SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

Dr. Willard Larkin
Program Manager for Life Sciences,
AFOSR/NL
801 North Randolph Street
Arlington, VA 22203-1977

☐ Reprint (Orig + 2 copies)

☐ Technical Report (Orig + 2 copies)

☐ Manuscript (1 copy)

☐ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: F49620-97-1-0041

REPORT TITLE: Final Technical Report for Grant Working memory, controlled attention and task switching

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Randall W. Engle

20000417 215

DTIC QUALITY INSPECTED 2

Tulsa 2313 BX

AFRL-SR-BL-TR-00-

0109

REPORT DOCUMENTATION PAGE

-0188

Public Reporting burden for this collection of information is estimated to average 1 hour p ns, searching existing data sources, gathering and maintaining the data needed, and complet omment regarding this burden estimates or any other aspect of this collection of information, inclu ion Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 31 March 2000	3. REPORT TYPE AND DATES COVERED Final Technical Report
4. TITLE AND SUBTITLE Working memory, controlled attention and task switching		5. FUNDING NUMBERS F49620-97-1-0041
6. AUTHOR(S) Randall W. Engle and Michael J. Kane		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Georgia Tech Research Corporation Georgia Institute of Technology Atlanta, Georgia 30302		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research Bolling AFB DC, 20332-6448		10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words) Eight sets of studies are reported that support a model my colleagues and I have developed which views WM as a system consisting of those long-term memory traces active above threshold, the procedures and skills necessary to achieve and maintain that activation and limited-capacity, controlled attention. The specific features of this model include: (1) Domain-free, limited-capacity controlled attention, (2) Domain-specific codes and maintenance (articulatory loop and visuo/spatial sketchpad are two examples but the potential number of such codes is large), (3) Individual differences in both 1 and 2, but individual differences in capacity for controlled processing are general and probably the mechanism underlying general fluid intelligence. While people can, with practice and expertise, circumvent the abiding limitations of controlled attention in quite specific situations, the limitations re-emerge in novel situations and even in the domain of expertise if the situation calls for controlled processing. (4) Limited-capacity controlled processing is required for maintaining temporary goals in the face of distraction and interference, and for blocking, gating and/or suppressing distracting events., (5) The dorsolateral prefrontal cortex (PFC) and associated structures mediate the controlled processing functions of WM. We also argue that individual differences in controlled-processing represent differences in functioning of the PFC.

14. SUBJECT TERMS			15. NUMBER OF PAGES 43
			16. PRICE CODE
17. SECURITY CLASSIFICATION OR REPORT **UNCLASSIFIED**	18. SECURITY CLASSIFICATION ON THIS PAGE **UNCLASSIFIED**	19. SECURITY CLASSIFICATION OF ABSTRACT **UNCLASSIFIED**	20. LIMITATION OF ABSTRACT **UL**

NSN 7540-01-280-5500
(Rev.2-89)
Prescribed by ANSI Std. 239-18

Standard Form 298

298-102

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO

Final Report

Grant: ***F49620-97-1-0041***
Grantee: ***Randall W. Engle***
Title: ***Working memory, controlled attention and task switching.***

Unrestricted distribution, unclassified

31 March 2000

Prepared for:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
BOLLING AFB DC, 20332-6448

Eight sets of studies are reported that support a model my colleagues and I have developed which views WM as a system consisting of those long-term memory traces active above threshold, the procedures and skills necessary to achieve and maintain that activation and limited-capacity, controlled attention. The specific features of this model include:

(1) Domain-free, limited-capacity controlled attention.

(2) Domain-specific codes and maintenance (articulatory loop and visuo/spatial sketchpad are two examples but the potential number of such codes is large).

(3) Individual differences in both 1 and 2, but individual differences in capacity for controlled processing are general and probably the mechanism underlying general fluid intelligence. While people can, with practice and expertise, circumvent the abiding limitations of controlled attention in quite specific situations, the limitations re-emerge in novel situations and even in the domain of expertise if the situation calls for controlled processing.

(4) Limited-capacity controlled processing is required for maintaining temporary goals in the face of distraction and interference, and for blocking, gating and/or suppressing distracting events.

(5) The dorsolateral prefrontal cortex (PFC) and associated structures mediate the controlled processing functions of WM. We also argue that individual differences in controlled-processing represent differences in functioning of the PFC.

A number of intellectual influences have served to shape our thinking about working memory (WM) and its evolution as a construct from that of short-term memory. One is that, in thinking about the nature of these constructs, both behavioral and biological, we need always be mindful of the tasks that we use to measure them. It is important to understand that tasks vary in validity as measures for their putative constructs. Further, there is no such thing as a pure measure of any construct, including short-term and WM.

Another influence is an approach to developmental psychology often called 'neo-Piagetian,' although the ideas can easily be traced back to Baldwin (1894), one of Piaget's early influences. Baldwin and others argued that memory span tasks reflect ability to maintain attention, a fundamental aspect of intellectual abilities, both for the developing human and across individuals at given stages of development (Case, 1985; Pascual-Leone & Baillargeon, 1994; Piaget, 1926). At least in studies with normal adults, however, simple digit and word span tasks do not consistently and reliably predict such mainstays of higher-level cognition as reading or listening comprehension (Perfetti & Lesgold, 1977).

Daneman & Carpenter (1980) developed a WM span measure called the reading span task that correlated reasonably well with language comprehension measures. Subjects read (or listened to) a list of 2-6 sentences. Afterward, the subject recalled the last word of each sentence. This critical measure, the number of sentence-final words recalled, is, on the face of it, very much like a simple word-span task. Turner & Engle (1989) developed a similar task in which the subject solved a string of arithmetic operations and then read aloud a word that followed the string. After a series of such operation-word strings, the subject recalled the words.

Both the reading-span task and the operation-span task are really dual tasks that require the subject to do something (read a sentence or solve an operation string) and, separately and interleaved with this task, to keep track of an evolving and growing list of words. These span tasks, and others developed since, apparently reflect an ability that is fundamentally important to higher-level cognition because measures of working-memory capacity reliably predict performance in a wide variety of real-world cognitive tasks. Significant relationships with measures of WM capacity have been reported for reading comprehension (Daneman & Carpenter, 1980, 1983), language comprehension (King & Just, 1991; MacDonald, Just & Carpenter, 1992), learning to spell (Ormrod & Cochran, 1988), following directions (Engle, Carullo & Collins, 1991), vocabulary learning (Daneman & Green, 1986), notetaking (Kiewra and Benton, 1988), writing (Benton, et al., 1984), reasoning (Kyllonen & Christal, 1990), and complex learning (Shute, 1991; Kyllonen & Stephens, 1990).

Two questions have guided our work over the past twelve years. The first is **'What accounts for individual differences on measures of WM capacity?'** More specifically, 'What is measured by the complex tasks that is <u>also</u> important to higher-level cognitive tasks?' The second question is **'What do results of studies on individual differences in WM capacity tell us about the nature of WM in general?'** Our attempts to answer these questions have used a combination of both regression studies and experimental or ANOVA-based studies using extreme groups. The extreme groups are individuals who score in the upper quartile (high span) and lower quartile (low span) on a variety of WM capacity tasks such as the reading span and operation span tasks.

We think of 'working memory' as a system consisting of: (a) a store (short-term memory) in the form of long-term memory traces active above threshold, (b) processes for achieving and maintaining that activation, and (c) controlled attention. However, when we refer to 'working-

memory capacity', we mean the capacity of just one element of the system: controlled attention. We do not mean the entire WM system, but rather the capabilities of the limited-capacity attention mechanism which Baddeley and Hitch (1974; see also Baddeley & Logie, 1999) called the central executive. Thus, we assume that WM capacity is not really about storage or memory per se, but about *the capacity to focus and sustain attention*. We think this concept corresponds to what Norman and Shallice (1986) called the "Supervisory Attentional System," and is related to what Posner and Snyder (1975) and Schneider and Shiffrin (1977) referred to as "controlled attention." The central executive is also likely related to the anterior attention system, i.e., the prefrontal cortex and anterior cingulate, proposed by Posner and Peterson (1990; see also Gevins, Smith, McEvoy and Yu, 1997).

Our attentional view thus differs from traditional notions of short-term memory (STM) or WM (e.g., Miller, 1956; Daneman & Carpenter, 1980), which emphasize how large or small working-memory capacity is, or how much or how little working memory can hold. The dependent variable measured in WM capacity tasks is, indeed, the amount of information that can be recalled in particular contexts. Nonetheless, we emphasize the importance of working-memory capacity for maintaining focus on even a *single* stimulus, goal, or response production, especially if there are significant sources of interference or distraction present. As we will make evident below, WM theories that focus on the *size* of WM capacity do not allow for the empirical predictions or data that are accommodated by our framework. By our view, individual differences on measures of WM capacity primarily reflect differences in capability for controlled processing. Thus, these individual differences will only be reflected in situations that either encourage or demand controlled attention (Conway & Engle, 1994; Engle, Conway, Tuholski & Shisler, 1995; and Rosen & Engle, 1997). Such situations include: (a) when task goals may be lost unless they are actively maintained in WM; (b) where actions competing for responding or response preparation must be scheduled; (c) where conflict among actions must be resolved to prevent error; (d) where there is an incentive to maintain some element of task information in the face of distraction and interference; (e) where there is incentive to suppress or inhibit information irrelevant to the task; (f) where error monitoring and correction are controlled and effortful; and (g) when controlled, planful search of memory is necessary or useful.

Our proposal, then, is that working-memory capacity reflects the ability to apply activation to memory representations, to either bring them into focus or sustain them in focus and this ability is particularly important in the face of interference or distraction. We have also argued recently that controlled processing capability is necessary in the case of suppression to either dampen activation of representations or otherwise remove them from focus (Engle, Conway, Tuholski and Shisler, 1995). This attention capability is domain-free and therefore individual differences in this capability will reveal themselves in a variety of tasks. We have also argued (Engle & Oransky, 1999; Kane & Engle, 2000a) that differences in WM capacity correspond to individual differences in the functioning of the prefrontal cortex.

Working Memory = Short-Term Memory + Controlled Attention

But, do the different tasks we refer to as measures of WM capacity really reflect the same underlying construct? Further, do such tasks measure something different from traditional STM tasks and, if so, what distinguishes them? Engle, Tuholski, Laughlin and Conway (1999) addressed these questions in work funded by the previous AFOSR contract. The study was directed at Cowan's (1995) distinction between STM and WM and the relationship of these constructs to general fluid intelligence (*g*F). Cowan considers STM a subset of WM. STM is a simple storage component meaning those memory units active above some ambient baseline, while WM refers to a system consisting of that storage component as well as an attention

component. This view is consistent with Baddeley & Hitch's (1974) original model, except the concept of STM is retained and viewed as consistent with the slave systems (the phonological loop and the visuo-spatial sketchpad). Thus, Cowan's view is that WM consists of the contents of STM plus controlled attention.

Engle et al. (1999) performed an analysis of the unique and shared variance in tasks thought to reflect STM and WM, the underlying factor structure of that variance and the extent to which a theory of the two constructs is supported by structural models of the variance in the tasks. Figure 1 shows a schematic of a measurement model based on those constructs. Controlled attention can be used to achieve activation of long-term traces through controlled retrieval, to maintain activation through various means or to dampen activation through inhibition. STM consists of those traces active above threshold, with loss of activation due to decay and/or inhibition. The short-term traces could be primarily in the form of phonological features or visual features or, indeed, many other features; however, the traces obey the same principles of forgetting, interference, etc., regardless of format. Some very small number of those traces receive increased activation as a result of salience to the current task goal or from endogenous activation resulting from emotional salience.

Cowan defined STM as a subset of WM. Thus, at a conceptual level, variance shared between WM tasks and STM tasks should reflect the short-term component. The variance left over (or residual) in WM tasks after removal of the variance shared by the two tasks should reflect the controlled attention or central executive component of WM. However, variance could be shared between STM and WM tasks at several other levels as well depending on: (a) the specific mental procedures, skills and strategies used to achieve and maintain activation, (b) the specific materials used in the tasks, and (3) how much the STM task requires of the central executive component. For example, if both WM and STM tasks used digits in a serial recall task, then individual differences in perceptual grouping or chunking skills, skill at phonological coding, and speed of phonological rehearsal would all contribute to shared variance between the WM and STM tasks. Likewise, skill at imaginal coding and speed of manipulating visual and spatial images would contribute to shared variance to the extent that both types of tasks make use of visual/spatial materials and require similar mental procedures. However, to the extent that STM and WM tasks require <u>different</u> procedures, there would be unique variance in each task associated with individual differences in the domain-specific skills and abilities.

Figure 1

Relationship of components of Working Memory system
Any given WM or STM task reflects all components to some extent

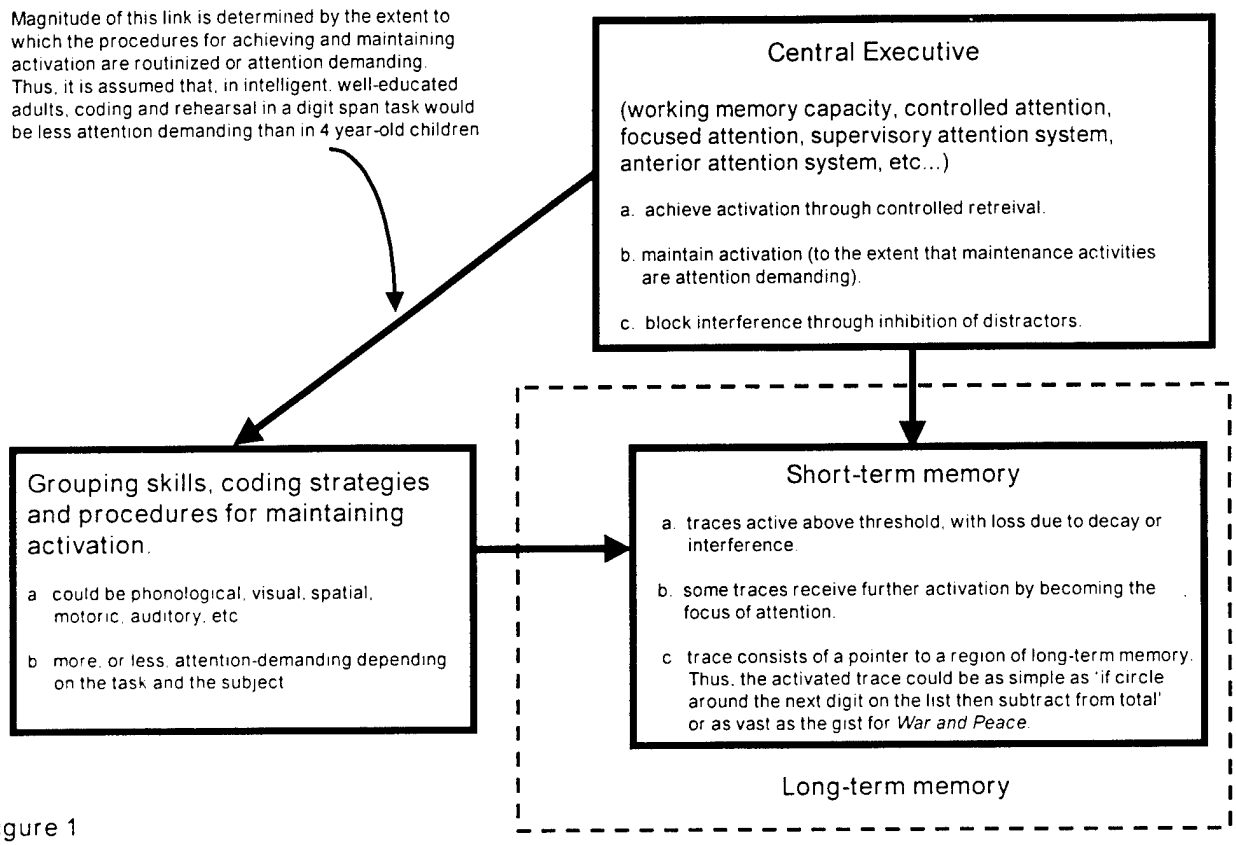


Figure 1

Another point at which variance would be shared between the two types of tasks is the extent to which they both rely on controlled attention. To the extent that the different mental procedures do not rely on limited-capacity controlled attention, the degree of shared variance would depend on whether the two tasks shared specific materials or procedures. However, if the procedures used for the STM and WM tasks both require controlled processing, they would tap variance common to the central executive. That would be true even if the materials used in the two types of tasks were from different domains and the procedures required different coding strategies. Thus, there are no 'pure' STM or WM tasks. Tasks are graded in the extent to which they are a 'good' STM or WM task depending on the overlap in the task content or materials, in the procedures used to perform the task and the extent to which the tasks require controlled processing.

