Journal of Experimental Psychology: General
2011, Vol. 140, No. 4, 573–585

0096-3445/11/$12.00 DOI: 10.1037/a0024587

The Stability of Working Memory: Do Previous Tasks Influence Complex Span?

M. Karl Healey and Lynn Hasher
University of Toronto and Rotman Research Institute

Elena Danilova
University of Toronto

Schmeichel (2007) reported that performing an initial task before completing a working memory span task can lower span scores and suggested that the effect was due to depleted cognitive resources. We showed that the detrimental effect of prior tasks depends on a match between the stimuli used in the span task and the preceding task. A task requiring participants to ignore words reduced performance on a subsequent word-based verbal span task but not on an arrow-based spatial span task. Ignoring arrows had the opposite pattern of effects: reducing performance on the spatial span task but not on the word-based span task. Finally, we showed that antisaccade, a nonverbal task that taxes domain-general processes implicated in working memory, did not influence subsequent performance of either a verbal or a spatial span task. Together these results suggest that while span is sensitive to prior tasks, that sensitivity does not stem from depleted resources.

Keywords: working memory, resource depletion, ego depletion, interference, cognitive control

Human cognition is limited, and psychologists have long been interested in understanding these limits. Perhaps the most influential example of this tradition is George Miller's famous "magical number seven" (1956); a proposed limit on the number of items that can be stored in short-term memory. While the details continue to be debated (see Cowan, 2001, for an extensive review), the notion that there is a fundamental limit on the amount of information that can be recalled after a short delay remains a staple of short-term and working memory theories.[1] Some have proposed that there is a focus of attention that is limited in the number of items it can hold (Cowan, 2001, 2010; Oberauer, 2002), others that time-based decay of information in working memory limits storage (e.g., Baddeley, Thomson, & Buchanan, 1975; Towse, Hitch, & Hutton, 1998), others that the ability to form and break arbitrary bindings is critical (Oberauer, 2005), and still others that working memory (Kane, Conway, Hambrick, & Engle, 2007) and both working and long-term memory (Hasher, Lustig, & Zacks, 2007; Hasher, Zacks, & May, 1999; Healey, Campbell, Hasher, & Ossher, 2010) are limited by the efficiency of attentional mechanisms that control various sources of interference; there are other views as well (see chapters in Conway, Jarrold, Kane, Miyake, & Towse, 2007). All of the views share the notion that some aspect or aspects of the cognitive system place an upper limit on successful recall.

One of the key findings of the last several decades of working memory research is that the severity of the limit varies among individuals and that these individual differences correlate with a wide range of other cognitive abilities, such as reading comprehension, problem solving, and reasoning (e.g., Conway et al., 2005; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; De Beni, Borella, & Carretti, 2007; Kyllonen, 1996). Indeed, it has been suggested that the same mechanisms that limit working memory also place limits on fluid intelligence, a construct thought to reflect domain-general thinking and reasoning abilities critical to complex cognition (Healey, Zacks, Hasher, & Helder, 2011; Kane et al., 2007). In much of this work, complex span tasks have been used to measure working memory limits. These tasks interleave presentation of a short list of to-be-remembered stimuli (e.g., between two and seven words) with a processing task (e.g., solving simple equations). The number of items participants can recall in these tasks is thought to reflect the underlying working memory limit (e.g., the efficiency of interference control, the size of the focus of attention, and so forth).

Thus, variations among individuals in working memory limits have broad implications for cognitive functioning. But within an individual, is the working memory limit fixed or variable? Most theories treat working memory limits, either implicitly or explicitly, as a relatively stable characteristic of a given individual. An invariant working memory makes intuitive sense; given that many

This article was published Online First July 18, 2011.

M. Karl Healey and Lynn Hasher, Department of Psychology, University of Toronto, Toronto, Ontario, Canada, and the Rotman Research Institute, Baycrest Centre, Toronto, Ontario, Canada; Elena Danilova, Department of Psychology, University of Toronto.

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada to Lynn Hasher and by an Ontario Graduate Scholarship to M. Karl Healy. We thank Karen L. Campbell and John A. E. Anderson for comments on the manuscript and Elizabeth Howard for her invaluable assistance conducting the studies.

