The influence of complex working memory span task administration methods on prediction of higher level cognition and metacognitive control of response times

DAVID P. MCCABE
Colorado State University, Fort Collins, Colorado

Participants between the ages of 18 and 80 were tested on a complex working memory span task that was administered either using a typical experimenter-paced method or using a method in which the processing component was presented at a fixed, limited-pace presentation rate. Path analyses revealed that even after controlling for individual differences in general processing speed, the limited-pace task predicted unique variance in episodic memory, executive functioning, and fluid intelligence, whereas the experimenter-paced task did not. For the experimenter-paced task, slower responses on the processing component of the task were associated with better recall, but only when individual differences in processing speed were controlled. These findings suggest that metacognitive control of response times affects recall from working memory span tasks, as well as the relationship between span task recall and high-level cognition. These results support resource-sharing explanations of working memory and suggest that limiting processing times using computer pacing of complex span tasks can be an effective way to efficiently measure working memory capacity.

The concept of working memory (WM) has become central to understanding complex cognition. Although researchers debate the specific details of models of WM, it can generally be defined as the cognitive system responsible for maintaining information or task goals in an active state over brief periods of time. Part of the reason for the success of WM as a model of short-term cognitive processing is that individual differences in the ability to maintain and manipulate information in WM, referred to as *working memory capacity* (WMC; Engle, Tuholski, Laughlin, & Conway, 1999), have been linked to various measures of higher level cognition, including reading comprehension (Daneman & Carpenter, 1980; Turley-Ames & Whitfield, 2003), episodic memory (McCabe & Smith, 2002; Oberauer, 2005; Park et al., 2002), executive function (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001), and general fluid intelligence (Engle et al., 1999; Kyllonen & Christal, 1990).

Research on individual differences in WMC has primarily been based on examination of the relationship between performance on complex span tasks (e.g., reading span) and measures of higher level cognition. These complex span tasks typically require participants to encode and maintain several to-be-remembered items while completing some interpolated processing task, such as reading or arithmetic, that is intended to disrupt that maintenance. The participants are further required to recall the to-be-remembered items in serial order. Complex span tasks can be contrasted with simple span tasks, which only require the maintenance and retrieval of to-be-remembered items in serial order. Recall is poorer for complex span tasks than for simple span tasks, and complex span tasks typically show stronger correlations with higher level cognition (e.g., Ackerman, Beier, & Boyle, 2005; but see Unsworth & Engle, 2007, for a detailed discussion of the similarities between simple and complex span tasks). Consequently, complex span tasks have become popular tools in the investigation of individual differences in many areas of psychology (see Engle & Kane, 2004, for a review).

Some models of WM suggest that the reason that complex span tasks are more strongly related to complex cognition than are simple span tasks is that the former are more likely than the latter to engage the central executive component of WM (Baddeley, 2000; Engle et al., 1999). Indeed, individual differences in WMC have often been conceptualized as the efficiency of the central executive component of the WM system (Engle et al., 1999; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010), and this is the definition of WM capacity used in the present study. Of course, complex span tasks engage other abilities in addition to executive control processes. For example, Engle's model of WMC (see Engle et al., 1999) indicates that task-specific strategies related to grouping or maintenance strategies, which are unrelated to atten-

D. P. McCabe, david.mccabe@colostate.edu

tional control (e.g., phonological rehearsal), are distinct from central executive functioning but may influence recall from span tasks.

Methods of Administering Complex Span Tasks

Complex span tasks are typically administered individually, and the experimenter influences the speed with which the processing component is completed (Conway et al., 2005). For example, in the reading span task, participants are typically asked to read sentences aloud (e.g., *There are 7 days in every week*), decide whether a sentence makes sense, and then remember the final word in the sentence (Daneman & Carpenter, 1980). Immediately after the participants have made their decision for a given sentence, the experimenter advances the screen to reveal the next sentence. Trials typically range in length between two and five sentences, and the participants are asked to complete the processing component of the task as quickly as possible while maintaining accuracy. Although the experimenter has some influence on the speed at which the processing component of the task is completed, by controlling when the screen advances, the participants are able to complete the processing tasks at their own pace. Thus, the participants have some control over the amount of time spent on the processing tasks. Previous research indicates that participants often take additional time to rehearse or refresh the to-be-remembered items during complex span task performance (Engle, Cantor, & Carullo, 1992; St. Clair-Thompson, 2007). Because the dependent measure for a span task is recall performance, considering the factors that influence recall performance is crucial to accurately estimating and understanding the WMC construct and its relation to higher level cognition.

I will refer to the traditional administration method used for complex span tasks as *experimenter-paced*, because the experimenter controls when the screen advances to reveal the to-be-remembered item (cf. Friedman & Miyake, 2004; St. Clair-Thompson, 2007). This method can be contrasted with two other methods: a self-paced administration method, in which the participants completely control advancement of the screen following each processing item (e.g., by pressing the space bar when they have completed each processing item; Friedman & Miyake, 2004; Waters & Caplan, 1996), and a limited-pace administration method, in which the amount of time available for processing during the task is restricted in an effort to reduce the amount of processing and maintenance that the participants can engage in (Lépine, Barrouillet, & Camos, 2005; Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000).

Two primary issues were addressed in the present study, both associated with the administration method of complex WM span tasks. The first issue of interest is how administration methods influence the relation between the WM task performance and measures of complex, or higher level, cognition. Specifically, the predictive power of experimenter-paced tasks was compared with that of limited-pace tasks. The other primary issue addressed in the present study is the control of response times during complex WM span tasks. Specifically, response times for the processing component of experimenter-paced tasks were examined for evidence of strategic slowing of processing times in response to increases in memory load (i.e., trial length).

The Influence of Administration Method on the Predictive Power of Complex Span Tasks

Recently, several studies have been reported in which the influence of administration method on the predictive power of span tasks was examined. For example, experimenter-paced tasks show stronger correlations with higher level cognition than self-paced tasks, indicating that when participants are allowed extra time to strategically allocate attention to maintenance (in the self-paced conditions), the correlation between the span task and higher level cognition is weakened (Friedman & Miyake, 2004; St. Clair-Thompson, 2007). The finding that self-paced versions of complex span tasks show weaker predictive power than experimenter-paced versions is important, but as Friedman and Miyake (2004) noted, most researchers who administer complex span tasks follow a procedure similar to the experimenter-paced version (as opposed to the self-paced version). Although experimenters typically ask participants to complete the processing component of the tasks as quickly as possible while maintaining high levels of accuracy, it is quite natural for the participants to attempt to optimize performance by slowing down to improve accuracy if they are able to do so (i.e., a speed–accuracy trade-off; Engle et al., 1992). In addition to the participants' strategically slowing down to improve recall performance, another source of unwanted variability in experimenter-paced span tasks is that some instructions might explicitly state that the participants should complete the processing component of the task as quickly as possible, whereas others may not. Moreover, even if instructions do suggest that the participants complete the processing component quickly, variability in the training and/or assertiveness of experimenters across labs could potentially influence the participants' strategy use during span tasks. For the present purposes, the important point is that studies in which administration method has been examined have shown that allowing participants additional processing time on complex span tasks can reduce the strength of the correlation between these tasks and measures of higher level cognition, presumably because they allow strategies to be used for maintenance, and individual differences in strategy use add noise to the correlations of these tasks with outcome measures (Engle & Kane, 2004; Turley-Ames & Whitfield, 2003).

With respect to the relationship between processing and maintenance during complex span tasks, the present study is guided by the general principles of many resource-sharing explanations of WM. Specifically, the theoretical framework adopted here is based on resource-sharing explanations that propose that processing and maintenance functions rely on a single pool of attentional resources (Barrouillet, Bernardin, & Camos, 2004; Engle et al., 1999; McCabe et al., 2010), and thus, *resource sharing* refers specifically to models of this type. Although these resource-sharing theories have conceptualized the atten-

tional resources somewhat differently, they share the common notion that processing and maintenance rely on a single pool of attentional resources. These resource-sharing theories predict that WM performance (i.e., recall from complex span tasks) should be reduced when less time, and therefore less attention, is available to concurrently complete processing and maintenance functions during span task performance.