We (Kane & Engle, 2000b) and others (c.f., Duncan, 1995) have argued that the construct of working-memory capacity is isomorphic with the capacity for controlled processing which has a strong relationship to general fluid intelligence or *g*F. Fluid intelligence is reflected in the ability to solve novel problems and is putatively non-verbal and relatively culture free (Horn & Cattell,

1967). The Raven's Standard Progressive Matrices (Raven, Court & Raven, 1977) and the Cattell Culture Fair Test (Cattell, 1973) were used as our *g*F measures. Both tests consist of visual patterns and the choice of a pattern that would complete the larger target pattern or that would fit a sequence of patterns.

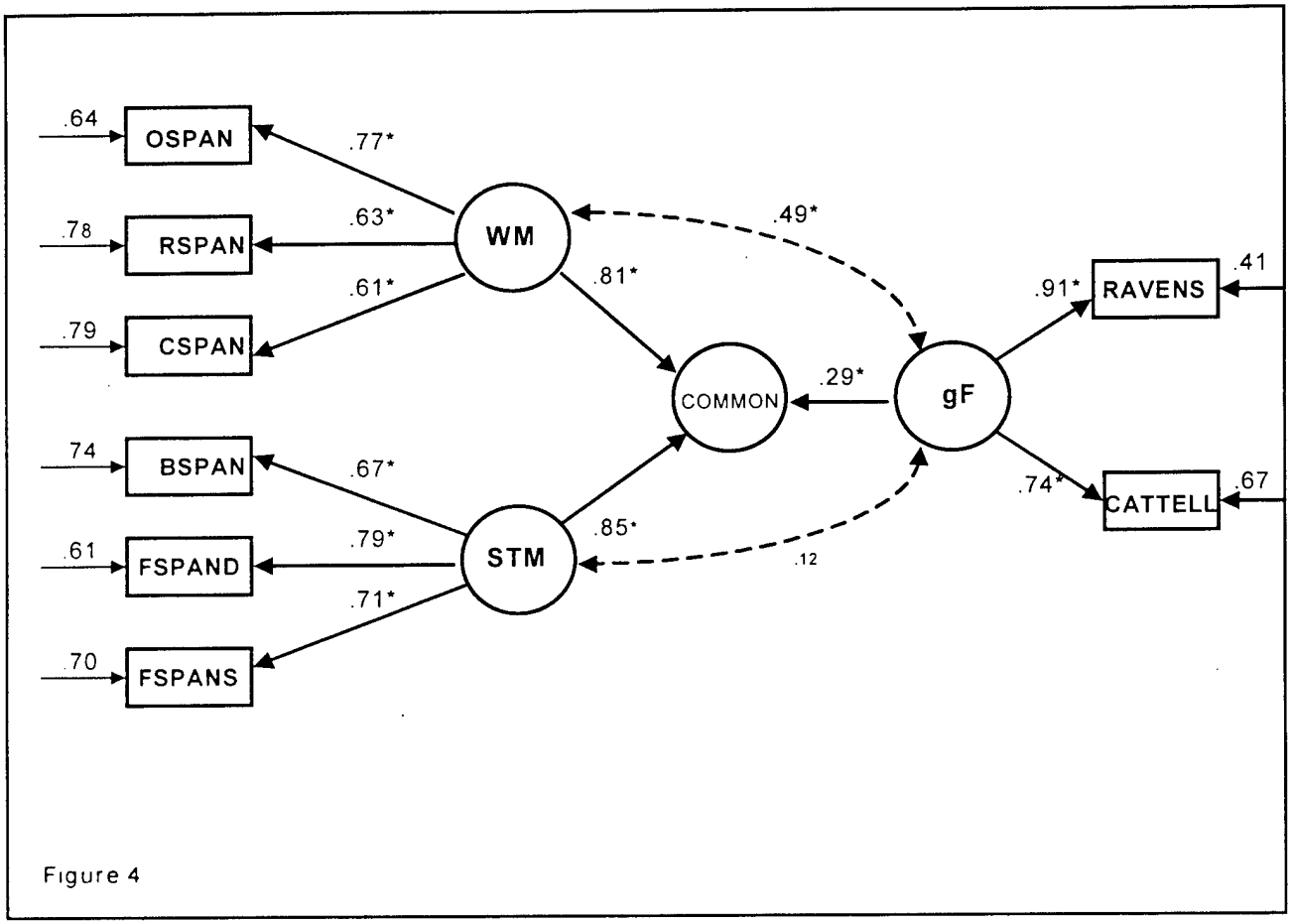
Engle et al. (1999) tested 133 subjects on a variety of tasks which logic and research had led us to believe would be good WM tasks. These included reading span, operation span and counting span (a form of the task used by Case, Kurland and Goldberg, 1982, but modified to require controlled counting). We also used tasks we thought would be good STM tasks including forward-word span with dissimilar words, forward-word span with similar (rhyming) words and backward-letter span with dissimilar letters. We also tested all subjects on full versions of the Ravens and Cattell Culture Fair tests. Several other tasks were also used, including the ABCD and Continuous Opposites tasks from the CAM4 battery (Kyllonen & Crystal, 1990), keeping track task (Yntema, 1963), the primacy and recency portion of immediate free recall, the random generation task and scores were obtained on Verbal and Quantitative Scholastic Aptitude Test but those variables are not discussed here.

One question was whether all the memory tasks reflected a single construct or whether a two-factor model (i.e., STM and WM) was necessary to account for the pattern of variance among the tasks. A second question was whether, at the latent variable level, after variance common to WM and STM was removed, the WM residual variance (which should reflect controlled attention) would correlate with the residual for *g*F, and, whether the STM residual would also correlate with the *g*F. Of course, the STM residual should not correlate with the *g*F residual if it reflects only error as our logical analysis would suggest.

Confirmatory factor analysis and structural equation modeling were performed on the 3 tasks we thought would be good STM tasks and the 3 we thought would be good WM tasks. The goodness-of-fit results showed that the two-factor model was a significantly better fit than the single-factor model. Thus, the model required a separate construct for WM capacity and for STM even though the two constructs were strongly related. The next question was whether the two constructs showed any relationship to general fluid intelligence (*g*F) as indexed by the Ravens and Cattell Culture Fair tests. The analysis showed that the model required a strong connection between the WM latent variable and general fluid intelligence but did not require a connection between short-term memory and *g*F.

If our arguments about the relationship among the constructs of WM, capacity for controlled attention and *g*F are correct then we should be able to test whether controlled attention is related to general fluid intelligence. Starting with all the variance in complex WM measures, we should be able to remove that variance due to domain-specific materials and procedures. The residual variance in WM, after removal of the variance common to STM, should reflect the capacity for controlled processing. The WM residual should, therefore, strongly relate to the construct of general fluid intelligence. Thus, the line of logic is as follows: (a) complex-span tasks of WM capacity reflect the construct STM ***plus*** the construct of controlled attention, (b) the construct of controlled attention has a strong relationship with the construct of general fluid intelligence, (c) there is little or no relationship between the construct STM and the construct *g*F. Therefore, if, at the latent variable level, we partial out the variance common to the constructs WM and STM we would expect the residual variance in WM capacity to reflect controlled attention. This method of removing common variance to see what connections need to be there above and beyond the common factor has been used by Salthouse and his colleagues very effectively in cognitive aging research (Salthouse, 1991).

Figure 2



The results of our common-factor analysis are shown in Figure 2. The arrows leading to the rectangles reflects the residual variance and the numbers associated with the arrows between rectangles and circles reflects the standardized regression coefficient between the task (rectangle) and the latent variable. Solid lines reflect a significant link and dotted lines reflect a non-significant link. The numbers associated with the two curved lines reflect correlations between the residuals of the STM and WM latent variables after the variance common to the two latent variables is removed.

When the variance common to the STM and WM latent variables was removed, the correlation between the residual of WM and *g*F, which should reflect capacity for controlled attention, was sizable and highly significant. The link between the STM residual and *gF*, which should be entirely error, was, in fact, not significant. This finding is even more impressive if we consider that, despite the fact that STM and WM latent variables are highly correlated, this analysis shows them to be quite separable. Another note about the analysis is that earlier we argued that even the STM tasks would co-vary with the WM tasks to the extent that controlled processing was necessary. Thus, the analysis depicted in Figure 2 actually is a conservative estimate of the correlation between the construct for controlled attention and *g*F. The common

factor, which also has a significant link to *g*F, removes the portion of that variance common to the WM and STM tasks.

The Engle et al. (1999) latent-variable study yields two important conclusions: (1) WM and short-term memory are highly related but separable constructs, and (2) when we partial out the variance ***common*** to the STM and WM, the link between the residual of WM and *gF*, which should theoretically be controlled attention, is high and highly significant. This lends support to the idea that the component of the WM tasks that is important to higher-order functioning is controlled attention.

Alternative explanations?

As we said earlier, the reading span and operation span are dual tasks and so should tap the subject's capability to focus or sustain attention when performing these task components. One component can be thought of as a processing task, reading the sentence or solving the arithmetic string, and the other can be thought of as a storage task, recalling the gradually lengthening list of words or digits. One possible alternative reason that WM capacity tasks correlate with higher-order measures of cognition is that individuals differ in skill on the processing component as a result of experience; this processing efficiency frees up resources to be used to rehearse the items in the storage task. Engle, Cantor & Carullo (1992) measured the time it took Ss to perform the processing component of the reading span and operation span tasks (reading the sentences or performing the arithmetic operations). They then used these times as measures of processing efficiency and/or the extent to which subjects traded off time on the processing component for time on the storage component. If the processing-skill explanation of the relationship between WM measures and measures of higher-order cognition is correct, then when the measure of processing time is partialed out of the relationship between the complex span score and a measure of reading ability, the correlation should significantly decrease if not disappear. However, Engle, et al. showed that partialing out the processing-time measure led to no decrease in the correlation between the span score and the Verbal Scholastic Aptitude Test score.

A study by Conway and Engle (1996) approached this same issue differently. Instead of statistically controlling for processing skill, they attempted to equate, across subjects, the processing demands of the span task. The logic was quite simple. If the relation between the WM measure and reading comprehension is a result of a trade-off between the processing and storage components, then equating the difficulty of the processing component across subjects should eliminate the span X reading comprehension correlation. On the other hand, if the correlation between the WM measure and comprehension is a result of a controlled attention capability beyond the trade-off, then equating subjects on the processing task should have no affect on the correlation. Conway and Engle pre-tested subjects on operations of the type used in the operation-span task but differing greatly along a dimension of difficulty. The pre-test determined the point on the difficulty dimension at which each subject solved operations accurately 75%, 85% and 95% of the time. Each subject then received an operation-word-span task in which the sets of operations were created specifically for that subject. One set was created separately at each of the 75%, 85%, and 95% level, specifically for each subject. One question was whether the correlation between the operation-span score and reading comprehension would disappear since the procedure equated the processing skill component. The second was whether the correlations between the resulting span scores and reading comprehension would differ as the processing component of the span test varied in difficulty from 75% to 95% accuracy. The answer to both questions was a resounding *no*. The correlation between the span score and Verbal Scholastic Aptitude test for the easy 95%

condition was .62 and for the difficult 75% condition the correlation was .54. These compare to a correlation of .59 for the same sample of subjects on an operation span test in which all subjects receive the same set of operations. Conway and Engle also measured the time each subject took to perform the operation-processing component in the various tasks and statistically removed this solution time from the above correlations. The effects were virtually identical, with a partial correlation of .60 for the easy condition and .52 for the difficult condition. Clearly, skill on the processing component had no bearing on the significant and high correlation between this measure of WM capacity and a measure of reading ability.

These findings show, quite convincingly, that the correlations between measures of WM capacity and higher-order cognitive tasks do not result from skill in the specific processing components of the WM tasks or individual differences in strategic allocation of attentional resources. The findings support our view that the critical feature of the tasks accounting for the correlations with higher-order cognition is some aspect of controlled attention, and that it represents a rather abiding characteristic of the individual.

Working Memory, General Fluid Intelligence, and Controlled Attention

The evidence discussed thus far certainly suggests that individual differences in WM and in general fluid intelligence are significantly related to one another and that they probably do not depend on differences in general knowledge or specific procedural skills. Our discussion of latent variable results and individual differences work on WM tasks themselves tentatively suggest that controlled attention capabilities lie at the heart of individual differences in working-memory capacity.

A controlled-attention view of WM capacity is consistent with Baddeley's (1986, 1993) proposal that the central-executive component of WM may be analogous to the Supervisory Attentional System (SAS) described by Shallice and colleagues (Norman & Shallice, 1986; Shallice & Burgess, 1993). The SAS is hypothesized to be a conscious control mechanism that resolves interference between activated action schemas. In particular, when a pre-potent action is environmentally triggered, but conflicts with the individual's goal state, the SAS biases the action-selection process by providing additional activation to a more appropriate action schema and by inhibiting the activation of the inappropriate schema. The SAS thus allows attentional control over action by providing a means with which to override interference from powerful environmental stimuli and habitual responses.

Initial evidence for the theoretical connections among WM capacity, the central executive, and attentional control came from Conway and Engle (1994), who demonstrated that high and low working-memory span participants differed in their speed of retrieval from long-term memory only under conditions of interference. In a modification of the Sternberg (1966) memory-scanning task, participants learned different-sized sets of unique digit-letter pairs (e.g., *2-RW*; *4-BKNZ*). Then, after extensive learning that matched the span groups, participants took a speeded recognition test. High and low spans were equivalently fast to recognize studied versus mismatched foil pairs. However, in another experiment each letter appeared in two different memory sets (e.g., *2-RW*; *4-BKRZ*), and so a considerable level of response competition, or set confusion, was likely to occur at retrieval. Here, low spans were significantly slowed relative to high spans, and relative to low spans in the non-interference condition. Low spans thus appeared to be more vulnerable to interference than did high spans, who may have used attention to inhibit the interfering link between the shared letter sets at retrieval.

Recent studies have provided more direct evidence that interference differences between high and low span participants reflect controlled attention differences. For example, Rosen and Engle (1997) tested high and low spans in a category fluency test in which they recalled as many animal names as they could for 10 – 15 min. High spans generated more animal names than did low spans, and the difference between groups increased across the recall period, a finding indicative of span differences in susceptibility to output interference. That is, successful fluency across long intervals requires strategically searching for low dominance exemplars while blocking the re-retrieval of the already-recalled, high dominance exemplars such as *dog, cat, cow*, and *horse*. Most important for present purposes, high spans' superiority in fluency was eliminated in a second experiment in which participants recalled animal names while also performing a secondary digit-tracking task. When high spans had their attention divided, their fluency dropped to the level of low spans' fluency. Moreover, low spans were unaffected by the secondary-task load. These findings suggest that high spans used controlled processing to attain their high fluency performance, because their performance dropped when attention was occupied by the secondary task. In contrast, low spans did not appear to use controlled processing during fluency recall (perhaps relying on automatic spreading activation among exemplars), because their poorer performance remained the same under no-load and load conditions.