Correspondence concerning this article should be addressed to M. Karl Healey who is now at the Department of Psychology, University of Pennsylvania, 3401 Walnut St., Rm. 303C, Philadelphia, PA, 19104, or to Lynn Hasher at the Department of Psychology, University of Toronto, 100 St. George St., Toronto, ON M5S 3G3, Canada. E-mail: healeym@sas.upenn.edu or hasher@psych.utoronto.ca

[1] There is a long and ongoing debate about whether there are distinct cognitive and neural systems that correspond to the labels *short-term memory* and *working memory* (see Jonides et al., 2008, for a recent review). Our view is that there is sufficient overlap between the two terms to allow us to use them interchangeably in the present context.

theories postulate a causal link between working memory and fluid intelligence (e.g., see chapters in Conway et al., 2007), a variable limit would seem to imply variable intelligence. Despite this profound implication, there is a paucity of work on within-individual variation. Therefore, we posed a critical, but largely unanswered, question: Is a person's working memory limit stable, or are the mechanisms that limit working memory labile?

Of course, experimental studies have identified numerous variables and manipulations that influence performance on span tasks. However, many manipulations alter how much information is successfully recalled without actually altering the underlying limiting mechanism. Consider chunking: An individual's span can be dramatically increased if prior knowledge permits linking several distinct items together into a single unit. Chunking does not directly increase the storage capacity of the system but rather reduces the number of distinct units that need to be stored, leaving the underlying limit unchanged. Proactive interference effects provide a similar example. Imagine that two groups are given identical lists of words to remember; for Group A, this is the first list they have learned, but Group B has already learned 10 similar lists. Group B would be expected to recall fewer words than Group A because of proactive interference from the prior lists (Underwood, 1957). Group A and Group B do not fundamentally differ in the ability to control interference; Group B simply has more interference to control (see Lustig & Hasher, 2006).

Changes in recall accuracy caused by factors such as ease of chunking and amount of proactive interference result not from changes to the memory system itself, but from external factors that alter the demands (e.g., the number of chunks to be stored or the amount of interference to be resolved) placed on an unchanged system. Therefore, we refer to such changes as *extrinsic*. By contrast, *intrinsic* changes reflect differences in the memory system itself. The clearest example is a brain lesion; an episodic memory task places exactly the same extrinsic demands on an individual with hippocampal amnesia and a healthy control, but the two differ in their ability to deal with those demands. Similarly, individual differences in memory ability likely reflect intrinsic differences. For example, while exposing different groups to different levels of interference on a memory task will create extrinsic differences in recall accuracy, even within a group some individuals will recall more than others. Such individual differences likely arise due to variation in some fundamental aspect of the memory system, such as the ability to control or regulate interference (Healey et al., 2010).

The distinction between the amount of interference on a given task and an individual's ability to control interference helps highlight the extrinsic/intrinsic distinction: Variations between tasks in the *amount* of interference produce extrinsic differences in accuracy but variations among individuals in the *ability to control* interference produce intrinsic differences. In general, if memory accuracy changes as characteristics of the task (memoranda, encoding/retrieval conditions) change, then the change may be extrinsic. If accuracy changes in the absence of any changes to the task conditions, it is likely an intrinsic difference.

The extrinsic/intrinsic distinction becomes critical when interpreting the relationship between span scores and other cognitive measures such as fluid intelligence: if two participants had different span scores because one happened to study easily chunked lists and another studied difficult-to-chunk lists (an extrinsic difference), one would not expect to see corresponding differences in their fluid intelligence scores. If, however, two participants had different span scores because one had more efficient attention-regulation processes (an intrinsic difference), one may indeed expect fluid intelligence differences. That is, extrinsic changes are unlikely to impact other aspects of cognition, whereas intrinsic changes are. The question is, are there any manipulations that produce intrinsic changes in an individual's working memory limits?