Recently, Barrouillet and colleagues (Barrouillet, Lépine, & Camos, 2008; Lépine et al., 2005) proposed that the most effective method of measuring the limited attentional capacity of the WM system is to limit the amount of time available for maintenance and processing by having the computer present items at a faster pace than would typically be allowed by using an experimenter-paced task (e.g., by using a limited-pace task). The rationale for using a limited-pace method of administration is based on the time-based resource-sharing model (Barrouillet et al., 2004), which suggests that maintenance and processing during complex span tasks are constrained by the number of elementary cognitive operations that must be completed in a given period of time. Indeed, the time-based resource-sharing theory operationalizes cognitive load as the number of retrievals that can be completed in a specified period of time (assuming the difficulty of those retrievals is held constant). Stated differently, one can consider cognitive load as the amount of attentional capture by the processing and maintenance requirements of the task (Barrouillet et al., 2008). Thus, when a limited-pace task restricts the time provided for processing and maintenance relative to an experimenter-paced or a self-paced task, the influence of individual differences in strategic processing on performance should be reduced, thereby reducing span task recall as well (Lépine et al., 2005; see also Engle & Kane, 2004; St. Claire-Thompson, 2007; Turley-Ames & Whitfield, 2003).

The use of limited-pace tasks to minimize individual differences in strategic processing can be contrasted with approaches that attempt to minimize strategic processing by tailoring the pace of the processing component to each individual participant's ability on the processing component. This has been accomplished by using either experimenter-paced tasks (Conway et al., 2005) or tasks with a custom pace based on prescreening each participant's ability on the processing component of the task (e.g., providing 2.5 *SD*s of additional time per processing operation relative to the baseline average; Unsworth, Heitz, Schrock, & Engle, 2005). Although all of these administration methods are intended to limit strategic processing, thereby reducing noise contributing to the correlations between span tasks and higher level cognition, it is unclear whether these approaches are equally successful in this regard. If experimenter-paced or custom-paced tasks provide considerable time to engage strategic retrieval or rehearsal processes, according to the principles of the time-based resource-sharing model, cognitive load should be reduced and will measure WMC less efficiently.

Despite the apparent importance of administration method with respect to effectively measuring individual differences in WMC, to date, in no study have experimenter-paced and limited-pace administration methods been directly compared for the same task in order to compare their predictive power. Lépine et al. (2005) did compare limited-pace and experimenter-paced tasks, although the processing components of the tasks differed (because they were interested in issues of the complexity of the processing task in addition to administration method). Specifically, the experimenter-paced tasks required completion of traditional math problems (e.g., $6 + 2 + 3 = 13$?), whereas the limited-pace task required adding or subtracting 1 from some root number on successive frames at a fixed pace (e.g., $8 - 1 + 1$; cf. Barrouillet et al., 2004). Lépine et al. reported that novel limited-pace complex span tasks showed stronger correlations with reading comprehension than more traditional experimenter-paced tasks. In the present study, identical processing operations were included in the experimenter-paced and limited-pace tasks, allowing a direct comparison of administration method in terms of predictive power. Moreover, because Lépine et al. only included sixth-grade children, and children can differ from adults in terms of the complexity of the retrieval structures used to buttress strategies, as well as strategy execution (Gaultney, Kipp, & Kirk, 2005; Lemaire & Lecacheur, 2002), it is unclear to what extent their findings would generalize to adults. Finally, Lépine et al. compared experimenter-paced and limited-pace tasks within subjects, and the order of the tests was not counterbalanced, leading to potential carryover or practice effects. Practice effects can, under some conditions, attenuate the correlation between span task performance and higher level cognition (Blalock & McCabe, 2008). In the present study, we examined administration method between subjects, which avoids the influence of practice or carryover effects.

One of the primary purposes of the present study was to examine whether the experimenter-paced or limited-pace task was more strongly related to measures of higher level cognition in an adult life-span sample of adults. Previous research in which self-paced and experimenter-paced tasks have been examined has shown that allowing additional time for processing suppresses the predictive power of span tasks (e.g., Friedman & Miyake, 2004; Lépine et al., 2005; St. Clair-Thompson, 2007). Thus, it was expected that the limited-pace task, which provided less time for processing and maintenance than the experimenter-paced task, would be more strongly related to measures of higher level cognition, because it would limit the amount of strategic processing that the participants could engage in during the task. On the basis of time-based resource-sharing theory, limiting time to engage in strategic processing should make the task more effective in terms of measuring individual differences in WMC by increasing cognitive load and limiting the influence of individual differences in strategic processing (see Lépine et al., 2005). However, as was mentioned previously, limited-pace tasks may be more strongly related to higher level cognition because they are more sensitive to individual differences in general processing speed. Indeed, the time-based resource-sharing theory defines *cognitive load* as the number of retrievals required in a specified time

period, which suggests that speed of processing should be strongly related to the limited-pace task performance. However, the time-based resource-sharing theory also posits that it is the efficiency with which attentional refreshing of to-be-remembered items can be maintained during task pauses that ultimately determines one's WM performance, a process that is not expected to be captured by simple perceptual speed measures. Thus, to the extent that speed of processing and attentional refreshing are distinct variables, the limited-pace task should account for additional variance in complex cognition beyond that accounted for by speed of processing.

In order for the role of processing speed in influencing the predictive power of span tasks to be examined, regression models were computed in which general speed of processing was controlled, and the relations between span task recall and multiple measures of higher level cognition (episodic memory, executive function, and fluid intelligence) were assessed for both administration methods.

Metacognitive Control of Response Times During Complex Span Tasks

Another purpose of the present study concerned metacognitive control of response times during WM span tasks. Many resource-sharing theories of WM (e.g., Barrouillet et al., 2004; Engle et al., 1999) propose that the efficiency of the WM system is constrained by a general, domain-free attention ability, or energy, to be shared between maintenance and processing during span task performance. Because long-term memory retrieval, attentional refreshing, and other maintenance processes require time to complete, response times should increase when memory load increases. That is, because more to-be-remembered items must be maintained with longer trial lengths, increased memory load should lead to slower processing so that more attentional resources can be devoted to maintenance. This slowing has typically been considered a reactive control strategy (Braver, Gray, & Burgess, 2007), such that response times trade off with maintenance during online task performance as attentional resources are depleted (i.e., as memory load increases during a trial). This is an inherent characteristic of resource-sharing theories that assume that a single general attentional resource is shared between processing and maintenance functions. However, resource-sharing theories have not explicitly addressed in great detail the potential metacognitive aspects of increases in response times (although this is an explicit central executive function of some WM models; e.g., Baddeley, 1986). In the present study, trials were presented in an ascending fashion, beginning with the shortest trials and proceeding through the longer trials. Because the participants were aware of the list length of trials prior to presentation, it is possible that they could use a proactive control strategy (Braver et al., 2007), such that they responded more slowly as trial lengths become longer on the basis of anticipated memory load. In order to examine this issue and to ensure that any increases in response times reflected proactive metacognitive control, only the response times for the processing component prior to presentation of the first item of each trial were examined as a func-

tion of list length.[1] Thus, because there was no memory load for the processing decision prior to presentation of the first to-be-remembered item on each trial, increases in response times as a function of trial length would reflect anticipatory or proactive metacognitive control. Evidence for proactive control of response times would suggest that participants are aware of WM limitations and control response times to compensate for increasing maintenance demands for longer trial lengths.

Comparing age differences in proactive metacognitive control was also important, because older adults have sometimes shown a tendency to allocate attentional resources less efficiently than younger adults during task performance (Dunlosky & Connor, 1997; Salthouse & Coon, 1994; Touron & Hertzog, 2004) but at other times have shown similar levels of efficiency (Dunlosky & Hertzog, 1997; Miller & Stine-Morrow, 1998). Given the centrality of complex span tasks in investigations of age differences in WM, understanding the effect of aging on strategy use on complex span tasks is of considerable importance.

The issue of metacognitive control of processing times was also examined through comparison of age-related differences in recall performance on the experimenter-paced task. Previous research in which younger adult samples were used has shown that slowing processing during self-paced complex span tasks improves recall accuracy relative to experimenter-paced tasks (Friedman & Miyake, 2004; St. Clair-Thompson, 2007). However, in a life-span sample, like the one used in the present study, aging is expected to both slow the speed at which processing is completed and reduce the level of recall. Thus, processing times may not show a strong relationship with recall performance in a life-span sample, because age-related slowing might suppress the relation between processing time and recall. Suppression effects are often complex, making them difficult to discover and interpret, but they represent a clear danger with respect to interpreting relations between variables (Cramer, 2003). Because aging is associated with general declines in speed of processing (Salthouse, 1996), but slower processing should reduce cognitive load once the general age-related decline in speed of processing is controlled for, slower processing may be related to improved recall performance. That is, age differences in processing speed may suppress the processing–maintenance trade-offs predicted by resource-sharing explanations, leading researchers to erroneously conclude that processing and storage are independent from one another.