Investigating Working Memory and Controlled Attention in the Present AFOSR Project

Overview of Studies

Below we present 8 studies in which we explore the relation between WM capacity and attention control. In Study 1, three experiments were conducted to test the role of controlled attention in fluency across different stimulus domains. In Study 2, two experiments were conducted to test the contributions of WM capacity and controlled attention to resisting memory interference in a proactive interference task. In Study 3, we tested whether high WM subjects were more flexible in their allocation of attention in space than were low WM subjects in a visual orienting task. In Study 4, we tested whether high WM subjects (and high intelligence subjects) were better able to selectively focus attention amidst distraction in the Stroop task. In Study 5, two experiments tested whether, in an antisaccade task, high WM subjects were better able to suppress reflexive eye movements according to task goals compared to low WM subjects. In Study 6, 2 experiments tested whether high WM subjects were better able to rapidly switch between task sets, abandoning a prior task set in favor of uploading a new set, compared to low WM subjects. In Study 7, 2 experiments tested whether high WM subjects were better able to sustain a memory representation of a non-verbal stimulus across filled and unfilled delays. In Study 8, a large-scale latent variable study tested whether domain-specific knowledge could compensate for WM related differences in language comprehension.

STUDY 1: WM Capactiy, Divided Attention, and Fluency Across Domains

Research in intelligence domain (e.g., Phillips, 1997), the aging domain (e.g., Troyer, Moscovitch & Winocur, 1997) and the neuropsychology domain (e.g., Milner, 1964; Moscovitch, 1994) suggests that different fluency tasks may make different demands on attention control and prefrontal cortex circuitry. That is, dividing attention may impair letter fluency tasks, where subjects report all the words they can beginning with a particular letter, more that it impairs semantic fluency tasks, where subjects report all the words they can from a particular taxonomic category. Likewise, damage to prefrontal cortex (PFC) structures may impair letter fluency more than semantic fluency.

However, Rosen & Engle (1997) clearly found substantial WM-related differences in semantic fluency, as well as substantial dual-task costs in semantic fluency for high WM subjects. Moreover, in our review of the PFC literature, Kane and Engle (2000) noted that failures to find PFC-damage deficits in fluency appeared to have more to do with the *duration* of the fluency task than its domain: Minimal PFC-related deficits are seen in fluency tasks lasting fewer than 2 min, but substantial PFC-related deficits are seen in tasks lasting 5 min or more. These duration findings make sense if a significant source of difficulty in these tasks is selectively blocking the re-retrieval of items that were already recalled earlier in the session (see Rosen & Engle, 1997). Particularly across long recall periods, fluency may be enhanced by selectively directing attention away from, or inhibiting, the high-probability words that were recalled earlier in the session. After a subject exhausts the first words that easily come to mind, executive attention should be necessary to search memory further and to avoid repeating the recall of previously recalled, and therefore highly activated, exemplars.

Thus, we conducted three experiments to test for divided attention effects in different types of fluency task (semantic, letter, and spatial), and to test whether dual-task costs change with the passage of time through the fluency task.

Experiment 1 Method. WM span was not measured for subjects in either Experiment 1, 2, or 3, but they were drawn from the undergraduate population of Georgia Institute of Technology, a population with an average WM span that is quite large. Subjects spent 10 min generating the names of as many animals as they could for 10 min, and thus Experiment 1 used a *semantic* fluency task. Half the subjects were required to perform an additional task at the same time as the fluency task, namely a finger-tapping task designed by Moscovitch (1994) to mimic the effects of PFC lesions. Subjects continuously tapped a pseudo-random sequence on a computer keyboard: index finger – ring finger – middle finger – pinkie. The keyboard was interfaced with a computer program that recorded the subject's tapping speed and accuracy, and also provided auditory on-line feedback (beeps) when subjects slowed their tapping below an individually calculated speed limit.

Experiment 1 Results. Divided attention significantly impaired fluency across all 10 min. of the task In the first minute of recall, the secondary tapping task reduced fluency by 22.9%; in the last, 10th minute of recall, tapping reduced fluency by 37.9%. Thus, although the dual-task cost increased with time in the task, as we predicted, it was even significant (and substantial) during the first minute of recall. Note here that the tapping task used here was very different from the digit-tracking task used by Rosen and Engle (1997) to divide subject's attention during fluency. Thus, with either a verbal or non-verbal secondary task, dividing attention substantially impairs semantic fluency.

Experiment 2 Method. The method followed exactly that of Experiment 1, except that a new set of subjects were engaged in a *letter* fluency task, under focused or divided attention, in which they recalled as many words as they could beginning with the letter "S."

Experiment 2 Results. Again, divided attention significantly impaired fluency across all 10 min of the task. In the first minute, tapping reduced fluency by 36.5%, and in the 10th minute it reduced fluency by 49.3%. Again, the dual-task impaired fluency more as the session proceeded, but the divided-attention cost was significant and substantial even in the first minute of recall. Somewhat consistent with prior work, however, was that divided attention appeared to impair letter fluency more than semantic fluency.

Experiment 3 Method. The method followed Experiments 1 and 2, except that a new set of subjects were engaged in a figural fluency task under single- or dual-task conditions.

In the figural task, subjects saw an arrangement of 6 dots onscreen, with each each dot labelled 1 through 6. The task was to generate as many figures as possible from the same 6 dots. A "figure" could represent anything from a single line segment between two of the dots, to a completely enclosed figure encompassing all 6 of the dots. Subjects responded verbally, just as in Experiments 1 and 2, by calling out the numbers of the dots, in order, that comprised each figure. Each figure had to be visually unique, if it were actually drawn. So, if a subject called out "1-2, stop" to indicate a line segment between points 1 and 2, the subject could not later call out "2-1, stop" and have it count as a new figure. All subjects first engaged in a brief practice session with a different set of dots before proceeding to the actual test session.

Experiment 3 Results. Again, divided attention significantly impaired fluency across the entire session. In the first minute of recall, dividing attention reduced fluency by 25.5%. In the 10th minute it reduced fluency by 50.9%. Thus, as in Experiments 1 and 2, but here with a spatial task, dividing attention with a finger-tapping task impaired fluency more as the session proceeded. But here, as in the other Experiments, the dual task costs even in the first minute were significant and substantial.

Our three experiments suggest that attention control is an important determinant of successful fluency across different task domains, including the spatial domain. That is, dividing attention during a fluency task significantly reduces recall. Moreover, controlled attention appears to become more critical to success as the recall period proceeds and already-recalled items must be selectively blocked from consideration.

STUDY 2: WM Capacity, Divided Attention, and Proactive Interference.

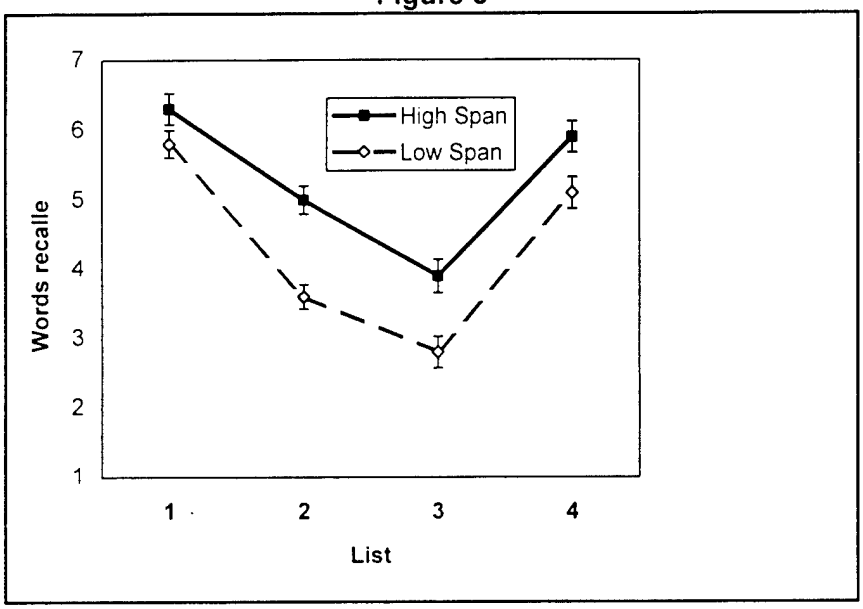
Kane and Engle (in press) reached similar conclusions from a proactive interference (PI) task, in which subjects studied and recalled three consecutive words lists from the same category (e.g., animals). During such tasks recall typically drops across each successive list, as the potential for PI from previous lists increases (e.g., Craik & Birtwistle, 1971; Wickens, Born & Allen, 1963). We hypothesized that interference resistance was in part responsible for the span differences in fluency reported by Rosen and Engle (1997), and observed in our Study 1, because resisting interference should require attention control. We therefore expected to find that high spans would be less vulnerable to PI than would low spans, and that only high spans' PI effects would increase under dual-task conditions.

Experiment 1 Methods. One hundred ninety-two undergraduates from the University of South Carolina participated in Experiment 1, identified from a much larger pool of subjects who had participated in the Operation-Word Span test of WM capacity (Turner & Engle, 1989). Ninety-six subjects were selected from the top quartile of the distribution (high spans), and 96 were selected form the bottom quartile (low spans). The PI buildup task was modeled after the Craik and Birtwhistle (1971) version of the Wickens (1970) PI buildup/release procedure. Subjects were visually presented with four different word lists, each followed by a 30 s distractor task, followed by the vocal free recall of the list for 30 s. Each list consisted of 10 words, presented one at a time on a computer monitor, at a rate of 1 word/2 sec. All subjects studied three related lists, followed by an unrelated list (e.g., three lists of "animals" followed by a list of "occupations"), to assess PI release. However, because neither WM nor load affected PI release, we will not discuss those results. Half of the subjects in each span group performed the PI buildup task under attentional load conditions. Subjects under load were required to

continuously tap a pseudo-random sequence with their fingers: index finger – ring finger – middle finger – pinkie (see Moscovitch 1994). In order to monitor finger-tapping performance, liquid crystal finger-pads were taped onto each subject's fingers. These finger-pads were interfaced with a computer program that recorded the subject's tapping speed and accuracy, and also provided auditory on-line feedback (beeps) when subjects slowed their tapping below an individually calculated speed limit.

Experiment 1 Results. Under "no load" conditions, low spans recalled significantly fewer words than did high spans (see Figure 3). More importantly for present purposes, relative to high spans, the low spans showed a more severe drop in the number of words recalled as lists 1-3

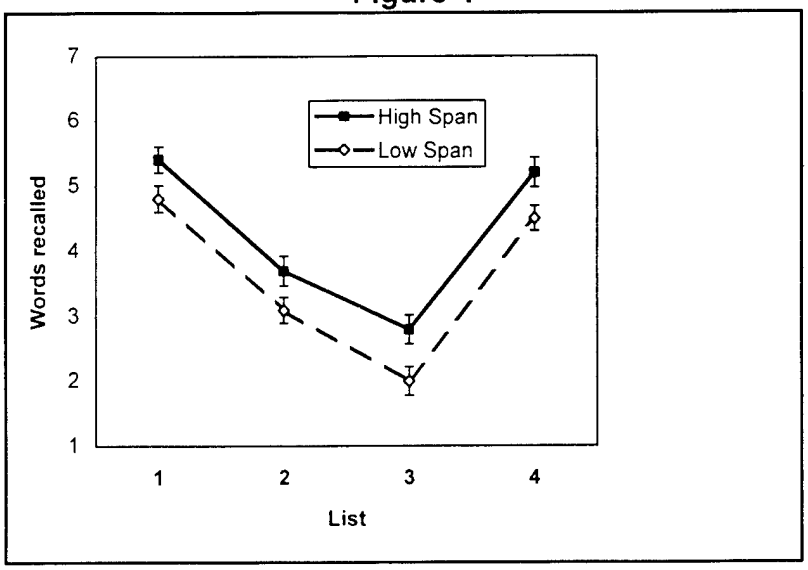
Figure 3



progressed; these findings indicated that low spans were more susceptible to PI buildup than were high spans. Under load (see Figure 4), high and low spans showed *equivalent* PI effects. Only high spans' PI effects increased under load, and so attentional load equalized high and low spans' interference susceptibility. That only high spans were impaired by load indicates that only they typically use controlled attention to combat PI; if low spans do not use attention to begin with, then they cannot show a further deficit when attention is divided.

Unfortunately, the finger-tapping apparatus often failed to collect data accurately, so we could not examine tapping performance. We remedied this in Experiment 2, where subjects simply tapped on the computer keyboard. Experiment 2 also addresses two more theoretically-motivated questions: 1) Does PI result primarily from an encoding failure, a retrieval failure, or both, across successive lists? 2) How attention demanding is encoding compared to retrieval? Both of these questions were investigated by manipulating the locus of the load task. Instead of tapping through all the PI-task components as in Experiment 1, subjects here tapped either while encoding each list, while retrieving each list, or neither.

Figure 4



The question of the locus of PI effects (at encoding or retrieval) was motivated by an older literature that debated whether PI resulted primarily from response competition from previously stored episodes at the time of target retrieval (e.g., M. Anderson & Neely, 1996; Postman & Underwood, 1973), or rather from a deficit in encoding each successively encountered episode (e.g., Hasher & Johnson, 1975; Postman, Stark & Burns, 1974). By manipulating load at either the encoding or retrieval phase of each list, we hoped to gain insight into where attention might be operating in the service of PI resistance.

The question of the attention demands of encoding versus retrieval arises from a recent set of studies by Craik and his colleagues (Craik, Govoni, Naveh-Benjamin & Anderson, 1996; see also Baddeley, Lewis, Eldridge & Thomson, 1984), which demonstrated that dividing subjects' attention while encoding word lists drastically reduced recall. In contrast, dividing subjects' attention during retrieval had very little impact on the number of words remembered. Craik et al. concluded from these studies that controlled attention is critical to memory encoding, and although retrieval may not completely attention-free, it can be maintained rather automatically once it is initiated. However, such conclusions contradict Rosen and Engle's (1997) findings that an attentional load impaired high spans' retrieval of information from long-term memory (e.g., their "fluency" in recalling animal names for 10 minutes; see also Baddeley et al.). Note that retrieval across such long periods is probably influenced in part by output interference increasing across the session (see Gruenewald & Lockhead, 1980), suggesting that attention at retrieval may be particularly important under interference conditions.

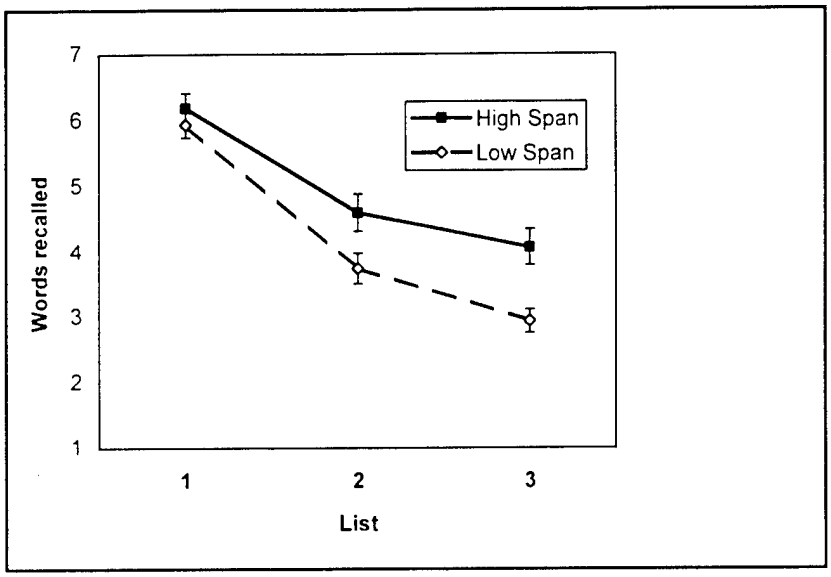
Experiment 2 Methods. One hundred eight high spans and 108 low spans (identified as in E1) from the University of South Carolina and Georgia State University participated in Experiment 2. The PI task was identical to that of Experiment 1, except that: 1) subjects studied only three lists instead of four (there was no PI release list); 2) subjects under load tapped either while encoding each list or while retrieving each list.

Experiment 2 Results. As in E1, under no-load conditions (see Figure 5) low spans again showed larger PI effects than did high spans. However, note that high and low spans did *not* differ in recall in the absence of interference, that is, on list 1. With respect to divided-attention effects, load conditions equalized the PI effects between high and low span subjects: Under

encoding load, however, high and low spans showed equivalent PI effects, with high spans' PI effects increasing under both load conditions and with low spans' PI effects not changing at all under either load condition.