Most manipulations known to influence working memory span, such as the chunking and interference effects previously discussed, likely reflect extrinsic changes. For example, classic word length and phonological similarity effects are likely examples of extrinsic changes (e.g., word length determines how many words will fit within an unchanging phonological loop; Baddeley et al., 1975). Another example of extraneous changes is seen in the release from proactive interference paradigm (e.g., Wickens, 1970). Here, performance increases are due to the decreasing need to resolve interference, not increased efficiency of interference resolution. Mnemonists can remember tremendously long lists of specific types of stimuli, but there is evidence that this feat relies on strategies that circumvent, rather than alter, the underlying limiting mechanisms (Ericsson, Delaney, Weaver, & Mahadevan, 2004). Another example is the drop in memory performance associated with reminding people of their membership in a group stereotyped as poor performers on a given task (i.e., stereotype threat; Chasteen, Bhattacharyya, Horhota, Tam, & Hasher, 2005; Hess, Hinson, & Hodges, 2009; Inzlicht & Ben-Zeev, 2000; Spencer, Steele, & Quinn, 1999). Evidence suggests that participants experiencing stereotype threat devote attention to suppressing anxiety engendered by the threat, thereby increasing the extrinsic demands on the system and lowering the amount of attention they can devote to the memory task (Johns, Inzlicht, & Schmader, 2008; Schmader & Johns, 2003).

Are there examples of true intrinsic changes in working memory ability? Schmeichel (2007) reported a reduction in span scores that may be consistent with an intrinsic change. In an inintial task, Schmeichel had participants view a silent video. One corner of the video screen showed a women being interviewed; another corner showed a series of random words. Experimental participants were asked to ignore the words and focus only on the woman; control participants were given no instructions about the words. Participants then completed a complex span task. Those who had been asked to ignore the words had lower span scores than the control participants. Schmeichel argued that ignoring words requires the same executive control processes that limit working memory and that engaging those processes during the video task transiently reduces their efficiency, which leads to poor performance on the subsequent span task. If this interpretation is correct, it would constitute an intrinsic change in the working memory limit as it is not the demands on the system that change but rather the ability of the system to handle the demands.

Schmeichel's study was motivated by a theory of self-control from the social psychology literature. While the terminology is often very different, the concept of self-control is actually very similar to the concept of executive control in cognitive psychology (see Robinson, Schmeichel, & Inzlicht, 2010). Schmeichel's task required participants to control the tendency to look at the irrelevant words. Another common self-control task involves restraining any outward expression of emotion while watching highly emotional video clips (Baumeister, Bratslavsky, Muraven, & Tice, 1998). These self-control tasks are conceptually very similar to measures of executive control such as the Stroop and antisaccade tasks, which require restraint of prepotent responses (Butler, Zacks, & Henderson, 1999; Roberts, Hager, & Heron, 1994) and have been argued to measure the same underlying mechanisms that limit working memory (Kane, Conway, Bleckley, & Engle, 2001; Miyake et al., 2000).

Thus, working memory and self-control may both be limited by closely related mechanisms. Unlike most theories of working memory, however, theories of self-control explicitly state that limits vary within individuals. Specifically, completing one self-control task is said to transiently reduce the effectiveness of self-control mechanisms, thereby impairing an individual's ability to perform a second self-control task immediately afterward. The negative impact of one task on a subsequent task has been labeled *resource depletion* (Baumeister et al., 1998; Baumeister, Vohs, & Tice, 2007). We prefer to avoid the term *depletion* when referring to the basic effect of reduced performance on the second task, as it prematurely assumes the explanation, in favor of the more theory-neutral, if less euphonious, term *subsequent-task effects*. We use the term *resource depletion* to refer specifically to the proposed resource-based explanation of the effect.

The possibility that working memory is so labile that simply being asked to ignore words for a few minutes can dramatically reduce its efficiency suggests important consequences for current theories of working memory limits, especially given the close and possibly causal link between working memory and fluid intelligence. It is therefore critical to determine whether any subsequent-task effects on complex span truly reflect resource depletion (i.e., intrinsic changes to the system).

Experiment 1

Schmeichel (2007) reported two other studies that could be considered conceptual replications of the study described previously. However, the magnitudes of the subsequent task effects observed were sometimes small or nonsignificant and varied both across studies and with the method used to calculate span score. Given the potential implications of working memory depletion, any ambiguity in the results suggests that replication of the basic finding is critical. We provide such a replication in Experiment 1 using the most psychometrically valid span administration and scoring procedures (Conway et al. 2005). We explore alternative, nondepletion-based explanations for the effect in subsequent experiments.