Overview of the Present Study

In the present study, a complex span task, called *Stroop span* (McCabe, Robertson, & Smith, 2005; McCabe & Smith, 2002; McCabe, Smith, & Parks, 2007), was administered to a sample of adults between the ages of 18 and 80 using one of two administration methods. Stroop span is structurally identical to typical complex span tasks, like reading span and computation span, in that participants are required to engage in a processing component that required controlled attention, to make a decision, and to encode to-be-remembered items presented as part of the processing component. Specifically, the processing com-

ponent of the task required that the participants determine whether the meaning and the font color of a Stroop color word matched, and respond aloud. The participants also had to remember the color of each of the words and, after between two and six of these words were presented, were asked to recall the colors in serial order. Half of the words were congruent (e.g., the word *red* in red font), and half were incongruent (e.g., the word *blue* in green font), and thus, the answer was *yes* for the congruent items and *no* for the incongruent items. The validity of the task as a WMC measure is supported by previous research showing that Stroop span performance is related to a performance on a traditional complex span task (i.e., computation span; McCabe et al., 2005), episodic memory (McCabe et al., 2007), executive functioning (McCabe et al., 2005), and the ability to avoid memory errors (McCabe & Smith, 2002; McCabe et al., 2007).

Obviously, there are differences between Stroop span and other more traditional complex span tasks. For example, the processing task in Stroop span is simpler than more typically used complex span tasks (e.g., reading and computation span), and certainly more novel than these more typical tasks. Nonetheless, the argument presented here is that any complex span task will effectively measure cognitive load to the extent that it consumes time that would otherwise be used for additional strategy use (e.g., subvocal rehearsal). In that respect, traditional span tasks like reading and operation span appear to be amenable to strategic rehearsal (Bailey, Dunlosky, & Hertzog, 2009; Kaakinen & Hyönä, 2007). Certainly, Stroop span also allows strategic rehearsal when additional time is available for processing, but it is not apparent that Stroop span would be more or less amenable to these sorts of strategies than other span tasks.

In the experimenter-paced Stroop span task, participants were asked to complete the processing component as quickly as possible while still maintaining high accuracy on the processing task, whereas in the limited-pace Stroop span task, the processing component was presented at a fixed, computer-controlled pace. In order to assess the predictive power of the two different versions of Stroop span, tests of several additional cognitive constructs were administered as outcome measures, including tests of fluid intelligence, executive functioning, and episodic memory. In previous studies, moderate to strong relationships have been found between complex span task performance and fluid intelligence (Engle et al., 1999; Kyllonen & Christal, 1990), executive functioning (McCabe et al., 2010; Miyake et al., 2001), and episodic memory (McCabe et al., 2007; Park et al., 1996). Using multiple outcome measures of higher level cognition in the present study has the advantage of providing converging evidence regarding the relationship between experimenter-paced and limited-pace tasks and higher level cognition.

METHOD

Participants

One hundred seventy-eight adults (107 female and 71 male) between the ages of 18 and 80 participated in this study. The participants were recruited from the St. Louis community and were randomly assigned to either the experimenter-paced or the limited-pace condition. Average age, sex, education, self-reported health, and vocabulary scores are presented in Table 1. None of these values differed significantly (at $p < .05$) between the limited-pace and experimenter-paced groups. Age was not related to sex, education, or self-reported health, although it was negatively correlated with scores of the mini-mental status exam (MMSE; Folstein, Folstein, & McHugh, 1975) and positively correlated with vocabulary scores (as measured by the Shipley Institute of Living Scale; Zachary, 1986). These findings are consistent with the characteristics of other studies of this type (e.g., Park et al., 2002).

Materials and Procedure

The participants were tested in two sessions, each lasting approximately 2.5 h. Two breaks were given during each session to reduce fatigue and interference between tasks. Other tasks not included as part of the present study were administered as well, as part of a large study of aging and cognition (see McCabe et al., 2010). Sessions were spaced at least 1 week apart, but never more than 3 weeks apart.

Stroop span (WMC). As described above, the Stroop span task involved the presentation of Stroop color words that were either congruent (e.g., the word *red* in red font) or incongruent (e.g., the word *blue* in green font). The words and colors red, yellow, green, and blue were used. For each word, the participants had to decide whether the word and the color matched; the answer was *yes* for the congruent items and *no* for the incongruent items. The participants also had to remember the color of each of the words and, after a series of these words were presented, were asked to recall the colors in serial order. Half of the words were congruent, and half were incongruent on a given trial if there was an even number of items (e.g., span length 4), and there was one extra congruent or incongruent word on trials with an odd number of items (e.g., span length 3).

The experimenter pressed the space bar to begin a trial when the participants were ready. Stroop color words were presented on the computer screen on a black background (Arial Black, 96 pt.), and the participants gave verbal responses indicating whether the color and the word matched and gave their recall responses verbally as well. After all of the words on a given trial were presented, a series of question marks appeared on the screen prompting recall. The task started with trials of length 1 and proceeded through trials of length 6 if the participants continued to correctly recall the trials. There were three trials at each length. After length 2, if the participants failed to recall all of the words on a trial correctly for two of the three trials at that length, the task was discontinued. The number of trials correct through the length at which the participants correctly recalled two of the three trials was the dependent measure (Daneman & Carpenter, 1980).

Prior to completing the span task, the participants were given practice making decisions about whether Stroop color words matched for 24 trials (12 congruent and 12 incongruent, randomly intermixed). For the first 8 practice trials, the words were presented at a rate of one

Table 1
Demographic Characteristics of Participants

| | Stroop Span Administration Method | | | | |
| | Limited Pace | | Experimenter Paced | | |
Variable	M	SD	M	SD	Age r
N	86	–	92	–	
Age	54.5	20.3	50.5	20.0	
Percentage female	67	–	53	–	−.03
Education	15.2	2.6	15.1	2.4	.00
Self-reported health	4.1	0.7	4.3	0.8	−.08
MMSE	29.1	1.1	29.2	0.9	−.26**
Shipley vocabulary	33.7	3.7	34.2	3.8	.25**

Note—MMSE, mini-mental state examination.   **$p < .01$.

every 3 sec; for the next 8, one every 2.5 sec; and for the last 8, one every 2 sec. The slower early trials during practice were included so that the participants could become acquainted with the novel processing task prior to faster processing being required. During the actual span task, in the experimenter-paced version of the task, the participants were asked to make their decision (i.e., whether the color and the word matched) as quickly as possible while maintaining near perfect accuracy. In the limited-pace condition, the participants were told that the word would be presented for 2 sec, and they had to make their decision within that time period. All of the participants were capable of making this decision within the 2 sec during the practice phase.

Executive functioning. The executive function factor was based on a battery developed by Glisky, Polster, and Routhieaux (1995) that has been used to measure executive functioning in previous research (Glisky & Kong, 2008; McCabe et al., 2010; Van Petten et al., 2004). The Wisconsin card sorting task (WCST) requires participants to sort cards on the computer on the basis of one of three dimensions, according to a rule that changes after every 10 correct trials (Heaton, 1993). The outcome measure was the number of perseverative errors, in which the participants continued to sort on the basis of a rule that had changed. The letter fluency task requires participants to generate as many unique words as they can in a minute for each of three letters (the letters F, A, and S; Thurstone, 1938). The Mental Control test (WMS-III; Wechsler, 1997a) requires participants to articulate well-learned sequences (e.g., days of the week) both forward and backward as quickly as possible. A speed–accuracy score was calculated. The Mental Arithmetic test (WAIS; Wechsler, 1997b) requires participants to solve verbal math problems of varying levels of difficulty. A speed–accuracy score was calculated. Finally, backward digit span (WMS-III; Wechsler, 1997a) requires participants to recall digit strings in the reverse order from that in which they were read aloud to them. The numbers of trials correct for the length at which both trials were correct was the outcome measure.

Episodic memory. The episodic memory factor was based on three tasks, all requiring immediate free recall. These included immediate recall of brief stories (i.e., Logical Memory I from the WMS-III; Wechsler, 1997a), immediate free recall of a 16-word list (Trial 1 from the California Verbal Learning Test [CVLT]; Delis, Kramer, Kaplan, & Ober, 2000), and immediate free recall of a 40-word list (4 words from each of 10 thematic lists from Roediger, Watson, McDermott, & Gallo, 2001). Recall was spoken for the Logical Memory and CVLT tasks and written for the 40-word list recall task.

Processing speed. Processing speed was measured using three tasks, including letter comparison (Salthouse & Babcock, 1991), pattern comparison (Salthouse & Babcock, 1991), and digit–symbol substitution (Wechsler, 1997a). Letter and pattern comparison tasks require participants to compare pairs of letter strings or simple line patterns as quickly as possible, to determine whether they are the same or different and to write an S or a D in a space in between each pair. Two pages are completed, 30 sec are given for each, and the number of correct responses for each of the two pages is the measure of performance. Digit–symbol substitution requires participants to write digits next to symbols as quickly as possible according to a key at the top of the page denoting which digits were to be paired with which symbols. The number correctly completed in 90 sec was the measure of performance.