Figure 5

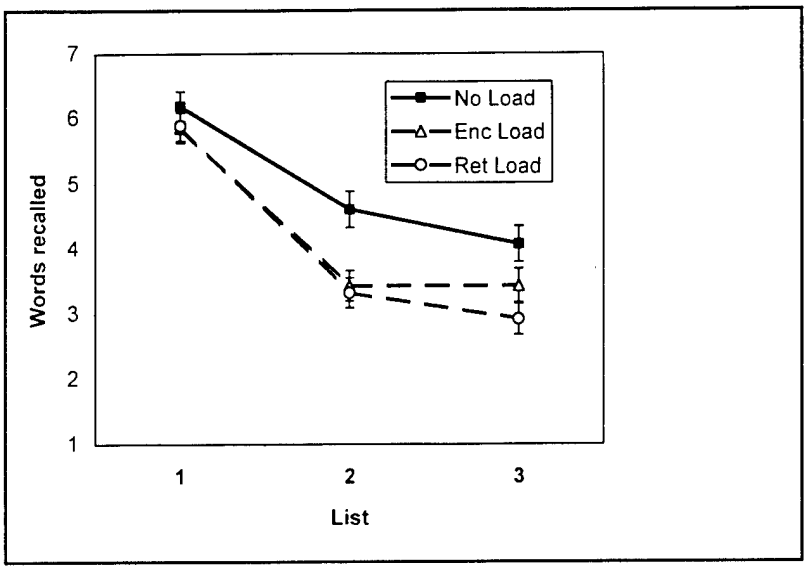
With regards to our first question, then, of the locus of PI effects, our data indicate that there is some contribution at both encoding and retrieval. For high spans, dividing attention with the tapping task increased their PI effects when the load was applied at either stage. High spans appeared use controlled attention at both encoding and retrieval to resist PI buildup. We further



speculate that low spans' lack of controlled attention at both encoding and retrieval contributed to their exaggerated PI effects under standard no load conditions.

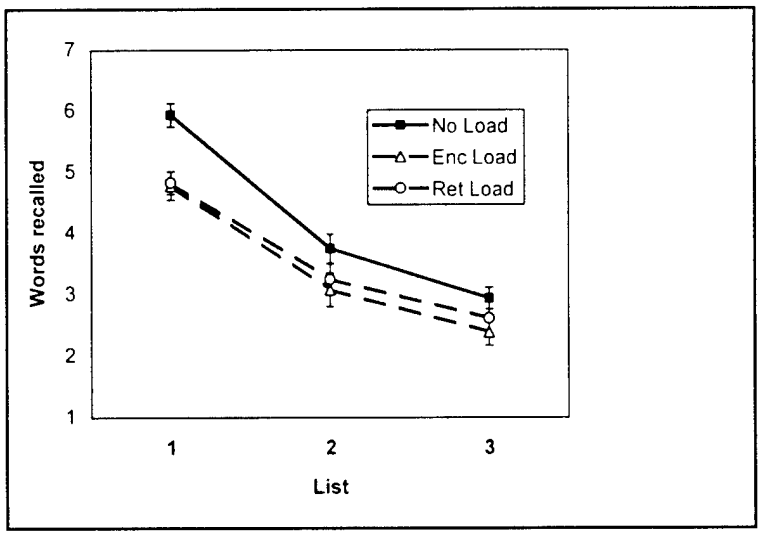
With regards to our second question about the attentional control of encoding versus retrieval processes, we can examine the loss in recall due to the load task both on recall list 1, before PI built up, and on lists 2 & 3, after PI had built (for high spans see Figure 6, for low spans see Figure 7).

Figure 6



On list 1, before PI influenced recall, high spans showed no effect of either encoding or retrieval load. High spans, then, appear to encode and retrieve rather automatically when there is little threat of PI. In contrast, low spans on list 1 showed significant load effects for both encoding and retrieval. Low spans thus appear to require controlled attention for both encoding and retrieval when PI is minimal. Note how the pattern changed by list 3, however, where PI was considerable. There, high spans under load showed a significant load effect. Under PI conditions, then, high spans were significantly impaired by divided attention requirements, suggesting that they normally use attentional control to resist PI. Low spans, in contrast, while they are impaired by the load task, are equally (if not less) impaired by load as they were on list 1. Thus, the presence of PI does nothing to change low spans' attentional allocation.

Figure 7



Together with the findings discussed previously, Kane and Engle (in press) supported the notion that WM capacity is related to controlled attention. Higher WM individuals demonstrated better

(or more) use of attention to resolve interference during memory encoding and retrieval than did lower WM individuals. However, if WM capacity reflects a relatively low-level general attentional capability, then span differences in controlled attention should be detectable outside the context of memory tasks. High and low span individuals should differ in their performance of more "molecular" attention tasks that make minimal storage demands and that have no requirements to explicitly recall information from long-term memory.

STUDY 3: WM Capacity and the Spatial Allocation of Visual Attention

Bleckley, Durso, Crutchfield, and Khanna (2000) tested the attentional hypothesis of WM capacity by studying the allocation of visual attention. They specifically tested whether high and low span subjects might differ in the *flexibility* with which they allocate attention in space. High WM span people were predicted to attend selectively to the areas of the visual field that are important to the task while ignoring other areas, whereas low WM span individuals were predicted to allocate their attention less flexibly.

Bleckley, et al. (2000) borrowed the procedure of Egly and Homa (1984) who showed that subjects could allocate attention simultaneously to two discontinuous locations: to a central location and to one of three concentric rings surrounding the center. In Egly and Homa's study, subjects were asked to identify a briefly presented center letter and to localize a displaced letter simultaneously presented on one of the rings. When participants were given a valid cue indicating on which ring the letter would appear, localization accuracy, of course, improved. Of primary interest is what happened when the wrong ring was cued. When the letter appeared on a ring *outside* of the cued ring, performance suffered. This finding is consistent with either a spotlight model or any view of flexible allocation, because the target occurred outside of the focus of attention. More important is the finding that when the letter appeared on a ring *inside* of the cued ring, performance also suffered. Because accuracy decreased when the target appeared either outside the cued ring or inside the cued ring, Egly and Homa, ruled out a simple spotlight model of attention allocation. A spotlight would have been suggested if performance remained high when the target appeared inside the expected ring. Instead, subjects appeared able to restrict their attentional focus to the cued ring, at the cost of locations inside and outside of that ring.

In the Bleckley et al. (2000), study, when the displaced letter occurred outside the cued ring both low span and high span individuals located fewer letters correctly than on their respective validly cued trials. For letters appearing on a ring *inside* of the one that was cued, high spans again performed poorly compared to their validly cued performance. This indicates a flexible, "ring-shaped" allocation of visual attention for the high spans (as in Egly & Homa, 1984). The low span subjects, however, showed no invalid cue costs when the letter appeared on a ring inside the cued ring; they were just as accurate in identifying stimuli on rings inside the cued ring as they were on the cued rings themselves. This suggests a spotlight of attention (a la Posner, et al., 1980) or a gradient of attention (LaBerge, 1983) for the low spans subjects. Low spans do not appear to be able to restrict visual attention to as tightly a defined location as are high spans.

These results suggest that treating WM as controlled attention not only has heuristic value but captures a fundamental understanding of WM and attention. Without such an attentional conception of WM, it is difficult to explain why a dual-task memory measure such as the Operation-Span task would predict the shape in which visual attention can be allocated.

STUDY 4: WM Capacity, Goal Neglect, and Focused Attention in the Stroop Task

In order to further test our attentional framework for understanding WM capacity, we tested high and low WM subjects in the Stroop (1935) task (Kane, Sanchez & Engle, 1999). The "Stroop effect" refers to an increase in response latencies (and error rates) to name the color in which a word is printed when that word names an incompatible color, compared to when that word does not name a color. The Stroop effect is generally taken to reflect efficacy in dealing with strong, competing response tendencies, and thus to reflect the selective attending to one stimulus dimension and/or the inhibition of a competing stimulus dimension. However, prior work with the Stroop task suggested that its interference effects often do not correlate with intelligence measures (e.g., Jensen, 1965; Spilsbury, 1992), nor do they consistently distinguish patients with PFC damage (i.e., damage to the "anterior attention system") from those with parietal damage (e.g., Ahola, Vilkki & Servo, 1996; Butters, Kaszniak, Glisky, Eslinger & Schacter, 1994). If WM capacity so strongly linked to controlled attention, *g*F, and to the PFC, then these Stroop findings might be considered something of an embarrassment to our view.

However, recent computer-simulation models suggest that a critical determinant of Stroop performance is the maintenance of unusual task goals or demands in the face of interfering and competing events (Cohen et al., 1990; Cohen & Servan-Schreiber, 1991). That is, individual differences in Stroop interference may arise from differences in the ability to keep the task goal (i.e., "*ignore the word and respond to the color*") sustained in an active, easily accessible state in WM. It so happens that most studies attempting to link Stroop interference to *g*F and the PFC have all minimized such sustained-attention demands by having all interference-type trials appear in one blocked and continuous sequence. These blocked presentations may minimize WM involvement because the task requirements remain consistent from trial to trial – once beginning the interference block, there is never a trial on which word information is relevant to response. Here, then, the task environment serves to reinforce the task demands; every stimulus reminds the subject to ignore the word information.

Individual differences in interference should thus be exacerbated, and corresponding correlations with *g*F magnified, if working-memory demands were increased by requiring switching between mixed trial types. Note that a mixed-trial Stroop task is functionally a switching task – sometimes subjects can respond on the basis of the representation that achieves activation first (i.e., the word information), but sometimes they cannot. Although the explicit task instructions and goals are the same between mixed and blocked procedures, the implicit procedures that subjects follow may be quite different.

We therefore hypothesized that if subjects were tested with discrete-trial versions of the Stroop task, in which randomized computer-presented Stroop words were sometimes incongruent (i.e., "*BLUE*" printed in green), sometimes neutral (i.e., "*JLXR*" printed in green) and sometimes congruent (i.e.,"*GREEN*" printed in green), Stroop interference would correlate with WM and gF. In that case, word information is unpredictably congruent or incongruent with responses, so maintaining the *ignore* demands within WM should be more difficult. Indeed, Stroop research outside of the individual-differences domain has shown that interference effects increase in magnitude as the proportion of congruent trials in the task increases (for a review see MacLeod, 1991).

Methods. Sixty-one high span subjects and 73 low span subjects (as identified by the Operation-Span task) participated in a computerized Stroop task. Subjects saw one stimulus per trial on a computer screen, and their job was to name the color of each stimulus quickly and

accurately. Stimuli appeared in red, blue, or green. The stimuli themselves were the words RED, BLUE and GREEN, as well as three non-words, as controls (JKM, XTQZ, and FPSTW). Subjects' vocal response on each trial tripped a voice-activated relay, which allowed a connected computer to record the response latency; the color named by the subject was keyed in by the experimenter following each trial.

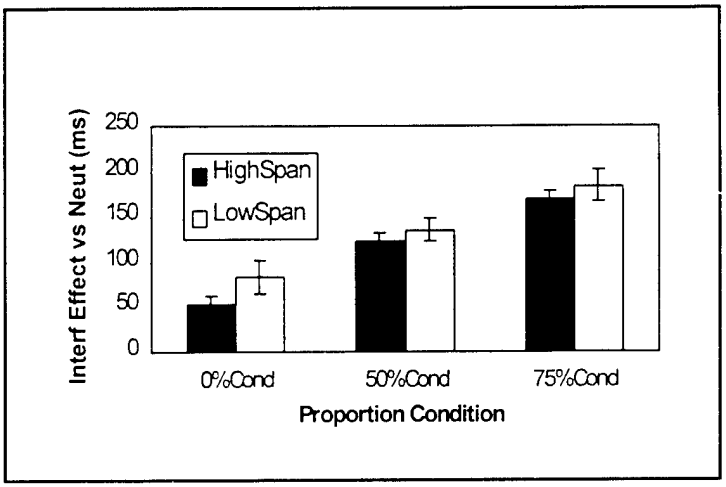
A "*READY*" signal appeared onscreen before each trial, and subjects initiated the trial by pressing a key. The screen was then blanked for 1000 ms, followed by a fixation signal appearing in the center of the screen ("*****") for 200 ms. The target stimulus then appeared in the center of the screen until the subject named its color aloud. After the experimenter logged the response, the *READY* signal re-appeared.

Subjects saw 288 experimental trials, after completing 36 practice trials. The composition of the 288 experimental trials differed among three between-subject groups. A *0%-Congruent* group was comprised of 21 high spans and 22 low spans. They saw 252 incongruent trials, in which the color and word were in conflict (e.g., "RED" presented in blue), and 36 neutral trials, in which a non-word appeared in color (e.g., "JKM" presented in blue). A *50%-Congruent* group was comprised of 21 high spans and 26 low spans. They saw 108 incongruent trials, 36 neutral trials, and 144 congruent trials, in which the color and word matched ("BLUE" presented in blue). A *75%-Congruent* group was comprised of 19 high spans and 25 low spans. They saw 36 incongruent trials, 36 neutral trials, and 216 congruent trials. Because these three groups had a different number of experimental trials in each condition, we analyzed data from only a sub-set of "critical" trials. All subjects saw the same 36 critical incongruent trials and 36 critical neutral trials. The *50%-* and *75%-Congruent* subjects also saw the same 36 critical congruent trials.

At the end of the Stroop task, all subjects had 25 min. to complete as much as the Ravens Advanced Progressive Matrices test as possible. The Ravens test is a non-verbal, matrix reasoning test that, when used in a large battery of psychometric tests, tends to load highly (and almost uniquely) on a second-order, general intelligence factor. Thus, the Ravens is widely considered to be a good measure of general fluid ability. We hypothesized that past failures to find correlations between Stroop and intelligence measures were tied to using 0%-Congruent conditions. Thus we expect higher Ravens scores to predict smaller Stroop interference effects in the *50%-* and *75%-Congruent* conditions only.

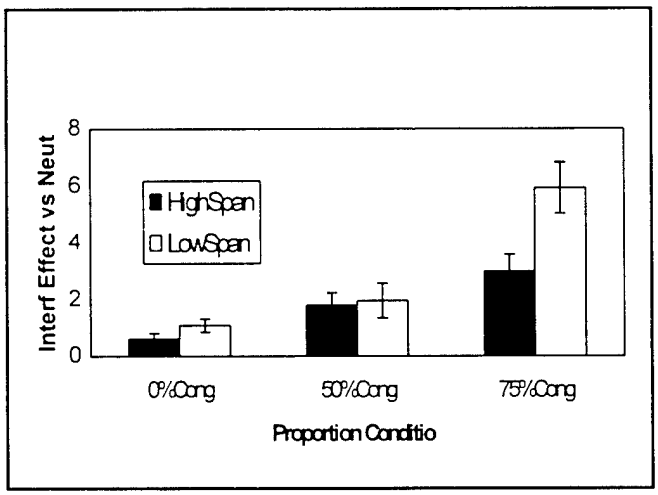
Results. Stroop interference effects for naming latencies were calculated by subtracting response times on critical neutral trials from critical incongruent trials. Interference effects for high and low spans, in each of the %-Congruent conditions, are presented in Figure 8. Interference effects grew with increasing proportion of congruent trials, but contrary to our predictions, high and low span subjects appear to show equivalent interference across all proportion conditions. Indeed, a 2 (Span) x 3 (Proportion Congruency: 0%, 50%, 75%) x 2 (Trial Type: Neutral, Incongruent) ANOVA confirmed these observations. Trial type interacted with Proportion congruency, $F(2,128) = 28.68$, $MSE = 2229.84$, $p < .001$, indicating larger interference effects with more congruent trials, but the Span x Trial-type x Proportion congruency interaction was not significant, $F(2,128) < 1$. In each of the Proportion congruency conditions considered separately, no span differences in interference emerged (Fs between 0.50 and 1.90; ps between .48 and.18).