Method

Participants. For each experiment reported here, we recruited a unique sample of undergraduates from introductory psychology classes at the University of Toronto. All participants had spoken English since early childhood and were compensated with either course credit or $10. Thirty-eight students participated in Experiment 1, 19 in each condition.

Materials and procedure. All of the experiments followed the same basic procedure: Participants first completed either a high-demand cognitive task or a low-demand version of the same task. They then filled out a brief mood scale and finally completed a complex span task. In Experiment 1, the initial task was to watch a video of a woman being interviewed and to either ignore (high-demand) or not ignore (low-demand) words that appeared in the lower right corner of the screen; the complex span task was operation span.

Video task. The video task was modeled after the one used by Schmeichel (2007). Figure 1A shows a representative frame of the video. Participants in both conditions were told that they would be watching a silent 6-min video of a young woman being interviewed. They were told that the study was about nonverbal personality assessments and that they would be asked to make person-perception judgments about the woman. The top left portion of the video frame showed a woman sitting at a table having a conversation with an off-camera individual. The bottom right portion of

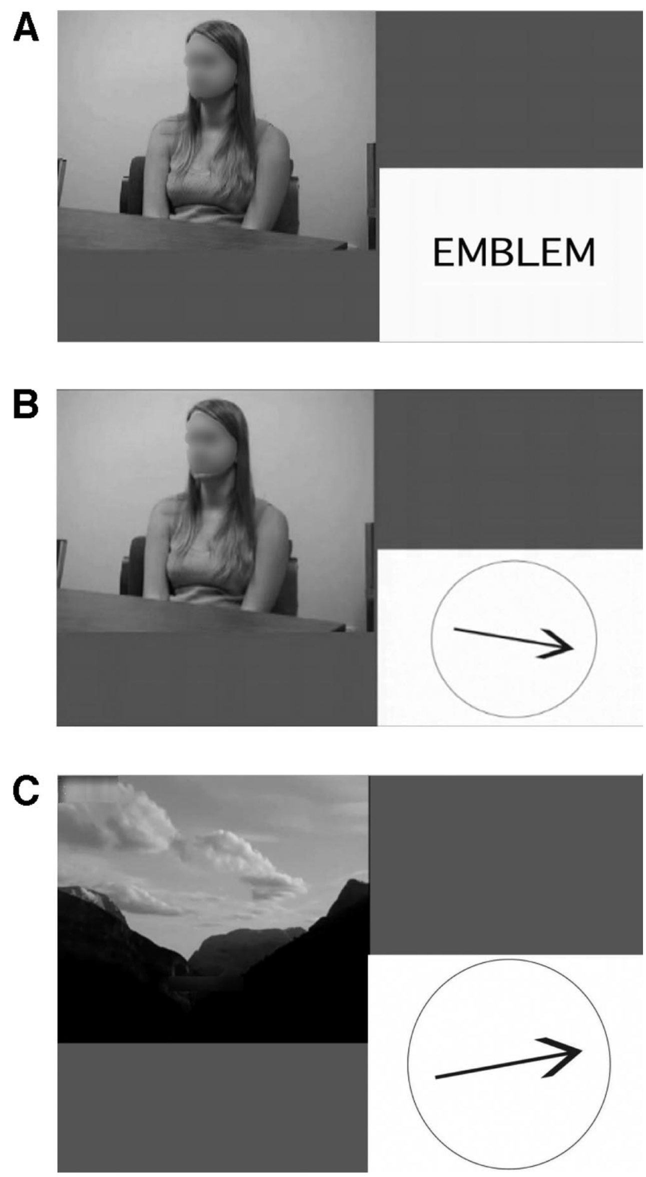

Figure 1. Example frames of the videos used in the various experiments. The examples are converted to black and white, but the original videos viewed by participants were in color. In addition, the woman's face has been blurred here to preserve her privacy but was visible in the original videos. A: Experiments 1 and 2. B: Experiment 3. C: Experiment 4.

the screen presented common one-syllable words in black font against a white background at a rate of 6 words per second.