Fluid intelligence. Fluid intelligence was measured using Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998), letter sets (Ekstrom, French, Harman, & Dermen, 1976), and space relations (Ekstrom et al., 1976). Raven's involves choosing one of several abstract figures to complete a series of abstract figures that adhere to a particular rule. The version used here included only the odd numbered items (from Salthouse, Atkinson, & Berish, 2003). Letter sets involve choosing one of four letter sets that does not adhere to the rule common to the other letter sets. Finally, space relations involve choosing the one of four three-dimensional objects that corresponds to an unfolded two-dimensional target object.

RESULTS

All effects are significant at $p < .05$ unless otherwise noted.

Stroop Span Recall

Although age was a continuous variable in the present study, for clarity of presentation, the sample was divided into younger (18–49 years old) and older (50–80 years old) groups for some analyses, including overall recall. Figure 1 shows the mean number of correct trials for Stroop span for younger and older adults, for both the experimenter-paced (open bars) and limited-pace (black bars) versions. Younger adults showed approximately equivalent recall, regardless of whether the task was experimenter paced or limited pace, whereas older adults showed better recall in the experimenter-paced than in the limited-pace condition. These observations were confirmed statistically by conducting an ANOVA with two between-subjects variables: age group (younger, older) and administration method (experimenter paced, limited pace). There was an overall effect of age group on recall from Stroop span [$F(1,174) = 37.82$, $MS_e = 207.25$, $\eta_p^2 = .18$], a nonsignificant effect of administration method [$F(1,174) = 2.98$, $MS_e = 16.33$, $p = .08$, $\eta_p^2 = .02$], and a significant interaction between the two [$F(1,174) = 7.85$, $MS_e = 43.03$, $\eta_p^2 = .04$]. Planned comparisons of administration method within each age group indicated that administration method had no effect on recall in the younger adults ($F < 1$), but recall was significantly greater in the experimenter-paced condition for the older adults [$F(1,86) = 11.79$, $MS_e = 55.68$, $\eta_p^2 = .12$]. The nonsignificant effect of administration method for the younger adults was likely because the average processing time for each item for the younger adults was substantially less than the fixed pace of 2,000 msec in the limited-pace condition ($M = 1,707$ msec, $SD = 399$). However, note that on the basis of their means and standard deviations, at least some of the younger adults in the experimenter-paced condition
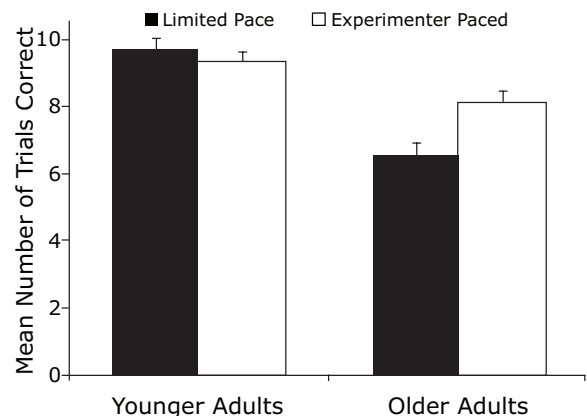

Figure 1. Average number of correctly recalled trials for participants in the younger adult and older adult groups in the limited-pace and experimenter-paced conditions.

took more than 2,000 msec on at least a third of the trials, indicating that they often took more time than they should have needed to complete the processing operations, given the equivalence in recall as a function of administration method. This finding, in which additional time is taken for processing without a benefit to recall, has been referred to as the *labor-in-vain* effect (Nelson & Leonesio, 1988) and appears to have operated in this context. For the older adults, their average processing time in the experimenter-paced condition was quite close to the 2,000-msec fixed pace in the limited-pace task ($M = 1,986$ msec; $SD = 411$), so it is somewhat unsurprising that they performed substantially more poorly in the limited-pace condition, because some of the participants did not have enough time to complete the processing task while maintaining the to-be-remembered items, at least on some subset of the trials. Later in the Results section, a more fine-grained analysis of the relationship between age and processing time in the experimenter-paced task is considered.

Correlations Between Stroop Span and Other Cognitive Constructs

With respect to the reliability of the tasks, the limited-pace Stroop span task (split-half reliability = .86) was slightly more reliable than the experimenter-paced task (split-half reliability = .73), although both of these values exceed the criterion for good internal consistency reliability (i.e., >.70; Nunnaly, 1978). Factor scores were computed for each construct measured in the present study by entering the measures representing each construct into a principal-components analysis using SPSS. The resulting factor scores ranked the participants according to a standardized score based on the weighted combination of performance on measures of each construct. Note that each of these confirmatory factor analyses led to a single-factor solution, so all of the participants were ranked on a single factor for each outcome measure. These scores were computed separately for each of the administration conditions (i.e., experimenter paced and limited pace).

The correlations between the experimenter-paced and limited-pace Stroop span tasks and cognitive factors, as well as age, are presented in Table 2. The correlations above the diagonal are for the limited-pace task, and the correlations below the diagonal are for the experimenter-paced task. Generally speaking, limited-pace Stroop span performance was more strongly related to age and cognitive factor scores than was experimenter-paced Stroop span performance. A direct comparison of the strength of the correlations for the limited-pace and experimenter-paced tasks revealed a significant difference for the correlations with age ($z = 2.43$, $p < .02$) and processing speed ($z = 3.12$, $p < .01$) and marginally nonsignificant differences for executive function ($z = 1.80$, $p = .07$) and fluid intelligence ($z = 1.90$, $p = .06$), but a nonsignificant difference for episodic memory ($z = 1.24$, $p = .21$). In each case, correlations were stronger for the limited-pace task than for the experimenter-paced task, although these differences obviously did not reach conventional levels of significance in many cases. We consider the relationship between each version of Stroop span and these other cognitive constructs in more detail in the analyses that follow.

Path Models Predicting Higher Level Cognition Factors Using Experimenter-Paced and Limited-Pace Stroop Span Tasks

From the bivariate correlation analyses, it is unclear whether the generally stronger correlations between the higher level cognitive factors and the limited-pace Stroop span task, as compared to the experimenter-paced version, were simply due to the increased sensitivity of the limited-pace task to individual differences in general processing speed. Indeed, some have argued that using limited-pace tasks will make those tasks more sensitive to individual differences in processing speed (Engle & Kane, 2004), which could undermine the tasks' sensitivity to individual differences in WMC. In order to examine this issue more closely, path models were computed to examine whether Stroop span recall still accounted for variance in each higher level cognitive factor after controlling for age-related differences in speed of processing. If computer pacing simply makes these tasks more sensitive to individual differences in processing speed, controlling for processing speed should eliminate or substantially reduce the correlation between the limited-pace Stroop span task and higher level cognitive factors. Models with the experimenter-paced task were included as well for comparison purposes.

The path models are shown in Figures 2, 3, and 4. In each path model, in order to account for individual and

Table 2
Correlations Between Age, Experimenter-Paced and Limited-Pace Stroop Span Tasks, and Processing Speed, Episodic Memory, Executive Function, and Fluid Intelligence Factor Scores

Measure	Age	SS	PS	EM	EF	*Gf*
Age	–	−.60	−.69	−.33	−.39	−.61
Stroop span (SS)	−.31	–	.71	.47	.58	.59
Processing speed (PS)	−.73	.39	–	.43	.59	.65
Episodic memory (EM)	−.32	.31	.36	–	.47	.47
Executive function (EF)	−.32	.33	.47	.48	–	.62
Fluid intelligence (*Gf*)	−.61	.37	.69	.44	.54	–

Note—All correlations are significant at $p < .01$. Limited-pace task correlations are presented above the diagonal ($n = 92$), and experimenter-paced are presented below the diagonal ($n = 86$).

age-related differences in processing speed, age was correlated with processing speed, and then processing speed was correlated with Stroop span recall and the measure of higher level cognition. Finally, there was also a direct path from Stroop span recall to the measure of higher level cognition, which was of primary interest. Thus, the model controls for processing speed and allows an examination of the relation between span task recall and higher level cognition. In all cases, the model fits were acceptable, as is indicated by CFIs of greater than .95 (Hu & Bentler, 1995).