Figure 8



However, substantial span differences emerged in color-naming accuracy. Stroop interference effects for errors were calculated by subtracting the number of errors on critical neutral trials from critical incongruent trials. Interference effects for high and low spans, in each of the %-Congruent conditions, are presented in Figure 9. A 2 (Span) x 3 (Proportion Congruency) x 2 (Trial Type) ANOVA indicated that low spans showed larger interference than did high spans, $F(1,128) = 5.96$, $MSE = 3.88$, $p < .05$, interference effects grew with proportion congruency, $F(2,128) = 19.22$, $MSE = 3.88$, $p < .001$, and the Span x Trial-type x Proportion-congruency was significant, $F(2,128) = 3.24$, $MSE = 3.24.$, $p < .05$. That is, span differences in Stroop interference increased with the proportion of congruent trials in the design. Note that span differences in interference were actually significant only in the 75%-Congruent condition, $F(1,42) = 7.02$, $MSE = 7.02$, $p < .05$, and not in either the 0% condition, $F(1,41) = 1.08$, $MSE = 1.11$, $p = .31$, or the 50% condition, $F(1,45) < 1$. Thus, in the 75%-Congruent context, which did not repeatedly reinforce the goal of the task ("*ignore the word and read the color*") low spans showed evidence of increased goal neglect relative to high spans. Low spans were more likely to name the word, in error, than were high spans when the context provided no penalty for word naming on a majority of trials.

Figure 9



These data indicate that in the absence of environmental prompting, when successful performance depends upon actively maintaining goals in a highly accessible state, high spans will do so more effectively than will low spans. However, in a 0%-Congruent context that repeatedly reinforces the goal of the task by presenting only conflict stimuli, high and low spans performed equivalently. Here the task context reduced the need to actively maintain task goals because those goals were reinforced by every stimulus that appeared.

Ravens scores correlated significantly with Stroop response time measures only in the 75%-Congruent condition. For example, the correlation between Ravens and incongruent-trial response times was -.50, and between Ravens and response time interference effects was -.49. Thus, the higher the Ravens score, the better the Stroop performance. Corresponding (non-significant) correlations for the *50%-Congruent* condition were -.02 and .03, and for the *0%-Congruent* condition were .00 and -.04.

Note that our findings suggest a view of Stroop interference effects that is different from most. If Stroop interference simply arose from an output conflict between a word response and a color response, then interference effects and individual differences therein should not be sensitive to the proportion of congruent and incongruent trials. Our data suggest that the locus of individual differences in Stroop interference lies in the capability to sustain a goal representation in the face of interference from the environment. That is, if the environment repeatedly reinforces the task goal by presenting only incongruent stimuli, high and low WM individuals are equally able to ignore the word information in responding to color. However, if the environment repeatedly interferes with the task goal by presenting congruent stimuli, low span individuals are more likely to show interference from the word on rare incongruent trials. Consistent with the Cohen et al. (1990) model, we suggest that the active sustaining of task goals may be necessary and sufficient to avoid potent distractors such as incongruent Stroop words.

STUDY 5: WM Capacity and Attention Control in the Antisaccade Task

We have replicated low-level attention differences between high and low working-memory individuals in a visual-orienting task known as an "antisaccade" task (Kane, Bleckley, Conway & Engle, in press). In antisaccade tasks, subjects respond to the location of a visual target that is either "accurately" or "inaccurately" pre-cued (see Hallett, 1978). There are typically two possible target locations, on the right and left side of a computer screen. On each "antisaccade" trial the target is preceded by a cue flashed to the screen location *opposite* of the upcoming target. The relationship between the cue and the target is consistent and predictable across trials. Thus, optimal performance requires that subjects prevent shifting their attention and their eyes to the cue. Instead, they must always attend to the opposite side of the screen from the cue to find the target. Successful performance in the antisaccade task thus requires that the strong reflexive tendency to move the eyes to the cued location be blocked, or inhibited.

We recently completed an antisaccade study to follow up the Roberts, Hager & Heron (1994) finding that a WM load disrupts performance in some aspects of the antisaccade task, but not in others. When subjects in the antisaccade task had to simultaneously update the sums of auditorially presented digits, they could no longer block their eye movements to the invalid cue. Thus, Roberts et al. found that when accurate performance required subjects to actively maintain the goal, "*look away from the cue*," in WM in the face of a strong pre-potent tendency to automatically look toward the cue, adding a memory load greatly impaired performance. Subjects were less accurate in identifying the target stimuli, and they made more overt eye movements to the cued, but incorrect, screen position. In contrast, the addition of a memory load had no effect on subjects' abilities to make "prosaccades." In prosaccade trial blocks, the

location of the upcoming target was always pre-cued accurately (i.e., the cue and target appeared in the same location), and so there was no interference with actively sustaining the goal of the task.

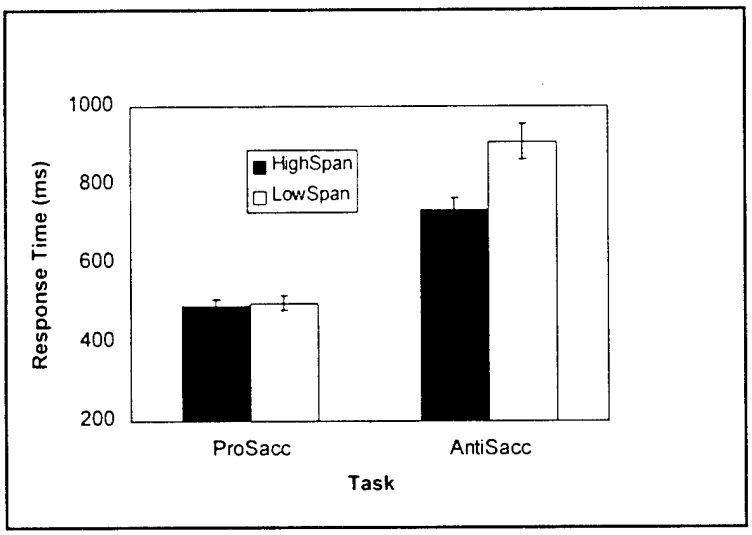
Here we expected that high and low spans would perform equivalently in a prosaccade task, but that high spans would perform better than low spans on an antisaccade task. To preview, that is precisely what we found (see Kane et al., in press).

Experiment 1 Methods. Two hundred three undergraduates from Georgia State University and Georgia Institute of Technology participated. These subjects were identified from a larger pool who had participated in the Operation-Span task: There were 107 high spans and 96 low spans.

Subjects identified the masked target stimulus on each trial by pressing a key corresponding to the target. The target on each trial was the letter *B*, *P*, or *R*. The entire experiment consisted of six trial blocks: two letter identification practice blocks, a prosaccade practice block, a prosaccade experimental block, an antisaccade practice block, and an antisaccade experimental block, with the order of the prosaccade and antisaccade blocks varying between subjects. In all blocks, the target letters *B*, *P*, and *R* occurred an equal number of times.

Each non-practice experimental block consisted of 72 trials. Each block began with the presentation of a yellow "*READY?*" signal at the center of the screen against a black background. The ready signal remained on screen until the subject pressed the keyboard's space bar, which was followed by a 400 ms blank screen. A cyan fixation signal ("***") then appeared at the center of the screen for an interval that varied unpredictably, as is typically done in antisaccade tasks (see Hallett & Adams, 1980; Roberts et al., 1994), here between 200 – 2200 ms. An equal number of trials had fixation durations of 200, 600, 1000, 1400, 1800, or 2200 ms. Immediately after the cyan fixation signal disappeared, a 50 ms blank screen was followed by a "=" cue that appeared for 100 ms to the right or left of fixation (with an eccentricity

Figure 10



of 11.33° of visual angle), one character space below the horizontal plane of the fixation signal.

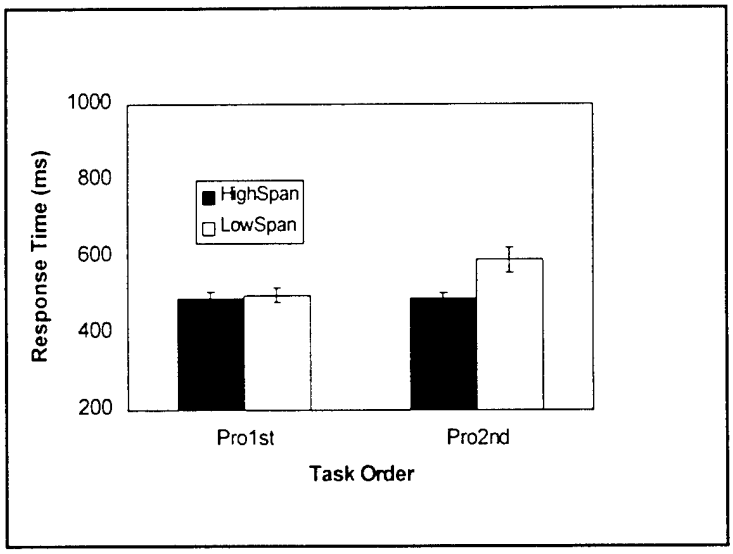
Then, a second 50 ms blank screen was followed by the second appearance of the cue, which appeared for 100 ms in the same eccentric location. Thus the cue appeared to briefly flash on and off, and so was a strong attractor of attention. In *prosaccade* trial blocks, the target appeared for 100 ms. in the character space directly above the one that had been occupied by the cue. Here, subjects could reflexively orient to the cue to locate the target.

In *antisaccade* trial blocks, the "=" cue always appeared on the *opposite* side of the screen from the upcoming target stimulus. So if the cue appeared on the left of the screen, the target then appeared on the right of the screen, and vice versa. For both prosaccade and antisaccade trials, the target was followed by a succession of backward-masking stimuli: an "H" for 50 ms, and then an "8" that remained until a response key was pressed. A 500 ms tone gave feedback immediately following after an incorrect response. The next trial began with a 400 ms blank screen.

Experiment 1 Results. We expected that high and low spans would differ minimally (if at all) in the prosaccade task, where fast and accurate target identification would be aided by a relatively automatic orienting response. In contrast, we expected high spans to significantly outperform low spans in the antisaccade task, where fast and accurate identification required the active blocking of, or recovery from, an automatic orienting response.

Significant task order effects lead us to examine span differences in prosaccade versus antisaccade performance by analyzing response latencies from subjects' first task block only, thus treating task as a between-subjects variable. For the prosaccade task, then, data were analyzed from 52 high spans and 45 low spans, and for the antisaccade task, data were analyzed from a different set of 55 high spans and 51 low spans. These data are presented in Figure 10.

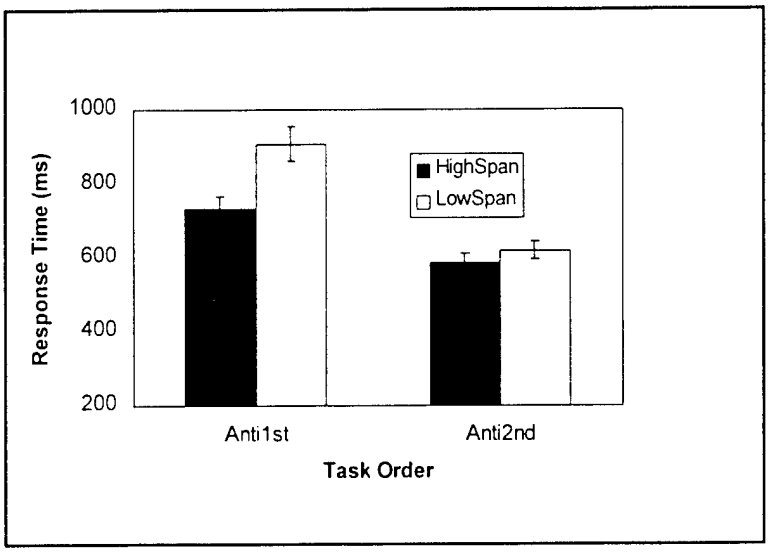
Figure 11



Responses in the prosaccade task were significantly faster than in the antisaccade task. Most importantly, high and low spans performed virtually identically in the prosaccade task (M difference = 8 ms) and quite differently in the antisaccade task, with high spans identifying targets significantly faster than low spans (M difference = 174 ms).

The effects of task order on target-identification speeds in prosaccade and antisaccade tasks are depicted in Figures 11 and 12, respectively. For subjects who experienced the prosaccade task first, there were no span differences in prosaccade performance, as discussed above. However, for subjects who experienced the prosaccade task second – after completing the antisaccade task – span differences emerged. Here, target-identification latencies for low spans who completed the prosaccade task after the antisaccade task were significantly longer than those of their high span counterparts. Moreover, low spans who completed the prosaccade task second were significantly slower to identify targets than were those who completed it first.

Figure 12



Task order also affected antisaccade performance, but it did so in the opposite direction (see Figure 12). That is, for subjects who experienced the antisaccade task first, large span differences were evident, as discussed above. However, for subjects who experienced the antisaccade task second – after completing the prosaccade task – span differences were absent. Moreover, low spans' antisaccade performance benefited significantly more from practice on the antisaccade task than did high spans' performance.

Why should low spans have responded more slowly on prosaccade trials following the antisaccade task? A possibility is that once low spans had repeatedly attempted the controlled task of looking away from the cue, they had more difficulty than high spans in abandoning that task set in favor of the more automatic task set allowed by prosaccade trials. Low spans may have perseverated more than high spans on the antisaccade requirement of trying to look away from the cue when the task changed to allow looking toward the cue. Of course, this speculative interpretation is consistent with our view that low spans are less able to control attention than are high spans. We replicate these findings in Experiment 2.

An interesting contrast to the prosaccade order effects was seen in the antisaccade task. Both high and low span subjects were *faster* when antisaccade was their second task of the experiment (i.e., when it followed the prosaccade task), than when it was their first task. Moreover, this "task-two" benefit was actually larger for low spans than for high spans, and when the antisaccade task was presented second it eliminated span differences in target identification times. Because this finding suggests that span differences in antisaccade

performance may be eliminated with minimal practice, Experiment 2 further explored the effect of practice on the antisaccade task.

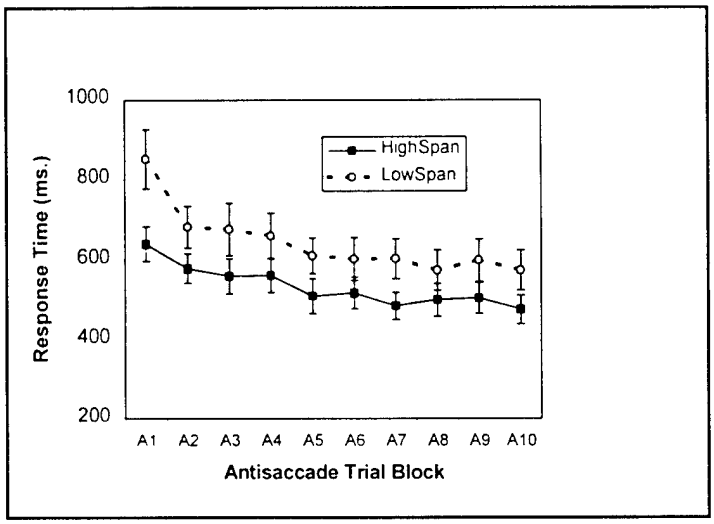
In Experiment 2, high and low WM subjects performed the target identification task from Experiment 1 while we monitored eye movements. We also presented 10 separate blocks of 36 antisaccade trials in order to examine practice effects on span differences in suppressing eye movements. The final trial block was a prosaccade block that tested the effects of extended antisaccade practice on prosaccade performance.

Experiment 2 Methods. The methods for Experiment 2 were identical to those of Experiment 1 with the following exceptions. We tested 40 undergraduates (20 high spans, 20 low spans) from Georgia State University and Georgia Institute of Technology. Due to problems with the eye-movement data collection system, data from 7 subjects were discarded, leaving 16 high spans and 17 low spans in the analyses.

Eye-movement data were collected using an Applied Science Laboratories (ASL) E-5000 eye tracker and pupilometer, an infra-red based, corneal reflectance system. A Flock-of-Birds® magnetic head tracking system was used to coordinate head movements and camera focus on the eye. The apparatus allowed for the detection of left-eye movements greater than 0.5°. Subjects first put on the MHT headband, and point of gaze was calibrated. Subjects began with one target-identification practice block of 36 trials. Subjects then practiced the antisaccade task for only 6 trials, and then began the 10 experimental blocks of 36 antisaccade trials each. Following the antisaccade trial blocks, subjects practiced the prosaccade task for 6 trials, and then began the 1 experimental block of 36 prosaccade trials. At the beginning of each experimental block, gaze was checked for proper calibration and recalibrated as necessary.