High-demand participants were given the following additional instructions:

> Please focus on the woman throughout the video, and do not read or look at any words that may appear on the screen. If you find yourself looking at the words, please redirect your gaze to the woman as quickly as possible.

Low-demand participants were not given any specific instructions about the words. That is, high-demand, but not low-demand, participants were required to control the prepotent tendency of stimulus onsets to capture attention as well as the tendency to read words once their attention had been captured.

Mood scale. As is common in the depletion literature, we administered the Brief Mood Inspection Survey (BMIS; Mayer & Gaschke, 1988) after the video task to determine whether any differences in working memory scores were related to mood. On the BMIS, mood is measured by having participants rate the extent to which they are currently feeling eight pleasant emotions (e. g., “content,” “peppy,” “happy”) and eight unpleasant emotions (e. g., “jittery,” “nervous,” “gloomy”) on a 7-point scale (1 = *definitely do not feel*; 7 = *definitely feel*). To calculate a final pleasant–unpleasant mood score, experimenters subtract the summed rating for the unpleasant emotions from the summed rating for the pleasant emotions, yielding a score ranging from 49 (pleasant) to −49 (unpleasant).

Operation span. The target words were common one- or two-syllable nouns with between four and seven letters. There was no overlap between the words used in the video task and the span target words. The equations—for example, $(2 \times 2) + 3 = 7$—were composed of two numbers that were multiplied or divided, a third number that was added to or subtracted from the result, and a final answer. The task consisted of 12 operation span trials, three trials at each set size (the number of equation/word pairs per trial) ranging from two to five. The same random trial order was used for each participant.

On each trial, an equation and a target word were presented together on the same screen—for example, $(2 \times 2) + 3 = 7$ SMOKE—and the participant read the equation aloud, verified its accuracy by saying “yes” or “no,” and then read the target word aloud. As soon as the participant finished reading the target word, the experimenter advanced the program to the next equation/word pair. Participants were instructed to begin reading the equations as soon as they appeared on screen and were reminded of this instruction whenever the experimenter noticed the participant pausing before beginning to read.

Once all of the equation/word pairs for a trial had been presented, the participant was prompted to recall all of the target words from that trial by saying them aloud in the same order in which they were presented. Partial credit scoring (Conway et al., 2005) was used: An item was considered recalled if its output position matched its presentation position, and credit was given for a correctly recalled item even if other items in the trial were not recalled. For all experiments, we report span scores as a percentage of the targets recalled.

Data screening and analysis. In each experiment, we identified and excluded outliers on the basis of a 1.5 times interquartile range rule (Tukey, 1977).[2] The number of excluded participants for each experiment is reported in Table 1; the sample sizes reported in the Method sections reflect final sample sizes after exclusions. We report Cohen’s *d* as a measure of effect size.

As mentioned earlier, in most studies in the depletion literature, mood scores were examined to rule out the possibility that mood differences between the cognitive demand conditions influence subsequent task performance. To streamline the Result sections of individual experiments, we report and analyze mood data from all of the experiments here. See Table 1 for the mood scores from each experiment (mood data were lost for three participants due to experimenter error). The magnitude and direction of the difference in mood scores between conditions fluctuated from study to study and was significant only in Experiment 5. The critical question is not so much whether mood differed between demand conditions but whether mood differences accounted for any differences in span scores. To test this possibility, we converted span scores to *z* scores within each study and then correlated the *z* scores with mood scores, collapsed across studies. Mood did not correlate with span *z* scores for the high-demand, $r(133) = -.08$, or low-demand, $r(136) = .04$, conditions, nor when it was collapsed across conditions, $r(269) = -.02$), indicating that mood did not influence span scores. As a final check for any influence of mood, we retested all of the high-demand/low-demand comparisons reported in the article with mood score included as a covariate. Mood score was not a significant covariate in any of these analyses, further confirming that mood does not drive subsequent task effects. Therefore, for brevity and clarity, we do not discuss mood data in subsequent experiments and report standard *t* tests and analyses of variance (ANOVAs) rather than analyses of covariance.