The model predicting episodic memory is shown in Figure 2. In Figure 2, Model A included span task recall from the limited-pace task, and Model B included span task recall from the experimenter-paced task. As is shown in Model A, Stroop span recall accounted for significant variance (.34) in episodic memory, even after controlling for age differences in the processing speed factor scores, whereas processing speed did not account for any significant unique variance in episodic memory (.19). In Figure 2, Model B, which included Stroop span recall from the experimenter-paced task, Stroop span recall did not account for significant variance (.20) in episodic memory after controlling for age differences in the processing speed factor, and general processing speed did account for significant unique variance in executive function (.28).

The model predicting executive functioning is shown in Figure 3. In Figure 3, Model A, which included Stroop span recall from the limited-pace task, Stroop span recall accounted for significant variance (.31) in executive function, even after controlling for age differences in the pro-

cessing speed factor, and processing speed also accounted for unique variance in executive function (.32). In Figure 3, Model B, which included Stroop span recall from the experimenter-paced task, Stroop span recall did not account for significant variance (.17) in executive function after controlling for age differences in the processing speed factor, and general processing speed did account for unique variance in executive function (.41).

The models predicting fluid intelligence are shown in Figure 4. In Figure 4, Model A, which included Stroop span recall from the limited-pace task, Stroop span recall accounted for significant variance (.26) in fluid intelligence, even after controlling for age differences in the processing speed factor, and processing speed also accounted for unique variance in fluid intelligence (.47). In Figure 4, Model B, which included Stroop span recall from the experimenter-paced task, Stroop span recall did not account for significant variance (.12) in executive function, even after controlling for age differences in the processing speed factor, but processing speed did account for unique variance in executive function (.64).

In summary, in all three models, the limited-pace Stroop span task accounted for unique variance in measures of higher level cognition, even after controlling for general speed of processing, whereas the experimenter-paced task did not. When models were calculated with span task processing times for the experimenter-paced task as a mediator between processing speed and span task recall (not reported here), the paths between span task recall and each measure of higher level cognition were still not significant. The finding that the limited-pace task accounted for significant variance in higher level cognition,
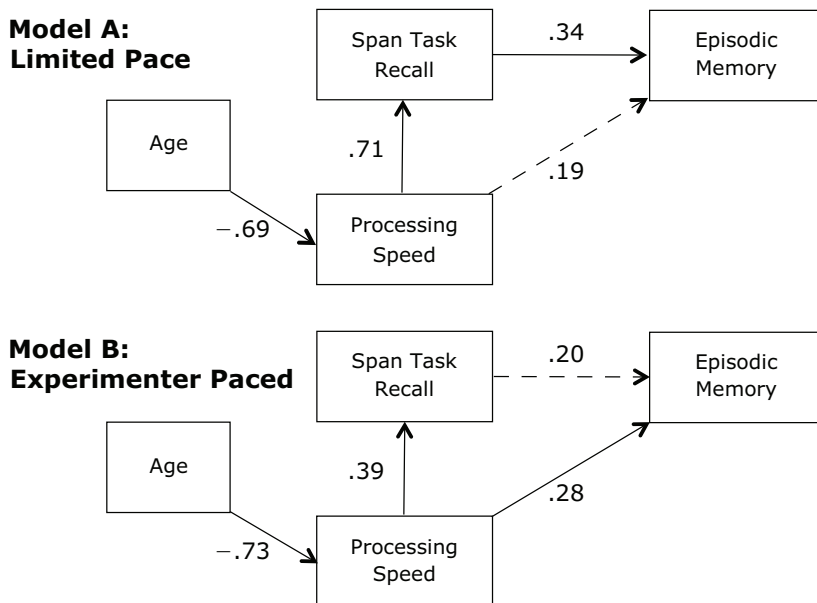

Figure 2. Path models indicating the relation between age, processing speed, span task recall, and episodic memory for the limited-pace and experimenter-paced versions of the Stroop span task. Solid lines indicate paths that are significant at $p < .05$; dotted lines indicate nonsignificant paths.

**Model A:
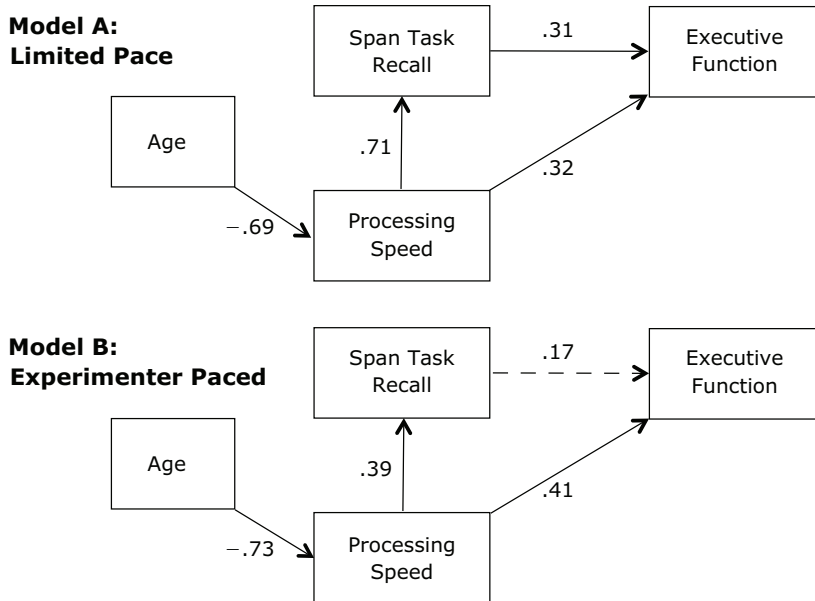Limited Pace**

**Model B:
Experimenter Paced**

Figure 3. Path models indicating the relation between age, processing speed, span task recall, and executive functioning for the limited-pace and experimenter-paced versions of the Stroop span task. Solid lines indicate paths that are significant at $p < .05$; dotted lines indicate nonsignificant paths.

even after controlling for processing speed, is somewhat counterintuitive if one assumes that the reason for the predictive power of the limited-pace task is primarily its increased sensitivity to individual differences in processing speed. Of course, consistent with the bivariate correlations, the limited-pace task was more strongly associated with processing speed (.71) than the experimenter-paced task (.39), and thus, processing speed had much greater indirect effects on higher level cognition in the models including the limited-pace task than in models including the experimenter-paced tasks. By contrast, in models including the experimenter-paced tasks, the direct effects of processing speed were greater than in the models including the limited-pace tasks, but there were no significant direct effects for recall from the experimenter-paced task.

**Response Times for the First Items of Trials
As a Function of Age and Trial Length for the
Experimenter-Paced Stroop Span Task**

In order to examine metacognitive control of processing time on the experimenter-paced task, average response times for the processing decision that was completed prior to presentation of the first to-be-remembered item at each trial length were examined as a function of trial length. Note that for these initial processing decisions, there is no memory load, so any slowing found as trial length increases is presumably due to a proactive strategy intended to reduce cognitive load in order to improve maintenance and, consequently, recall performance. Recall that cognitive load is defined by the time-based resource-sharing theory as the number-of-retrievals:time ratio. Thus, slow-

ing processing increases the available attentional resources for processing and maintenance functions, thereby reducing cognitive load. Although this proactive metacognitive control strategy is not a feature of the time-based resource-sharing model, in principle, it is not incompatible with this conceptualization of cognitive load.

Figure 5 shows the average response time per item for trial lengths 1–4 for each age group. As is shown in the figure, older adults were slower overall, and response times increased with age as a function of increasing trial length, but these variables did not interact. In order to confirm the above-mentioned observations, a mixed-model ANOVA examining average processing times was conducted, with trial length (1, 2, 3, 4) as a within-subjects variable, and age group (younger, older) as a between-subjects variable. Note that 38 younger adults and 31 older adults accurately completed trials through length 4 in the experimenter-paced condition, which meant that over 80% of the participants contributed to this analysis (69 of the 86 participants). There was an increase in response times as a function of trial length $[F(1,67) = 48.39, MS_e = 5,259,946, \eta_p^2 = .42]$, as well as an increase in response times for older adults as compared to younger adults $[F(1,67) = 8.74, MS_e = 4,448,896, \eta_p^2 = .12]$, but no interaction $(F < 1)$. An identical pattern of results, in terms of significance of statistical tests, was found when trial lengths 1–3 or 1–5 were analyzed as well, but the reported analysis was confined to trial lengths 1–4, because it maximized the number of data points while still including the majority of participants in each age group in the analysis. Moreover, a comparison
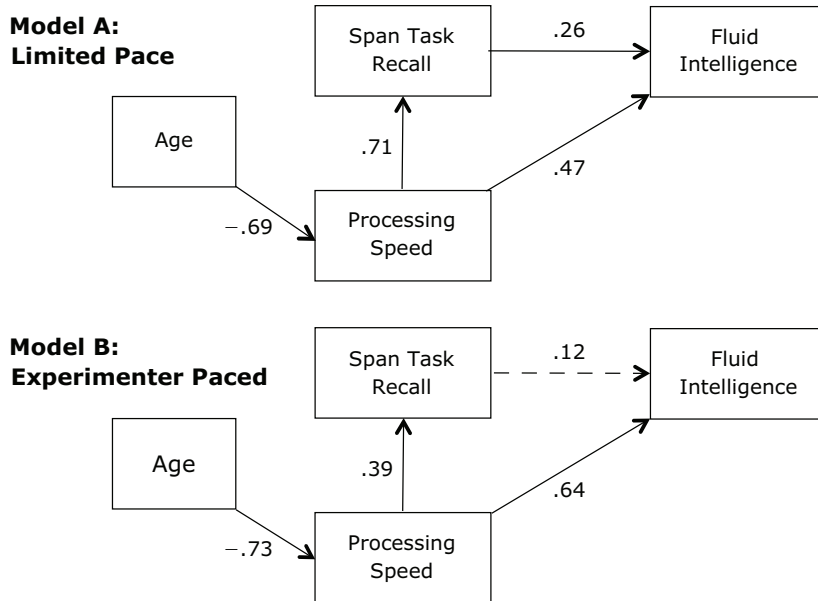
Figure 4. Path models indicating the relation between age, processing speed, span task recall, and fluid intelligence for the limited-pace and experimenter-paced versions of the Stroop span task. Solid lines indicate paths that are significant at $p < .05$; dotted lines indicate nonsignificant paths.