Experiment 2 Results. Figure 13 presents high and low span subjects' mean target-identification latencies for blocks 1 – 10 of the antisaccade task. High spans responded significantly faster than did low spans across all antisaccade blocks (Ms = 533 and 641 ms, respectively). Even though span differences appeared larger in block 1 than in subsequent blocks, the Span x Block interaction did not approach significance. Thus, in contrast to Experiment 1, where significant prosaccade practice eliminated span differences in subsequent antisaccade performance, here span differences persisted across several hundred trials of antisaccade practice. On the final block 11, which presented *prosaccade* trials, high spans identified targets significantly faster than did low spans (Ms = 460 and 551 ms, respectively).

Figure 13



Thus, as in Experiment 1, significant antisaccade practice was followed by substantial span differences in prosaccade task performance, with low spans taking much longer to identify a targets even when their eyes should have been reflexively drawn to its subsequent location.

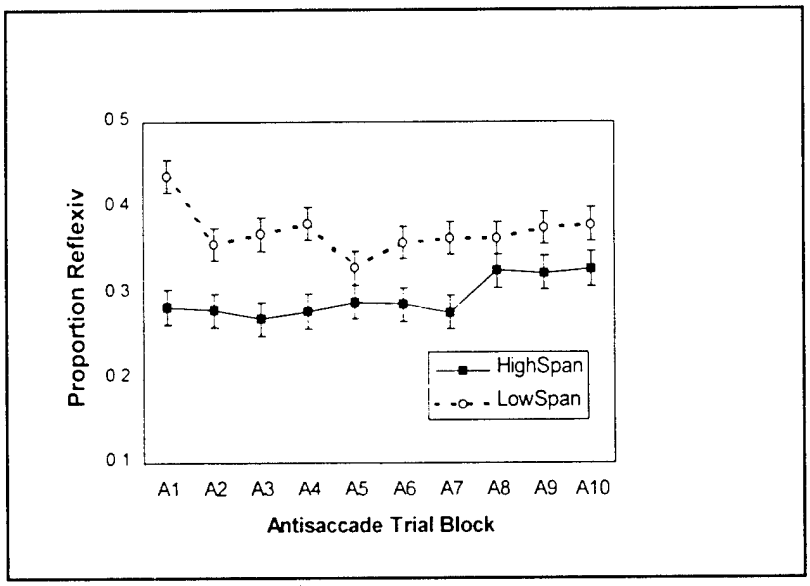
For each trial the saccade accuracy and latency were calculated from the eye movement data. Figure 14 displays the proportions of high and low spans' initial saccades on antisaccade trials that were reflexively drawn to the cue, in opposition to task instructions.

Low spans were significantly more likely than were high spans to initially move their eyes toward the abrupt-onset cue, which reliably appeared in a location that would *not* contain the target (M error rates = .371 and .280, respectively). Indeed, as in the target-identification data, the significant span difference persisted across practice on hundreds of antisaccade trials.

Moreover, on those trials on which a reflexive saccade occurred, significant span differences emerged in the time taken to recover. Low spans maintained fixations on the incorrect side of the screen over 150 ms longer than did high spans (overall Ms = 674 and 512 ms, respectively). Thus, compared to high spans, low spans not only made more saccade errors, but, after committing an error they also took much longer to correct it.

In block 11, the prosaccade block, low spans also made significantly more saccade errors than did high spans. Here, however, saccade errors reflect looking *away* from the cue instead of reflexively attending to the cue. Thus, the higher saccade error rate for low spans (M = .281) compared to high spans (M = .202), indicates that low spans were significantly more likely to look away from a "valid," prosaccade cue.

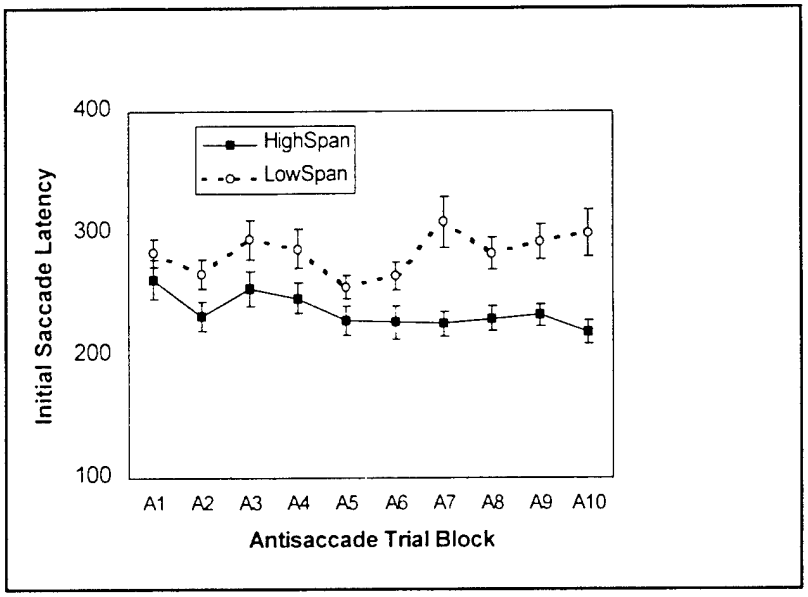
Figure 14



Low spans appeared to have more difficulty than did high spans in abandoning the task set from the previous antisaccade blocks and shifting set to the prosaccade task requirements, a difficulty that was also reflected in the target-identification data from this experiment and from Experiment 1.

Figure 15 presents mean latency for initiating saccades across antisaccade trial blocks, collapsed across correct "controlled" eye movements (i.e., toward the direction opposite the cue) and incorrect "reflexive" eye movements (i.e., toward the same direction as the cue). Overall, low spans initiated their eye movements significantly more slowly following the cue than did high spans across the entire session (Ms = 284 and 236 ms, respectively). On prosaccade trials (Block 11), low spans' saccades were initiated significantly more slowly (M = 286 ms) than were high spans' saccades (M = 203 ms). Thus, not only did low spans tend to make eye movement errors on prosaccade trials following extended antisaccade practice, but also the saccade latencies were quite long. These long latencies might suggest that low spans were making "controlled" saccades on many prosaccade trials (in fact, low spans' mean saccade latency in the prosaccade task, at 286 ms, was nearly identical to that for the antisaccade task, at 284 ms). Following antisaccade practice, then, low spans appeared to persist, more than high spans, in making controlled eye movements when no longer required.

Figure 15



Our findings from Experiment 2 replicate and extend the key findings from Experiment 1. First, high and low WM span subjects differed significantly in ability to identify visual targets that were signaled by antisaccade cues. That is, on trials in which a flashing cue predictably appeared in the opposite location as the upcoming target, low spans were slower and more error prone in identifying targets. Moreover, Experiment 2 demonstrated that this substantial span difference maintained across a total of 360 trials, with little sign of diminution over practice. These findings suggest that low spans are less able to block reflexive eye movements to abrupt-onset cues that conflict with task goals, and that low spans' difficulties are not limited to novel situations that involve minimal practice.

In addition, with respect to the target-identification task, Experiment 2 replicated the unexpected finding from Experiment 1 that prosaccade performance for low spans was particularly disturbed by prior practice on antisaccade trials. Compared to high spans, low spans were significantly slower (by more than 150 ms) and less accurate in their responding on the block of prosaccade trials, which followed after 10 blocks of antisaccade practice. Low spans may be less able to shift intentional set between tasks than are high spans.

The eye-movement data collected in Experiment 2 nicely reinforce the target-identification findings from both experiments. Specifically, on antisaccade trials, low spans were considerably more likely to make reflexive saccades towards the cue than were high spans. This difference in the ability to suppress saccades, although especially large in the first trial block, persisted over substantial practice. Moreover, once an error was committed, low spans took much longer than high spans to recover and move their eyes to the correct side of the screen. The same was true for initial saccade latency: Antisaccades were initiated more slowly by low spans than by high spans over the entire session.

Moreover, with respect to the prosaccade task, the eye-movement data suggested that low spans' difficulties following antisaccade practice are at least in part due to a perseveration on the antisaccade task goal. *Low spans were more likely than high spans to look away from the prosaccade cue, and they were slower to initiate saccades in this condition*. Even though the cue consistently appeared in the same location as the target, low spans appear to have been less able than high spans to reconfigure their task set to allow less controlled, more automatic responding.

The data from Experiment 2 also constrain further hypotheses regarding the other task-order effect from Experiment 1, namely that prosaccade practice eliminated span differences in antisaccade performance. Clearly, the findings from Experiment 2 discount the possibility that simply *any* kind of visual-orienting task practice will eliminate span differences in the antisaccade task, as span differences in target identification, saccade accuracy, and saccade latency remained significant across 10 blocks of antisaccade practice. Either the Experiment 1 effect was spurious, or something specific about prosaccade practice led low spans to improve in the antisaccade task. Further experiments will be required to determine which of these is correct.

Thus, in two experiments in which subjects with high and low WM capacity were tested on an analogue of the antisaccade task, high spans demonstrated better control over visual orienting. In antisaccade trial blocks, where eyes and attention were to be moved away from an abrupt-onset visual cue, optimal performance required that reflexive orienting responses be suppressed. Here, in accord with task demands, high spans were less likely than low spans to move their eyes toward the flashing cue (Experiment 2), and high spans were faster to correct their saccade errors (Experiment 2). High spans were also faster and more accurate in identifying visual targets that appeared in the opposite location as the cue (Experiments 1 and 2). In contrast, in prosaccade trial blocks, where subjects' reflexive responses did not conflict with task goals, high and low spans performed similarly, when the prosaccade task was performed *first*.

Thus, of primary interest here is that high span individuals outperformed low span individuals in a task demanding significant attention control but not a significant memory load. In a task requiring no complex mathematical processing or retention of random word lists (as in the OSPAN task), substantial differences were seen between individuals of high and low WM capacity. At least, span differences were seen in the antisaccade task, a task that not only required subjects to orient their eyes to a discrete location on cue, but also required them to actively maintain the task goals in the face of powerful interference from the environment. When such controlled processing was unnecessary for successful performance, i.e., on prosaccade trials, high and low spans performed equivalently. Note, however, that prosaccade performance did demand more than simple reflexes, as heavily masked stimuli were to be rapidly identified via a choice-RT task. WM capacity thus appears to be related to the controlled processing required in responding to interference. WM capacity, as measured by OSPAN and

other complex WM tasks, predicts performance even on very simple, low-level tasks that require little in the way of "complex" higher-order processing, as long as successful performance depends upon active maintenance in interference-rich conditions.

The task-order effects observed here on prosaccade task are intriguing. Performance on prosaccade tasks, unlike antisaccade tasks, is typically unaffected by the imposition of a memory load (Roberts et al., 1994), by advancing age (Butler et al., 1999), or by injury to prefrontal cortex (e.g., Fukushima et al., 1994). And here, in Experiment 1, we found that for unpracticed subjects, prosaccade-task performance was not related to WM capacity, either. Together these findings indicate that the prosaccade task may be performed with little involvement of controlled processing. However, our findings also demonstrate that this "automatic" task may be disrupted by the prior performance of a similar, but attention-demanding, task. Particularly for low spans, switching instructional set from the antisaccade task to the prosaccade task appeared to be quite difficult. Following practice on the antisaccade task, low spans made more antisaccade-type eye movements than did high spans on the prosaccade task (Experiment 2), and low spans were slower and less accurate than high spans in the prosaccade target-identification task (Experiments 1 and 2). However, even high spans showed some evidence of perseveration, with an increase in identification errors in Experiment 1, and with a non-negligible number of "anti" saccades in the final, prosaccade task block in Experiment 2.

Our results resemble those from Allport, Styles and Hsieh (1994; see also Harvey, 1984), who examined task switching in a series of experiments using various Stroop-like tasks, including the traditional color-word task (Stroop, 1935). In their Experiment 5, they found substantial set-shifting costs when naming the color of a color-word on one trial (high interference) was followed by reading the word of a color-word on the next trial (low interference). Thus, shifting set from a controlled task to an automatic task was markedly difficult, even though the tasks alternated predictably and occurred as much as 1100 ms apart. The converse effect was <u>not</u> found, however, in that shifting set from reading words to naming colors produced no cost whatsoever. In a similar vein, but outside the Stroop-task context, Meuter and Allport (1999) recently demonstrated switching asymmetries in bilingual subjects who switched between their dominant and non-dominant language in naming digits: Switch costs were larger from the non-dominant language into the dominant language than vice versa. Much like our data from Experiment 1, then, these findings demonstrate that switching from a more automatic task to a more controlled task causes minimal difficulty compared to switching from a controlled to an automatic task. Allport and colleagues discuss their findings in terms of *task-set inertia*, a kind of PI in which a non-dominant response mapping imposes a stronger set that is more difficult to overcome than is the set for a dominant response. Given our prior findings of WM-span differences in PI (Kane & Engle, in press), we recommend further exploration of the relations among WM, controlled attention, and task switching.

<u>STUDY 6: WM Capacity, Controlled Attention, and Task Switching</u>

A small (but growing) body of research indicates that people are slowed and less accurate to respond to stimuli if the goal of the task changes from trial to trial instead of remaining constant across (e.g., Allport et al., 1994; Meiran, 1996; Rogers & Monsell, 1995). This effect may be termed a "switch cost", and it presumably reflects a failure of the attentional control system to: (1) abandon a previous task set, and (2) prepare a new task set. The relative contribution of each of these two processes is a current point of controversy. However, we hypothesized that insofar as controlled attention is important for successful task switching, high and low spans

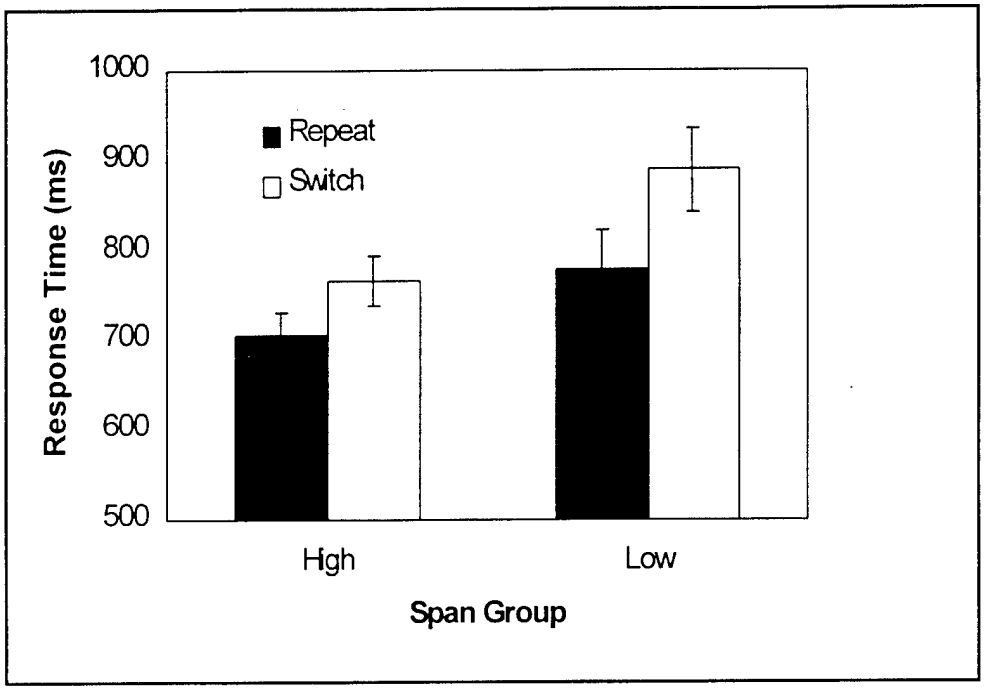
should differ in relatively simple switching tasks. We have completed two of a series of three task-switching experiments.