Results and Discussion

High-demand participants recalled an average of 54.3% ($SD = 12.2$) of the words, substantially fewer words than the 69.8% ($SD = 11.4$) recalled by low-demand participants, $t(36) = 4.06$, $p < .001$, $d = 1.35$, replicating the Schmeichel (2007) finding. See Table 2 for a summary of the difference between the high- and low-demand conditions for this and all experiments. Participants in the two conditions were exposed to exactly the same materials during the experiment, the only difference being that the high-demand participants were told to avoid looking at the words during the video task. This small change in instructions was sufficient to reduce average span accuracy in the high-demand condition by 15.6% relative to the low-demand condition.

Given that complex span scores are widely viewed as a reliable and important predictor of fluid intelligence (Kane et al., 2007), a drop in recall accuracy of 15.6% is quite startling. The interpretation of this reduction is therefore of critical importance. Does it

[2] To check for outliers, we sorted span scores within each condition into quartiles, calculated the distance between the upper and lower quartile boundaries (i.e., the interquartile range; IQR), and identified any value either above the upper quartile boundary or below the lower quartile boundary by more than 1.5 times the IQR as an outlier and excluded it from analysis. When applied to a normal distribution, the 1.5 times the IQR rule excludes approximately the most extreme 1% of observations. This procedure eliminated between zero and four participants per experiment. Including these participants did not change the pattern of data, altering the low-demand minus high-demand condition span score differences (see Table 2) by less than 0.04 in all cases.

Table 1
Number of Outliers and Mean and Standard Deviation of Mood Scores by Cognitive Demand Condition

Experiment	No. of outliers		Mood score			
			High-demand		Low-demand	
	High-demand	Low-demand	*M*	*SD*	*M*	*SD*
1	2	1	8.33	9.70	3.89	10.52
2	0	0	8.25	12.16	4.72	12.17
3	0	1	2.21	15.27	6.61	11.23
4	4	0	4.95	12.73	5.65	11.38
5	0	0	3.07	8.24	11.13	10.54
6	2	0	3.30	10.03	0.08	9.65

reflect a genuine reduction in working memory capacity, as suggested by the depletion perspective, which under some views of the relation between working memory capacity and fluid intelligence would imply that asking someone to ignore words for 6 min induces a general cognitive impairment? Or is there an alternative explanation for the present finding?

Experiment 2

The effect in Experiment 1 seems like a good candidate for an intrinsic reduction in working memory: Both groups completed exactly the same span task using exactly the same materials under what seem to be identical encoding/retrieval conditions. That is, the change appears to have occurred in the absence of any extrinsic alterations to the task conditions. However, consider again proactive interference effects, which we have argued constitute extrinsic changes. If two people study exactly the same list of words, a person who has already studied several lists will recall fewer words from the final list than someone who has not studied other lists (Underwood, 1957). The two people do not have fundamentally different memory systems; their systems simply have to contend with different levels of interference. That is, lingering effects of past stimuli can alter memory performance without fundamentally altering the memory system (Lustig, Hasher, & Zacks, 2007; Postman & Underwood, 1973). Could lingering extrinsic effects account for the results of Experiment 1?

In Experiment 1, the video task presented a list of words. The operation span task also presented lists of words. Could this similarity in material between the initial video task and the subsequent span task have produced the reduced span scores observed in Experiment 1? If so, replacing the distracting words in the video task with new stimuli that are not similar to the words used in operation span should eliminate the effect. Later we will discuss possible mechanisms through which ignoring stimuli in one task could impair memory for similar stimuli in a later task, but for now we focus on using stimulus-similarity to test a key prediction of depletion theory. Specifically, depletion theory predicts that subsequent task effects should be insensitive to changes in the particular stimuli used in the two tasks, provided that the underlying resource demands are maintained. Indeed, much of the depletion literature has focused on showing that depletion effects can be found with a range of superficially dissimilar tasks, provided that each requires self-control.