of the proportional differences in response times (older adult response times/younger adult response times) showed that they were similar across trial lengths 1–4 as well (1.16, 1.17, 1.15, and 1.12, respectively). Thus, it appears that both the younger and the older adults engage in proactive control of response times, slowing down for the first processing decision of longer trials lengths, presumably to devote more attentional resources to maintenance, thereby reducing cognitive load.

The Relation Between Response Times and Recall Performance on the Experimenter-Paced Stroop Span Task

Slowing of response times during the processing component of the span task should benefit recall performance in the experimenter-paced task by reducing cognitive load, thereby providing more time for maintenance-related processing. Thus, slower processing times should be positively correlated with recall. To test this idea, a series of path models were computed examining the relation of processing time to recall performance for the experimenter-paced task.

In order to assess whether suppression effects influenced the relationship between processing time and recall performance on the experimenter-paced task, two path models were computed. In the first model, the relationships between age and processing time and that between processing time and recall performance were considered without including general processing speed in the model. In the second model, age and individual differences in speed of processing were taken into account. Figure 6

shows both models. Model A includes a correlation between age and processing time and another between processing time and Stroop span recall. As is shown in the model, age was associated with slower processing times (.48), but the path between processing times and recall (.17) was not significant. Thus, this suggests that processing and maintenance functions in WM were independent. However, in Model B, in which age differences in the processing speed factor were controlled prior to assessing the relationship among processing speed, response times, and recall, there was a positive correlation between processing time and recall performance (.40), and processing speed also made a significant unique contribution to recall performance (.47). This comparison shows a classic suppression effect, in which a control variable (the processing speed factor) has a positive effect on the dependent variable through one path and a negative effect through another path (Cramer, 2003). In this case, general processing speed was negatively correlated with span task processing time but positively correlated with span task recall. Additionally, although the path between age and span task processing time was not significant, it was positive, indicating that the older adults were certainly not less strategic in their proactive metacognitive control strategy.

In summary, once age and individual differences in processing speed were accounted for, slower processing on the span task was positively correlated with span task recall, indicating that the metacognitive control of processing times allowed in the experimenter-paced task was beneficial to recall performance, supporting a resource-sharing explanation of WM.
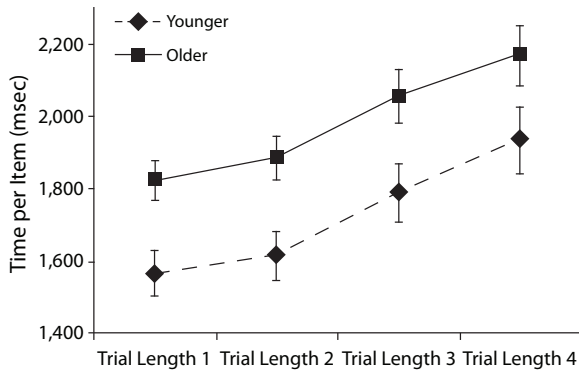
Figure 5. Average span task processing times for the experimenter-paced task as a function of age and list length for younger and older adults.

DISCUSSION

The present study provided three main findings of interest related to measuring WMC and examining age-related differences in WM: (1) Recall performance for the limited-pace administration method was more strongly related to higher level cognition than was recall performance for the experimenter-paced task. (2) Path models controlling for age-related differences in processing speed showed that much of the predictive power of the recall performance on the limited-pace task was mediated by processing speed, but there was still unique variance accounted for by span task recall for the limited-pace task, even after controlling for processing speed. Conversely, after controlling for processing speed for the experimenter-paced task, there was no unique variance in higher level cognition accounted for by recall in any of the outcome measures, although there were direct effects of processing speed on higher level cognition. (3) Recall was greater for the experimenter-paced task than for the limited-pace task for the older adults, but recall did not differ as a function of administration method for the younger adults. An analysis of processing times as a function of trial length for the experimenter-paced task indicated that as memory load (i.e., list length) increased, both the younger and the older adults slowed their processing in order to devote more time to maintenance. Moreover, path models showed that once age-related differences in general processing speed were accounted for, slower processing times on the experimenter-paced task were associated with greater recall.

The Effect of Administration Method on the Relation Between Span Task Recall and Higher Level Cognition

The finding that the limited-pace Stroop span task was more strongly correlated with higher level cognition than the experimenter-paced task is consistent with predictions from several resource-sharing explanations (Barrouillet et al., 2004; Engle & Kane, 2004; Lépine et al., 2005). According to Barrouillet et al.'s time-based resource-sharing model, computer pacing the task limits the amount of at-

tentional resources that can be devoted to maintenance, instead requiring that these limited attentional resources be devoted to the processing task. The finding that recall was poorer for the limited-pace task than for the experimenter-paced task for the older adults supports this idea that limited attentional resources were shifted away from maintenance and, instead, toward processing. The finding that the experimenter-paced task, which presumably accounts for the participants' level of ability on the processing component of the task to a great extent, was not as effective at predicting complex cognition raises questions regarding the effectiveness of using experimenter-pacing, or custom-pacing that provides considerable additional time for processing (see Unsworth et al., 2005), as a method of administration for complex span tasks.

The somewhat counterintuitive finding that the limited-pace task still accounted for variability in higher level cognition even though it was more sensitive to individual differences in processing speed than the experimenter-paced task suggests that the predictive advantage of the limited-pace task is not due to this task being more sensitive to measuring individual differences in processing speed. Furthermore, these results suggest that using a limited-pace method to administer a span task is a viable alternative to the typical experimenter-paced method and even has advantages over the typical experimenter-paced administration method.

One advantage of the limited-pace task, relative to the experimenter-paced task, was that it was a more effective measure of WMC, as assessed by the strength of the correlation of span task recall with higher level cognition. Of course, note that if a much slower pace had been used (e.g., 4 sec per processing decision), it would have reduced cognitive load considerably and would presumably make the task less effective than an experimenter-paced task. Thus, one issue to consider when using a limited-pace task is the choice of a particular processing pace that allows participants to complete the task but still reduces the overall amount of strategic processing that can be brought to bear during the task. Indeed, the design of the present study did not allow a decoupling of individual differences in processing speed and the amount of time available for processing, because the pace of the task was fixed. Future researchers could examine whether limited-pace tasks become less effective when additional time is available and whether the pace of the task interacts with individual differences in processing speed.

