending and producing words in context. *Journal of Memory and Language* 25:1–18.

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

Dempster, F., and A. Corkill. 1999. Neo-interference research and the development of intelligence. In *The development of intelligence.* M. Anderson, ed. Pp. 215–243. Hove: Psychology.

Dougherty, M. R. P., and J. Hunter. 2003. Probability judgment and subadditivity: the role of working memory capacity and constraining retrieval. *Memory and Cognition* 31:968–982.

Engle, R. W., J. J. Carullo, and K. W. Collins. 1991. Individual differences in working memory for comprehension and following directions. *Journal of Educational Research* 84:253–262.

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

Fan, J., B. D. McCandliss, T. Sommer, A. Raz, and M. I. Posner. 2002. Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience* 14:340–347.

Finn, P. R. 2002. Motivation, working memory, and decision making: a cognitive-motivational theory of personality vulnerability to alcoholism. *Behavioral and Cognitive Neuroscience Reviews* 1:183–205.

Gathercole, S., and A. D. Baddeley. 1993. *Working memory and language.* London: Psychology.

Gray, J. R., C. F. Chabris, and T. S. Braver. 2003. Neural mechanisms of general fluid intelligence. *Nature Neuroscience* 6:316–322.

Kane, M. J., L. H. Brown, J. C. McVay, P. J. Silvia, I. Mylin-Germeys, and T. R. Kwapil. 2007*a*. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. *Psychological Science* 18:614–621.

Kane, M. J., A. R. A. Conway, T. K. Miura, and G. J. H. Colflesh. 2007*b*. Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33:615–622.

Kane, M. J., and R. W. Engle. 2002. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual differences perspective. *Psychonomic Bulletin and Review* 9:637–671.

Kane, M. J., D. Z. Hambrick, S. W. Tuholski, O. Wilhelm, T. W. Payne, and R. W. Engle. 2004. The generality of working-memory capacity: a latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General* 133:189–217.

Kiewra, K. A., and Benton, S. L. 1988. The relationship between information processing ability and notetaking. *Contemporary Educational Psychology* 13:33–44.

Kimberg, D. Y., M. D'Esposito, and M. Farah. 1997. Effects of bromocriptine on human subjects depend on working memory capacity. *Neuroreport* 8:3581–3585.

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

Klein, K., and W. H. Fiss. 1999. The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, and Computers* 31:429–432.

Kunde, W., J. Müsseler, and H. Heuer. 2007. Spatial compatibility effects with tool use. *Human Factors* 49:661–670.

Kyllonen, P. C., and R. E. Christal. 1990. Reasoning ability is (little more than) working-memory capacity?! *Intelligence* 14:389–433.

Kyllonen, P. C., and D. L. Stephens. 1990. Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences* 2:129–160.

Logie, R. H. 2003. Spatial and visual working memory: a mental workspace. In *The psychology of learning and motivation*, vol. 42. D. E. Irwin and B. H. Ross, eds. Pp. 37–38. San Diego, CA: Academic Press.

Logie, R. H., S. Della Sala, M. Liacona, P. Chalmers, and V. Wynn. 1996. Group aggregates and individual reliability: the case of verbal short-term memory. *Memory and Cognition* 24:305–321.

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.

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

Redick, T. S., and R. W. Engle. 2006. Working memory capacity and attention network test performance. *Applied Cognitive Psychology* 20:713–721.

Reuter, M., U. Ott, D. Vaitl, and J. Hennig. 2007. Impaired executive control is associated with a variation in the promoter region of the tryptophan hydroxylase 2 gene. *Journal of Cognitive Neuroscience* 19:401–408.

Richeson, J. A., and J. N. Shelton. 2003. When prejudice does not pay: effects of interracial contact on executive function. *Psychological Science* 14:287–291.

Schmader, T., and M. Johns. 2003. Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology* 85:440–452.

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

Unsworth, N., and R. W. Engle. 2007*a*. The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological Review* 114:104–132.

———. 2007*b*. On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin* 133:1038–1066.

Unsworth, N., R. P. Heitz, J. C. Schrock, and R. W. Engle. 2005. An automated version of the operation span task. *Behavior Research Methods* 37:498–505.

Unsworth, N., J. C. Schrock, and R. W. Engle. 2004. Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30:1302–1321.