Experiment 1 Method. Twenty high spans and 19 low spans responded to paired stimulus trials (a prime trial followed by a probe trial). Half the paired trials required the same task demands to be repeated across prime and probe stimuli, and half required that a switch between tasks be made across stimuli. The two basic tasks subjects faced were: (1) deciding whether a stimulus word represented a living thing or not; or (2) deciding whether a stimulus word was presented above fixation or not. The same stimuli are used across the two tasks: One of the words, *horse, eagle, roach, shirt, paint*, or *table*, was presented either just above central fixation or just below fixation. Thus, a given stimulus was ambiguous with respect to task demands – the stimulus itself did not trigger the kind of response to be made. Any word could either be considered a living/nonliving thing, or be considered above/below fixation. Subjects' only cue for which task demands to use in responding to the stimulus was a brief signal that appeared just before the stimulus. The cue was either "*ALIVE?*", signaling the subject to respond to the animacy of the stimulus, or "*ABOVE?*" signaling the subject to respond to the location of the stimulus. Subjects pressed the 'yes' key to respond affirmatively to either task, and the 'no' key to respond negatively to either task. Thus, the stimuli and response mappings were identical for switch and non-switch trial pairs.

The sequence of events was as follows: A yellow "*READY*" signal remained centered onscreen until the subject initiated the trial pair by pressing a key. After a 700 ms pause, the first cue (*ALIVE?* or *ABOVE?*) appeared, centered onscreen, for 200 ms. The prime stimulus then immediately appeared and remained onscreen until response. The prime word was either a living thing (*horse, eagle, roach*) or a nonliving thing (*shirt, paint, table*) and it appeared one character space either above or below the previous task cue. After a blank pause of 300 ms, the second task cue (*ALIVE?* or *ABOVE?*) appeared for 200 ms, followed immediately by the probe word, which remained onscreen until response. The next trial pair then began when the subject pressed the space bar in response to the *READY* signal. On half the trial pairs, the task repeated across prime and probe trials (ABOVE → ABOVE; ALIVE → ALIVE), and on half the task switched (ABOVE → ALIVE; ALIVE → ABOVE). No stimulus word ever repeated across a prime-probe trial.

Experiment 1 Results. RT data from task-repeat and task-switch trials for high and low spans are displayed in Figure 16, collapsed across ALIVE and ABOVE tasks. Overall, high spans were slightly faster than low spans, $F(1,36) = 3.94$, $MSE = 46107.71$, $p = .06$, and task-switch trials were much slower than task-repeat trials, $F(1,36) = 123.55$, $MSE = 1118.61$, $p < .001$. Thus we replicated the basic task-switch cost effect demonstrated by others. Most importantly, the switching cost was significantly larger for low spans than for high spans (Ms = +110 vs. + 61 ms, respectively), $F(1,36) = 10.34$, $MSE = 1118.61$, $p < .01$. We also calculated switch costs as a proportion of task-repeat RTs for each subject in order to account for low spans' slower RTs in task-repeat trials. Even this more conservative measure indicated that low spans showed a significantly larger switch cost than did high spans (Ms = +14.1% vs. +8.7%, respectively), $F(1,36) = 7.61$, $MSE = 0.004$, $p < .01$.

Figure 16



Span differences in switching were consistent across ALIVE and ABOVE tasks individually. For the ALIVE task, a clear overall switch cost was evident between task-repeat trials (ALIVE → ALIVE) and task-switch trials (ABOVE → ALIVE) for both high and low spans, $F(1,36) = 54.49$, $MSE = 1173.01$, $p < .001$. However, the switch cost was significantly larger for low spans than for high spans, both in raw RTs (Ms = +78 vs. +39 ms, respectively), $F(1,36) = 6.20$, $MSE = 1173.01$, $p < .05$, and in proportional costs (Ms = +10.0% vs. + 5.4%, respectively), $F(1,36) = 5.11$, $MSE = 0.04$, $p < .05$.

For the ABOVE task, there was a significant switch cost between task-repeat trials (ABOVE → ABOVE) and task-switch trials (ALIVE → ABOVE), $F(1,36) = 113.44$, $MSE = 2129.06$, $p < .001$. And, the switch cost was significantly larger for low spans than for high spans in both raw RTs (Ms = +143 vs. +83 ms, respectively), $F(1,36) = 4.98$, $MSE = 2129.06$, $p < .01$, and in proportional costs (Ms = +19.3% vs. +12.7%, respectively), $F(1,36) = 4.98$, $MSE = 0.008$, $p < .05$.

In this experiment, high spans were better able to switch attention between two task sets than are low spans. In a second experiment, we sought to replicate this finding using a task that had been used in other investigations of set switching, namely, a numerical Stroop task. In the numerical Stroop task, subjects see a group of digits on each trial, and the number of digits may conflict with the digits' identity (e.g., 222). Subjects can thus be made to switch between responding to the size of the digit group (e.g., counting "3") or to the identity of the repeated digit (e.g., identifying "2"). Allport et al. (1994) found substantial switch costs in this task, but the costs were asymmetrical. That is, as discussed with reference to our antisaccade study, subjects had difficulty switching from the more controlled set of counting digits to the more automatic set of identifying digits. However, no cost was seen in switching from digit identification to digit counting. We found similar trends in our antisaccade task, where low spans were particularly impaired when switching from the more controlled antisaccade task into

the more automatic prosaccade task, but neither group had difficulty switching from prosaccade into antisaccade. We expected to find span differences in switch costs here, but perhaps only when switching from digit counting into digit identification.

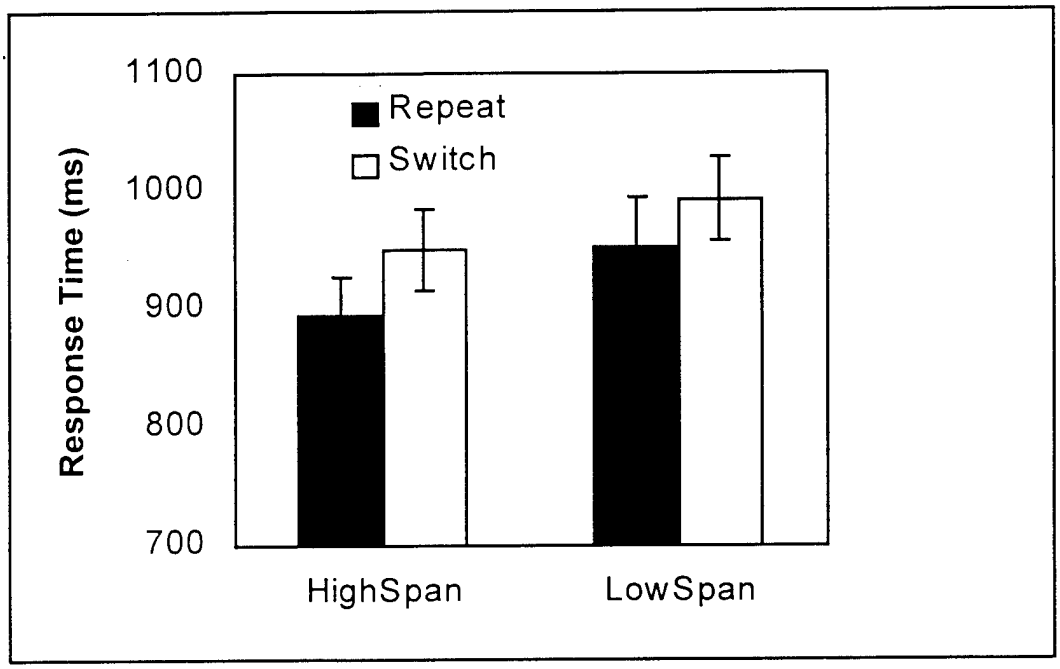
Experiment 2 Methods. We tested twenty-four high spans and 30 low spans (as indicated by their Operation-Span performance) in a Stroop-like switching task that closely resembled the task structure of Experiment 1. Here, however, the target stimuli were strings of 2 – 5 identical digits, presented in one horizontal row at fixation, with the digits 2, 3, 4, and 5 used exclusively. On each display, the number of digits always conflicted with the digit's identity. The subject pressed a key to indicate either the digit identity (on half the trials), or the total number of digits that appear in the string (on half the trials). As in Experiment 1, the task for each display is indicated by its preceding cue, the word *DIGIT* or *COUNT*, and half the trials were task-repeat and half are switch.

Each trial began with a red ready signal ("*READY?*"), which remained centered onscreen against a until the subject initiated the trial by hitting a key. A yellow "+" sign then appeared for 2500 ms, followed immediately by a green task signal for 250 ms ("*? DIGIT ?*" or "*? COUNT ?*"). The prime stimulus (green row of digits) then appeared immediately, one character space below the task signal. Both the task signal and the prime stimulus stayed onscreen until the subject responded with a key-press (the keys *v, b, m,* and *n* were labeled *2, 3, 4,* and *5*, respectively). Following a blank screen of 150 ms, a green task signal for the probe appeared for 250 ms ("*? DIGIT ?*" or "*? COUNT ?*"). The probe stimulus (green row of digits) then appeared immediately, one character space below the task signal. Both the task signal and the prime stimulus stayed onscreen until the S responds. Following a blank delay of 500 ms, the next red ready signal appeared. There were 24 trials each for *DIGIT* ➔ *DIGIT*, *COUNT* ➔ *COUNT*, *DIGIT* ➔ *COUNT*, and *COUNT* ➔ *DIGIT* sequences.

Experiment 2 Results. Probe response time data are presented in Figure 17. High and low spans responded equivalently fast overall, $F(1,52) = 1.01$, $MSE = 66105.63$, $p = .32$, task-switch trials were slower than task-repeat trials, $F(1,52) = 18.91$, $MSE = 3280.06$, $p < .001$, but high and low spans did not differ in switch costs (Ms = + 56 ms and + 41 ms, respectively): The Span x Switch-cost interaction was not significant, $F(1,52) < 1$, $p = .52$. Span differences were also not significant when switch costs were measured proportionally, as a function of subjects' speed on task-repeat trials, $F(1,52) < 1$, $p = .61$ (Ms = + 5.8% and + 4.6% for high and low spans, respectively).

When considering the digit-identification probe trials separately, there were no span differences in overall response speed, $F(1,52) = 1.72$, $MSE = 74098.14$, $p = .20$, but task-switch trials were significantly slower than task-repeat trials, $F(1,52) = 53.38$, $MSE = 7355.46$, $p < .001$. High and low spans showed equivalent switch costs in raw scores (Ms = + 126 ms and + 118 ms), $F(1,52) < 1$, $p = .817$, and in proportional costs (Ms = + 16% and + 14% for high and low spans, respectively), $F(1,52) < 1$, $p = .66$. Thus, when subjects had to switch from a relatively controlled task into a relatively automatic task, switch costs were robust, but in the absence of span differences.

Figure 17



When considering the digit-counting probes separately, there were again no span differences in overall response speed, $F(1,52) < 1$, $p = .55$. However, in contrast to the digit-identification task, switch costs here were negative; that is, subjects were actually slightly faster on task-switch trials than on task-repeat trials, $F(1,52) = 3.04$, $MSE = 5542.43$, $p = .09$. No span differences emerged in this task-switching "effect," either for raw scores (Ms = - 14 ms and - 35 ms for high and low spans, respectively), $F(1,52) < 1$, $p = .47$, or in proportional scores (Ms = - 1.6% and - 3.0% for high and low spans, respectively), $F(1,52) < 1$, $p = .63$. Thus, here, when subjects had to switch from a relatively automatic task into a relatively controlled task, switch costs were absent for both high and low span subjects.

On one hand, we replicated Allport et al. (1994) and Kane et al. (in press) in that switch costs were asymmetrical: switching from automatic tasks into controlled tasks proceeded easily, without cost, but switching from controlled tasks into automatic tasks incurred substantial cost. On the other hand, we did not replicate Experiment 1, in that low spans showed equivalent switch costs to high spans.

We were puzzled by the lack of replication across these two experiments. However, there may have been an important methodological difference between the Experiment 1 and the Experiment 2 tasks, beyond the stimuli and responses. Namely, in Experiment 1, half of all the switch trials allowed subjects to respond correctly *even if they failed to shift set*. That is, the probe display on these trials allowed the same response according to both stimulus dimensions. For example, in the ALIVE/ABOVE task from Experiment 1, the probe display represented a living thing AND it was presented above center. If the subject had made an ALIVE decision on the prime, and an ABOVE decision on the probe, a correct response ("yes"") to the probe could have been made whether or not the subject actually made an ABOVE judgment. On such trials the subject did need not switch attentional set in order to accurately respond with a "yes."

In contrast to Experiment1, the Experiment 2 task *never* allowed a correct response based on the previous task set, because the digit identity and the number of digits *always* conflicted. Thus, low spans may have shown larger costs than high spans in Experiment 1 because they could fail to switch set half the time and still perform well, thus setting them up to perform much more poorly when a switch was mandatory. In contrast, the task in Experiment 2 demanded that switching actually occur on every "switching" trial in order to respond correctly to the probe. Here the environment acted like a central-executive in some sense, in that it repeatedly reminded the subject of the task goal. Note the similarity here to our Stroop findings, reported above. In the Stroop task, span groups performed equivalently when the context included all incongruent trials, and so subjects were repeatedly reminded of the task goal ("*ignore the word*"). In contrast, low spans made substantially more errors on incongruent trials than did high spans when the context included 75% congruent trials. Here the context often allowed for "correct" responding even when the task goal was temporarily neglected, and so the task was sensitive to span differences in active maintenance of goal information. We expect that a similar pattern may hold for switching tasks.

Thus, we are currently conducting a third experiment using the Experiment 2 task. Here, however, we are testing half the high- and low-span subjects with a modified version in which half the trials allow correct responses without actually switching set. We expect that we will replicate Experiment 2 with the CUE group, but that we will find robust span differences in the modified CUE group that replicates Experiment 1.

STUDY 7: WM Capacity, Distraction, and Delayed Memory for Tones

The purpose of this study was to explore the hypothesis that working memory plays a role in the ability to maintain information over time, especially in the face of irrelevant and distracting information. Such a hypothesis has already received support from our Stroop and antisaccade experiments, with respect to the maintenance of task goals, or productions. The present experiment was planned and designed during the grant reported on here but was conducted under the current grant. It tested whether WM capacity predicts the ability to maintain the representations of external stimuli in the face of external distraction.

Experiment 1 Method. We used a delay matching-to-sample task in which subjects were instructed to remember the pitch of a target tone over a 5 s interval. Two types of trials were presented, distraction and non-distraction. Distraction trials included five additional tones that were presented during the delay. The experiment was divided into six blocks, alternating between trial types. In order to qualify for the study, subjects had to perform at 80% accuracy on non-distraction trials with a very short inter-stimulus interval (500 ms). The Operation-word span task was given to assess working memory span.

The auditory stimuli for the delay matching-to sample-task were tones of varying frequency. We will refer to the sample and target tones, respectively, as T1 and T2. Three different tone frequencies were used for T1 (810, 820, and 830 Hz). Frequencies for T2 were either 0,10, or 20 Hz above or below T1. Distractor tones were within 790 and 850 Hz, and were never the same frequency as the sample or target. All tones (samples, targets, and distractors) were presented 250 ms. The inter-stimulus intervals between the distractor tones were randomly selected, and were never less than 250 ms. The overall delay between the sample and target tones was always 5 s, for both distraction and non-distraction trials.

Experiment 1 Results. WM capacity, as measured by Operation span task, was significantly predictive of performance on the delayed-matching-to-sample task, for both no-distractor and

distractor trials. The slopes of the regression lines for accuracy as a function of WM span score are nearly identical for distraction and no distraction trials (although performance was significantly lower for trials that include distraction). However, WM span was *not* predictive of the ability to match the tones with only a 500 ms delay. This latter finding may be interpreted as evidence that WM capacity is not related to auditory discrimination abilities. All together, what we can conclude from these findings is that WM span predicts the ability to maintain a representation for pitch over a period of five seconds, regardless of whether that period is filled with external distractors.

Experiment 2 Method. The goal of Experiment 2 was to look at performance accuracy over varying delays between sample and target. Since Experiment 1 showed that WM span predicted the ability to maintain representation for the tone over 5 s, perhaps there is a lower bound for this relationship. With a shorter memory delay, span may no longer predict memory performance.