Metacognitive Monitoring and Control of Processing Time on Complex Span Tasks

The experimenter-paced task, although not as strongly correlated with higher level cognition, provided information about metcognitive control of processing times in WM that cannot be assessed in the limited-pace task. Reducing the control of processing times in the limited-pace task thereby reduces variability in strategic factors influencing span task performance, reducing the noise that strategy use can add to the correlations between span tasks and outcome measures (see Engle & Kane, 2004; Turley-Ames
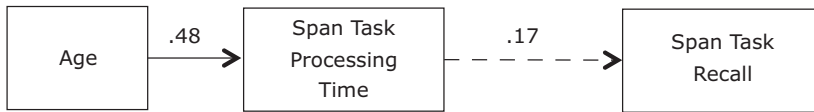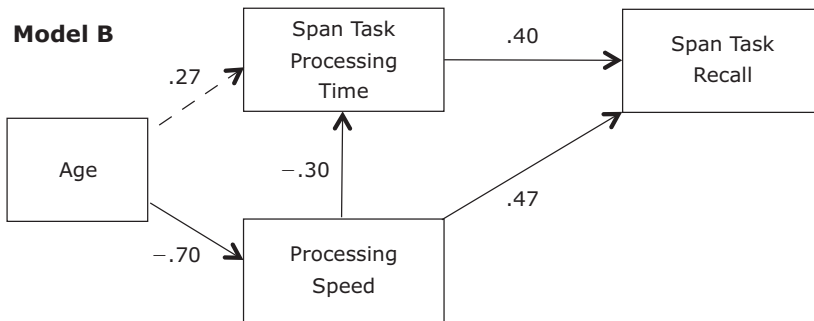
Model A

Model B

Figure 6. Path models indicating the relation between age, processing speed, span task processing time, and span task recall for the first trial of each trial length for the experimenter-paced version of the Stroop span task. Solid lines indicate paths that are significant at $p < .05$; dotted lines indicate nonsignificant paths.

& Whitfield, 2003). By contrast, the experimenter-paced span task (and self-paced span tasks; e.g., Friedman & Miyake, 2004; St. Clair-Thompson, 2007) provides fertile test beds for understanding the way in which participants use metacognitive control to deal with concurrent maintenance and processing demands in WM.

The examination of processing time as a function of trial list length (see Figure 5) provides on illustration of the role of metacognitive control on the experimenter-paced task. The younger and older adults slowed down for longer trial lengths, indicating that as the requirement for maintenance increased, the participants changed their attentional allocation to the processing task to deal with these increased task demands. Of course, it is not clear whether all of the participants engaged in this strategic slowing or the extent to which this strategy was successful across participants, but it is clear that allowing this strategic control during the task is one potential factor that leads to less effective measurement of WMC overall. In previous research, variability *within* a trial has been examined (Engle et al., 1992; Friedman & Miyake, 2004), under the assumption that if there was a limited pool of attentional resources to share between processing and maintenance, participants would slow down as memory load increased during performance within a particular trial. The idea behind this approach is that later in a trial, when more items need to be maintained, participants should slow their processing to deal with the additional attentional load. This examination of online attentional allocation during a trial has yielded conflicting results, with some studies supporting a processing-resources hypothesis and others failing to support it[2] (Engle et al., 1992; Friedman & Miyake, 2004; Towse, Hitch, & Hutton, 1998; Waters & Caplan,

1996). In the analysis reported here, I examined global response times as a function of trial length, which represents a different view of attentional control, in that it is assumed that participants strategically slow down processing for all processing items within a trial as trial length increases (see Figure 5).

One way to think about the difference between these two approaches to examining metacognitive control of response times is that slowing processing times during a trial requires reactive control (Braver et al., 2007), whereby online processing demands determine allocation of processing times, whereas slowing processing times globally for all trials of a given list length requires proactive control (Braver et al., 2007), whereby participants decide prior to a trial how to allocate processing time. Of course, it is not clear whether this control of response times is a conscious strategy or is simply a nonconscious strategy that is determined by task demands. This issue is one that would require additional research to understand more clearly.

The finding that the older adults were able to effectively allocate attentional resources in WM is consistent with other findings from episodic memory tasks showing that metamemorial processing is often unaffected by age (see Hertzog & Dunlosky, 2004, for a review). In order for attentional resources in WM to be effectively allocated, metacognitive control processes must operate effectively, and this appeared to be the case for both the younger and the older adults. It is also worth mentioning that even on complex span task trials with much more complex processing components (e.g., reading span, operation span), both younger and older adults use normative effective strategies, such as imagery or clustering, just as much

as younger adults, but both groups use normatively in-effective strategies (e.g., reading, repetition) most of the time (Bailey et al., 2009). Thus, it seems unlikely that the younger and older adults in the present study used different strategies on the experimenter-paced task during the additional time that they took to process longer trials, especially considering that age group did not interact with trial length for processing times (i.e., the two age groups increased their processing times in a similar fashion as trial lengths increased).

The finding that the participants slowed their process-ing as list length increased also reveals information about the participants' implicit theories of how WM operates. Specifically, these findings suggest that the participants adhered to a limited resource theory of WM. As was men-tioned previously, according to many limited resource the-ories of WM (Barrouillet et al., 2004; Engle et al., 1999; McCabe, 2008), there is a trade-off between processing and maintenance of information in WM, such that a single pool of attentional resources is shared between these two functions. Thus, limited resource theories predict that par-ticipants would benefit from slowing their processing to devote more attentional resources to maintenance, and this is exactly what happens during task performance. By con-trast, time-based forgetting theories of WM (e.g., Towse, Hitch, & Hutton, 1998, 2002) suggest that processing and maintenance are independent functions and that the amount of time that information has to be maintained is a potent factor determining forgetting. Thus, if the partici-pants' control of response times adhered to a time-based forgetting theory of WM, completing the processing com-ponent as quickly as possible in an effort to reduce the amount of time that information had to be maintained would be an effective strategy. However, the strategy that was effective in the present study (i.e., slowing processing to improve recall; see Figure 6) indicates that there was a trade-off of attentional resources between processing and maintenance and, furthermore, that the participants were influenced by knowledge of this trade-off (although this does not necessarily indicate explicit awareness of this trade-off) and allocated attention accordingly. This does not mean that time was unrelated to forgetting in the task or that the participants did not consider this factor, but the more potent factor influencing recall and the participants' control of processing time appeared to be the assumption that processing and maintenance relied on a single pool of resources.

Limitations of the Present Study

Another potential avenue of future research is whether the present results relating to limited pacing generalize to more typical complex span tasks (e.g., reading span). In both the present study and Lépine et al. (2005), which represent the only studies in which limited-pace span tasks have been compared with experimenter-paced tasks, fairly novel limited-pace verbal span tasks were used; therefore, it remains unclear whether the present results would generalize to more typical tasks. Like the reading span and computation span tasks, Stroop span includes an attention-demanding processing component that requires

a *yes–no* decision, the to-be-remembered information is integrated with the processing component, and processing is interleaved between the encoding of to-be-remembered items, followed by immediate serial recall of the to-be-remembered items. Thus, the primary difference between the Stroop span task and other complex span tasks is in the specific processing task employed, with reading span and computation span requiring more complex processing (e.g., reading comprehension, arithmetic computation) that takes slightly longer to complete. However, the use of a complex processing component during a span task does not appear to be necessary for that task to measure WMC effectively, and therefore, it should not be necessary that complex processing tasks be used for a span task to mea-sure WMC well (Barrouillet et al., 2004; Lépine et al., 2005). Therefore, because the Stroop span task is structur-ally identical to other span tasks that are commonly used and the task is predictive of higher level cognition (e.g., McCabe et al., 2005; McCabe & Smith, 2002; McCabe et al., 2007), it seems reasonable to assume that the effect of using a limited-pace method would generalize to more traditional tasks. However, this remains an issue that can only be addressed with further empirical research. Indeed, there may be factors (e.g., the amount of time afforded for rehearsal, the simplicity of the processing component) that can only be understood by conducting a study to address these issues directly. Thus, at this point, the present results support the idea that limited-pace tasks can be more effec-tive than experimenter-paced tasks, but the generality of this finding awaits further empirical support.

One issue that has not been considered widely with re-spect to the predictive power of span tasks is the complexity of the processing component of the tasks. Using a simpler processing component on complex span tasks might be preferred, because researchers can make a stronger claim that individual differences in central executive function drive the correlation between the task and higher level cog-nition, rather than more general individual differences in the ability to complete the complex processing component of a WM task. Alternately, more complex processing tasks might make it more difficult to implement effective strate-gies; therefore, it remains an open question as to whether processing task complexity has an important influence on the span-task–complex-cognition correlation.