The procedure for Experiment 2 was similar to the non-distraction trials in Experiment 1. Six inter-stimulus intervals (750, 1000, 2000, 3000, 4000 ms) were randomly presented. There were 48 trials for each interval. The pairs of tones used were exactly the same as in Experiment 1.

Experiment 2 Results. This experiment is still underway. Findings thus far (with 50 individuals) show that span does not predict memory performance when a 1000 ms delay is employed. However, when the delay is greater than or equal to 2000 ms, span does significantly predict performance. WM span is especially predictive when looking at accuracy for the trials in which there was only a 10 Hz frequency difference between the target and sample tones. Since high and low span individuals are equivalent in maintaining an active representation for tone frequency over 1000 ms, adding external distractors at this interval may hurt low spans' accuracy more than high spans' (in contrast to our Experiment 1 findings with a 5000 ms delay). We will be testing this hypothesis in a follow-up study.

STUDY 8: Domain Knowledge WM Capacity in Complex Cognitive Performance

Previous research suggests that there are at least two cognitive ability characteristics that contribute to success in the performance of a wide range of real-world cognitive tasks. The first characteristic is the construct referred to as working memory capacity. Working memory capacity is often defined as the ability to simultaneously store and process information, and it is typically measured with dual-task paradigms, such as the Turner and Engle (1989) operation span task. The goal of this task is to solve a series of math problems (the processing task) and to remember a word that follows each for later recall (the storage task). There is now a substantial amount of evidence to suggest that measures of working memory capacity predict performance in a wide range of complex cognitive tasks, including comprehension, arithmetic, reading graphs, solving logic problems, and following directions. The second ability characteristic that predicts cognitive performance is knowledge applicable to a particular task, or domain knowledge. Not surprisingly, the facilitative effect of domain knowledge on performance of cognitive tasks is also well documented.

The goal of the current research was to investigate the interplay between working memory capacity and domain knowledge in the performance of a complex cognitive task involving language comprehension, and the specific question of interest was whether domain knowledge would reduce or even eliminate the effect of working memory capacity on language comprehension. This question is motivated by a viewpoint referred to as the knowledge-is-

power hypothesis. The premise of the knowledge-is-power hypothesis is that acquisition of domain knowledge and expertise diminishes the importance of general ability characteristics (e.g., working memory capacity) in the performance of tasks representative of the domain. For example, Frensch and Sternberg (1989) stated, "It appears that beginners in any game seem to be relying on domain-general abilities, whereas experienced players utilize an extensive body of domain-relevant knowledge. One might expect, therefore, that measures of general intelligence...would be related to novices' but not to experts' game playing ability." However, it should be noted that there is very little evidence directly relevant to this hypothesis.

This study was planned and designed under the grant reported on here but was conducted under the current grant. The participants in the current study were 182 adults from wide ranges of WM capacity and knowledge about the game of baseball, and the task involved listening to and answering questions about simulated radio broadcasts of baseball games. Baseball was chosen as the knowledge domain because of the availability of research participants from wide ranges of knowledge about baseball in the general population, and comprehension was assessed at the conclusion of each broadcast by asking participants questions about what happened in the game. The knowledge-is-power hypothesis predicts that knowledge about the game of baseball should reduce the effect of baseball knowledge on comprehension of the simulated broadcasts. However, the results of the current study do not support this conclusion: There were significant effects of both WM capacity and baseball knowledge on a variable reflecting comprehension of the simulated broadcasts. That is, participants with high levels of WM capacity exhibited better comprehension than did those with lower levels of WM capacity, and participants with high levels of baseball knowledge exhibited better comprehension than did those with lower levels of baseball knowledge. However, as can be seen, there was no indication that baseball knowledge reduced the effect of WM capacity on comprehension. In other words, there was an effect of WM capacity on comprehension *even at high levels of baseball knowledge*. These results suggest that both general ability characteristics, such WM capacity, and domain knowledge contribute to success in the performance of complex cognitive tasks. Future research will investigate the possibility that task and situational factors (e.g., task difficulty) moderate the interplay between WM capacity and domain knowledge in the performance of complex tasks.

WM Capacity, Attention Control, Intelligence, and Prefrontal Cortex Function

The AFOSR project also allowed us to develop our theoretical framework by extensively reviewing related literatures in the areas of intelligence and cognitive neuroscience. We published an extended treatment of our theoretical framework in one book chapter (Engle, Kane & Tuholski, 1999), and in two brief commentaries (Conway, Kane & Engle, 1999; Kane, Conway & Engle, 1999). In addition, we have written several revisions of a theoretical paper spanning all of these literatures. Unfortunately, we were unable to publish it in either Psychological Review and Psychological Bulletin. Part of the difficulty was the sheer size of the manuscript, and so at the end of the granting period we divided the paper into two separate manuscripts, one dealing primarily with the behavioral literature regarding working memory, attention control, and fluid intelligence, and one dealing primarily with the cognitive neuroscience literature regarding these areas. The first manuscript was submitted to Psychological Bulletin, and the second to Psychonomic Bulletin and Review.

In the first paper (Kane & Engle, 2000b), we propose an "executive attention" framework for studying the relations among working-memory capacity, general fluid intelligence (*g*F), and the prefrontal cortex (PFC). Through a review of working memory, executive attention, and *g*F literatures, we argue here that executive attention and working-memory capacity are equivalent,

and individual differences in this construct correspond to *g*F. Tasks reflecting executive attention critically require that a memory representation be maintained in a highly active state, due to the presence of interference. Moreover, this active maintenance ability may serve to reduce the impact of distractors through blocking or perhaps inhibition. This aspect of executive attention – the general, domain-free capability to actively maintain information when confronted by interference – is crucial to broadly predicting higher-order cognitive performance across stimulus and task domains.

In the second paper (Kane & Engle, 2000a), we evaluate working memory, executive attention, and *g*F research conducted from a neuroscience perspective. Our critical review of single-cell, brain-imaging, and neuropsychological research indicates that PFC cells and circuitry are critical to executive attention functions, and that normal individual differences in executive attention may reflect individual differences in PFC functioning. The PFC is but one critical structure in a network of anterior and posterior "attention" areas. However, the PFC does have a unique executive-attention role in maintaining the activation of mental representations – including those of stimuli and goals – in contexts bearing significant sources of interference.

References

Ahola, K., Vilkki, J., & Servo, A. (1996). Frontal tests do not detect frontal infarctions after ruptured intracranial aneurysm. Brain and Cognition, 31, 1-16.

Allport, D. A., Styles, E.A., & Hsieh, S. (1994). Shifting attentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), Attention and performance XV (pp. 421-452). Hillsdale, NJ: Erlbaum.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), Memory (pp. 237-313). New York: Academic Press.

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

Baddeley, A. D. (1993). Working memory or working attention? In A. Baddeley & L. Weiskrantz (Eds.), Attention: Selection, awareness, and control (pp. 152-170). Oxford: 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.

Baddeley, A. D., & Logie, R. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 28-61). New York: Cambridge University Press.

Baddeley, A. D., Lewis, V. J., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. Journal of Experimental Psychology: General, 13, 518-540.

Baldwin, J. M. (1894). Mental development in the child and the race. New York: Macmillan.

Benton, S. L., Kraft, R. G., Glover, J. A., & Plake, B. S. (1984). Cognitive capacity differences among writers. Journal of Educational Psychology, 76, 820-834.

Butler, K. M., Zacks, R. T., & Henderson, J. M. (1999). Suppression of reflexive saccades in younger and older adults: Age comparisons on an antisaccade task. Memory and Cognition, 27, 584-591.

Butters, M. A., Kaszniak, A. W., Glisky, E. L., Eslinger, P. J., & Schacter, D. L. (1994). Recency discrimination deficits in frontal lobe patients. Neuropsychology, 8, 343-353.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33, 386-404.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. Psychological Review, 99, 45-77.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. Psychological Review, 97, 332-361.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. Journal of Experimental Psychology: General, 123, 354-373.

Conway, A. R. A., & Engle, R. W. (1996). Individual differences in working memory capacity: More evidence for a general capacity theory. Memory, 4, 577-590..

Conway, A.R.A., Kane, M.J., & Engle, R.W. (1999). Is Spearman's g determined by speed or working memory capacity? Book review of Jensen on Intelligence-g-Factor. Psycoloquy, 10 (074), article 16.

Cowan, N. (1995). Attention and Memory: An Integrated Framework. Oxford: Oxford University Press.

Craik, F. I. M., & Birtwistle, J. (1971). Proactive inhibition in free recall. Journal of Experimental Psychology, 91, 120-123.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. Journal of Experimental Psychology: General, 125, 159-180.

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

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 561-584.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. Journal of Memory and Language, 25, 1-18.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. Gazzaniga (Ed.), The cognitive neurosciences (pp. 721-733). Cambridge, MA: MIT Press.

Egly, R., & Homa, D. (1984). Sensitization of the visual field. Journal of Experimental Psychology: Human Perception and Performance, 10, 778-793.

Engle, R. W., & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. J. Sternberg (Ed.), The nature of cognition (pp. 515-555). Cambridge, MA: MIT Press.

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, and Cognition, 18, 972-992.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. Journal of Educational Research, 84, 253 - 262.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. Psychological Science, 6, 122-125.

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

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

Fukushima, J., Fukushima, K., Miyasaka, K., & Yamashita, I. (1994). Voluntary control of saccadic eye movement in patients with frontal cortical lesions and Parkinsonian patients in comparison with that in schizophrenics. Biological Psychiatry, 36, 21-30.

Gevins, A., Smith, M.E., McEvoy, L. & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. Cerebral Cortex, 7, 374-385.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category exemples. Journal of Experimental Psychology: Human Learning and Memory, 6, 225-240.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. Vision Research, 18, 1279-1296.

Hallett, P. E., & Adams, B. D. (1980). The predictability of saccadic latency in a novel voluntary oculomotor task. Vision research, 20, 329-339.

Harvey, N. (1984). The Stroop effect: Failure to focus attention or failure to maintain focusing? Quarterly Journal of Experiment Psychology: Human Experimental Psychology, 36(1-A), 89-115.

Hasher, L., & Johnson, M. K. (1975). Interpretive factors in forgetting. Journal of Experimental Psychology: Human Learning and Memory, 1, 567-575.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and chrystallized intelligence. Acta Psychologica, 26, 107 - 129.

Jensen, A. R. (1965). Scoring the Stroop test. Acta Psychologica, 24, 398-408.

Kane, M. J., & Engle, R. W. (2000a). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence. Manuscript submitted for publication.

Kane, M. J., & Engle, R. W. (2000b). An individual-differences approach to working-memory capacity, executive attention, and general fluid intelligence. Manuscript submitted for publication.

Kane, M. J., & Engle, R. W. (in press). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (in press). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General.

Kane, M. J., Sanchez, A., & Engle, R. W. (1999, November). Working memory capacity, intelligence, and goal neglect in the Stroop task. Poster presented at the annual meeting of the Psychonomic Society, Los Angeles, CA.

Kane, M.J., Conway, A.R.A., & Engle, R.W. (1999). What do working memory tests really measure? Behavioral and Brain Sciences, 22, 101-102.

Kiewra, K. A., & Benton, S. L. (1988). The relationship between information processing ability and notetaking. Contemporary Educational Psychology, 13, 33-44.

King, J. & Just, M. A. (1991). Indvidual differences in syntactic processing: The role of working memory. Journal of Memory and Language, 30, 580-602.

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.

LaBerge, D. (1983). The spatial extent of attention to letters and words. Journal of Experimental Psychology: Human Perception and Performance, 9, 371-379.

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin, 109, 163-203.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 1423-1442.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. Journal of Memory and Language, 40, 25-40.

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.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), The frontal granular cortex and behavior (pp. 313-334). New York: McGraw-Hill.

Moscovitch, M. (1994). Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. Neuropsychology, 8, 524-534.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory (vol. 4, pp. 1-18). New York: Plenum Press.

Ormrod, J. E., & Cochran, K. F. (1988). Relationship of verbal ability and working memory to spelling achievement and learning to spell. Reading Research and Instruction, 28, 33 - 43.

Pascual-Leone, J. & Baillargeon, R. (1994). Developmental measurement of mental attention. International Journal of Behavioral Development, 17, 161-200.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.

Phillips, L. H. (1997). Do "frontal tests" measure executive function?: Issues of assessment and evidence from fluency tests. In P. Rabbitt (Ed.), Methodology of frontal and executive function (pp. 191-213. East Sussex, UK: Psychology Press Ltd.

Piaget, J. (1926). The language and thought of the child. New York: Routledge & Kegan Paul.

Posner, M. I., & Peterson, S. E. (1990). The attention system of the human brain. Annual Review of Neuroscience, 13, 25-42.

Posner, M. I., & Snyder, C. R. (1975). Facilitation and inhibition in the processing of signals. In P.M.A. Rabbitt & S. Dornic (Eds.), Attention and performance V, New York: Academic Press.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. Memory and Cognition, 1, 19-40.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. Journal of Verbal Learning and Verbal Behavior, 7, 672-694.

Raven, J. C., Court, J. H., & Raven, J. (1977). Raven's Progressive Matrices and Vocabulary Scales. New York: Psychological Corporation.

Roberts, R. J., Jr., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. Journal of Experimental Psychology: General, 123, 374-393.

Rogers, R. D., & Monsell, S. (1995). The cost of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124, 207-231.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. Journal of Experimental Psychology: General, 126, 211-227.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. Journal of Memory and Language, 39, 418-436.

Salthouse, T. A. (1991). Theoretical perspective on cognitive aging. Hillsdale, NJ: Erlbaum.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Shallice, T., & Burgess, P. W. (1993). Supervisory control of action and thought selection. In A. Baddeley & L. Weiskrantz (Eds.), Attention: Selection, awareness, and control (pp. 171-187). Oxford: Oxford University Press.

Shute, V. J. (1991). Who is likely to acquire programming skills? Journal of Educational Computing Research, 7, 1-24.

Spilsbury, G. (1992). Complexity as a reflection of the dimensionality of a task. Intelligence, 16, 31-45.

Sternberg, S. (1966). High speed scanning in human memory. Science, 153, 652-654.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. Neuropsychology, 11, 138-146.

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

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to memory. Psychological Review, 77, 1-15.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. Journal of Verbal Learning and Verbal Behavior, 2, 440-445.

Yntema, D.B. (1963). Keeping track of several things at once. Human Factors, 5, 7-17.

Scientific publications resulting from the grant

1. Engle, R.W. (in press). What is working memory capacity? In Roediger, H. L., Nairne, J.S., Neath, I., Suprenant, A. M. The Nature of Remembering: Essays in Honor of Robert G. Crowder. Washington, D.C.: American Psychological Association Press.

2. Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (in press). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General

3. Kane, M.J., & Engle, R.W. (in press). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition.

4. Conway, A. R. A., Kane, M. J., & Engle, R. W. (1999). Is Spearman's g determined by speed or working memory capacity? Book review of Jensen on Intelligence-g-factor. Psycoloquy, 10(074).

5. Conway, A.R.A, Tuholski, S.W., Shisler, R.J. & Engle, R.W. (1999). The effect of memory load on negative priming: An individual differences investigation. Memory & Cognition, 27, 1042-1050.

6. Kane, M.J., Conway, A.R.A., & Engle, R.W. (1999). What do working memory tests really measure? Behavioral and Brain Sciences, 22, 101-102.

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

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

9. Engle, R.W. & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), The Nature of Cognition (pp. 514-555). Cambridge, MA: MIT Press.

10. Rosen, V. M. & Engle, R. W. (1998). Working memory capacity and suppression. Journal of Memory and Language, 39, 418-436.

11. Engle, R. W. & Conway, A. R. A. (1998). Working memory and comprehension. In R. H. Logie, & K. J. Gilhooly, (Eds.), Working memory and thinking (pp 67-92). East Sussex, UK: Psychology Press.

12. Rosen, V. M. & Engle, R. W. (1997). Forward and backward serial recall. Intelligence, 25, 37-47.

13. Rosen, V.M. & Engle, R.W. (1997). The role of working memory capacity in retrieval. Journal of Experimental Psychology: General, 126, 211-227.