A final issue to address is a consideration of the role of proactive interference on metacognitive control in the present study, which might be conceived of as an alterna-tive explanation to the resource-sharing explanation. Be-cause a stair-step method of presentation was used for the task in the present study (i.e., the shortest trials were pre-sented first, followed by the next longest trials, etc.), one might argue that it is the build-up of proactive interference (PI) during the task, rather than memory load per se, that causes participants to slow their processing as list length increases. However, the research on age differences in PI in complex span tasks indicates that although the PI in-duced by using a stair-step method significantly reduces older adults' recall performance on span tasks, it has lit-tle effect on younger adults' recall performance (Lustig, May, & Hasher, 2001; May, Hasher, & Kane, 1999; Rowe,

Hasher, & Turcotte, 2008). In fact, when an ascending stair-step method like the one used in the present study was compared with a descending stair-step method (i.e., the longest trials were presented first) in previous studies, older adults recalled more items from the descending condition, but younger adults actually recalled more items from the ascending condition. This has been attributed to practice effects outweighing PI effects for younger adults, but not for older adults (Rowe et al., 2008). The important point for the present purposes is that although PI induced by the stair-step method affects younger and older adults' recall performance differently, in the present study there were no age differences in slowing of processing times as list length and memory load increased. Thus, there is reason to prefer the resource-sharing (i.e., memory load) interpretation of these data, which suggests that a single mechanism common to both age groups is operating on performance, rather than a PI-based explanation. Note, too, that even if the participants slowed because of the buildup of PI, PI could also be viewed as a cause of increased memory load (i.e., larger trial lengths increase the pool of potential items to search in memory; Rowe et al., 2008) and, therefore, could be viewed as consistent, at least in principle, with the resource-sharing explanation.

Conclusions

Two general conclusions can be drawn from the study reported here, both of which support resource-sharing explanations of complex span task performance. First, using a limited-pace method to administer a complex span task was an effective method of reducing strategic processing during a span task relative to the experimenter-paced method, presumably because the former reduced individual differences in the strategic allocation of attention during the processing component of the span task. Thus, the limited-pace task provided a more effective estimate of WMC as shown by the stronger correlations with measures of higher level cognition. Second, using an experimenter-paced task allowed an examination of metacognitive control of response times during a complex span task and indicated that slowing of processing was adaptive, in terms of improving recall performance.

AUTHOR NOTE

Correspondence concerning this article should be addressed to D. P. McCabe, Department of Psychology, Colorado State University, Fort Collins, CO 80523-1876 (e-mail: david.mccabe@colostate.edu).

REFERENCES

ACKERMAN, P. L., BEIER, M. E., & BOYLE, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin*, **131**, 30-60.

BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.

BADDELEY, A. D. (2000). Short-term and working memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 77-92). New York: Oxford University Press.

BAILEY, H., DUNLOSKY, J., & HERTZOG, C. (2009). Does differential strategy use account for age-related differences in working memory performance? *Psychology & Aging*, **24**, 82-92.

BARROUILLET, P., BERNARDIN, S., & CAMOS, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, **133**, 83-100.

BARROUILLET, P., LÉPINE, R., & CAMOS, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review*, **15**, 528-534.

BLALOCK, L., & MCCABE, D. P. (2008, November). *Proactive interference and practice effects in working memory span task performance*. Poster presented at the 49th Annual Meeting of the Psychonomic Society, Chicago.

BRAVER, T. S., GRAY, J. R., & BURGESS, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 76-106). Oxford: Oxford University Press.

CONWAY, A. R. A., KANE, M. J., BUNTING, M. F., HAMBRICK, D. Z., WILHELM, O., & ENGLE, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, **12**, 769-786.

CRAMER, D. (2003). A cautionary tale of two statistics: Partial correlation and standardized partial regression. *Journal of Psychology*, **137**, 507-511.

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

DELIS, D. C., KRAMER, J. H., KAPLAN, E., & OBER, B. A. (2000). *The California Verbal Learning Test* (2nd ed.). San Antonio: Psychological Corporation.

DUNLOSKY, J., & CONNOR, L. (1997). Age-related differences in the allocation of study time account for age-related differences in memory performance. *Memory & Cognition*, **25**, 691-700.

DUNLOSKY, J., & HERTZOG, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multi-trial learning. *Journals of Gerontology*, **52B**, P178-P186.

EKSTROM, R. B., FRENCH, J. W., HARMAN, H. H., & DERMEN, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Services.

ENGLE, R. W., CANTOR, J., & CARULLO, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 972-992.

ENGLE, R. W., & KANE, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145-199). San Diego: Academic Press.

ENGLE, R. W., TUHOLSKI, S. W., LAUGHLIN, J. E., & CONWAY, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, **128**, 309-331.

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

FRIEDMAN, N. P., & MIYAKE, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory & Language*, **51**, 136-158.

GAULTNEY, J. F., KIPP, K., & KIRK, G. (2005). Utilization deficiency and working memory capacity in adult memory performance: Not just for children anymore. *Cognitive Development*, **20**, 205-213.

GLISKY, E. L., & KONG, L. L. (2008). Do young and older adults rely on different processes in source memory tasks? A neuropsychological study. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 809-822.

GLISKY, E. L., POLSTER, M. R., & ROUTHIEAUX, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology*, **9**, 229-235.

HEATON, R. K. (1993). *Wisconsin card sorting test: Computer Version 2, research edition*. Odessa, FL: Psychological Assessment Resources.

HERTZOG, C., & DUNLOSKY, J. (2004). Aging, metacognition, and cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 45, pp. 215-251). San Diego: Academic Press.

HU, L.-T., & BENTLER, P. (1995). Evaluating model fit. In R. H. Hoyle

(Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). London: Sage.

Kaakinen, J. K., & Hyönä, J. (2007). Strategy use in the reading span test: An analysis of eye movements and reported encoding strategies. *Memory*, **15**, 634-646.

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

Lemaire, P., & Lecacheur, M. (2002). Children's strategies in computational estimation. *Journal of Experimental Child Psychology*, **82**, 281-304.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review*, **12**, 165-170.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, **130**, 199-207.

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition*, **27**, 759-767.

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory & Language*, **58**, 480-494.

McCabe, D. P., Robertson, C. L., & Smith, A. D. (2005). Age differences in Stroop interference in working memory. *Journal of Clinical & Experimental Neuropsychology*, **27**, 633-644.

McCabe, D. P., Roediger, H. L., III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, **24**, 222-243.

McCabe, D. P., & Smith, A. D. (2002). The effect of warnings on false memories in younger and older adults. *Memory & Cognition*, **30**, 1065-1077.

McCabe, D. P., Smith, A. D., & Parks, C. M. (2007). Inadvertent plagiarism in young and older adults: The role of working memory capacity in reducing memory errors. *Memory & Cognition*, **35**, 231-241.

Miller, L. M. S., & Stine-Morrow, E. A. L. (1998). Aging and the effects of knowledge on on-line reading strategies. *Journals of Gerontology*, **53B**, P223-P233.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General*, **130**, 621-640.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 676-686.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, **134**, 368-387.

Oberauer, K., Süß, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity—Facets of a cognitive ability construct. *Personality & Individual Differences*, **29**, 1017-1045.

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

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology & Aging*, **11**, 621-637.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Section 4: The Advanced Progressive Matrices*. San Antonio: Harcourt Assessment.

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, **8**, 385-407.

Rowe, G., Hasher, L., & Turcotte, J. (2008). Age differences in visuospatial working memory. *Psychology & Aging*, **23**, 79-84.

St. Clair-Thompson, H. L. (2007). The influence of strategies upon relationships between working memory and cognitive skills. *Memory*, **15**, 353-365.

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

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, **132**, 566-594.

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, **27**, 763-776.

Salthouse, T. A., & Coon, V. E. (1994). Interpretation of differential deficits: The case of aging and mental arithmetic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1172-1182.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Touron, D. R., & Hertzog, C. (2004). Strategy shift affordance and strategy choice in young and older adults. *Memory & Cognition*, **32**, 298-310.

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory & Language*, **39**, 195-217.

Towse, J. N., Hitch, G. J., & Hutton, U. (2002). On the nature of the relationship between processing activity and item retention in children. *Journal of Experimental Child Psychology*, **82**, 156-184.

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

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex spans and their relation to higher order abilities. *Psychological Bulletin*, **133**, 1038-1066.

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

Van Petten, C., Plante, E., Davidson, P. S. R., Kuo, T. Y., Bajuscak, L., & Glisky, E. L. (2004). Memory and executive function in older adults: Relationships with temporal and prefrontal gray matter volumes and white matter hyperintensities. *Neuropsychologia*, **42**, 1313-1335.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *Quarterly Journal of Experimental Psychology*, **49A**, 51-75.

Wechsler, D. (1997a). *WAIS-III: Wechsler Adult Intelligence Scale–Third Edition*. San Antonio: Psychological Corporation.

Wechsler, D. (1997b). *WAIS-III/WMS-III technical manual*. San Antonio: Psychological Corporation.

Zachary, R. A. (1986). *Shipley Institute of Living Scale, Revised Manual*. Los Angeles: Western Psychological Services.

NOTES

1. I thank Klaus Oberauer for suggesting this approach to this analysis.

2. Note, too, that a within-trials analysis was not conducted in this study, because both congruent and incongruent Stroop color words were used for the processing items, and the incongruent words took substantially longer to process, which introduces tremendous variability into within-trials response times.