In this study, we pitted the discrepant predictions of the depletion and stimulus-similarity accounts against each other by removing the similarity among stimuli while maintaining the need for self-control. We did so by replacing the word-based operation span task with rotation span, a nonverbal complex span task in which arrows, rather than words, are used as the memoranda. If similarity between the words in the video task and the words in operation span produced the subsequent-task effect in Experiment 1, rotation span should show no subsequent task effect. If, however, the effect in Experiment 1 was due to depleted control resources, a similar effect should be seen for rotation span, which taps the same executive control processes as operation span (Kane et al., 2004).

Given that we are predicting a null effect, it is important to have sufficient power to detect a nonzero effect if, contrary to our hypothesis, one does exist. A sample size of 13 participants per condition gives a 1-β power value of .95 to detect an effect of the

Table 2
Subsequent-Task Effects (Low-Demand Span Recall Minus High-Demand Span Recall) and the Associated Effect Size for Each Experiment

Experiment	Task 1	Task 2	Low–high difference (%)	Effect size (*d*)
1	Interview/words	Operation span	15.55*	1.35
2	Interview/words	Rotation span	2.00	0.12
3	Interview/arrows	Rotation span	12.64*	0.73
4	Clouds/arrows	Operation span	−1.60	−0.13
5	Antisaccade	Operation span	5.79	0.37
5	Antisaccade	Symmetry span	−5.09	−0.35

* $p < .05$.

size observed in Experiment 1 with $\alpha = .05$ and a one-tailed test. We exceeded that minimum sample size and tested at least 22 participants per condition here and in subsequent experiments in which we predicted a null effect to ensure that we had adequate power to detect even smaller effects, should they exist.

Method

Participants. Twenty-five undergraduate students participated in the high-demand condition, and 25 participated in the low-demand condition.

Materials and procedure. The video task was identical to the one used in Experiment 1; that is, words served as the to-be ignored (or not-to-be-ignored) stimuli. After the video, participants completed the mood scale followed immediately by the rotation span task.

Our version of rotation span was based on the one used by Miyake, Friedman, Rettinger, Shah, & Hegarty (2001) and is designed to be a spatial analog of verbal complex span tasks. On each trial, participants had to remember the orientation of several arrows, the presentation of which was interleaved with a spatial processing task (mental rotation). For the mental rotation task, an uppercase letter (*G*, *R*, or *F*) was presented in the center of a roughly circular "blob" shape. The letter was rotated from its usual upright position by some multiple of 45° and was either a normal letter or a mirror-image version of a normal letter (i.e., rotated 180° around the vertical axis). The participants' task was to indicate if the letter was a normal or a mirror-image version by saying "yes" (normal) or "no" (mirror image). After the participant responded to the letter, an arrow appeared, originating from the center of the blob and pointing in one of eight directions (i.e., 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°). Participants were to remember the orientation of each arrow. A trial consisted of between three and six letter/arrow pairs. There were two trials at each set size.

Once all of the letter/arrow pairs for a trial had been presented, the participants were given a response sheet with six empty blob shapes and were asked to recall the orientations of the arrows from that trial by drawing them in the same order in which they were presented. We scored recall using the same partial credit scheme used for operation span.

Results and Discussion

Contrary to the findings of Experiment 1, there was no difference in rotation span scores between the high-demand (M = 67.4%, $SD = 19.8$) and low-demand (M = 69.4%, $SD = 15.1$) conditions, $t(48) = 0.40$, $d = 0.12$. That is, whereas asking participants to ignore words before completing operation span dramatically lowered their span scores, asking them to ignore words before completing rotation span had no effect. The absence of any effect is inconsistent with a depletion account of Experiment 1. For a depletion account to remain viable, one would have to argue that operation span requires substantial self-control, but rotation span requires very little. The results are, however, consistent with a stimulus-similarity account: ignoring words prior to memorizing words impairs memory, but ignoring words prior to memorizing arrow directions does not impair memory.

Comparison of Experiments 1 and 2

The stimulus-similarity account predicts that the effect of ignoring words in the initial task should interact with the type of memoranda in the span task, whereas the depletion account predicts subsequent task effects regardless of memoranda. As a direct test of these two hypotheses, we combined the results of Experiments 1 and 2 and tested for a demand condition (low versus high) by span memoranda (verbal in Experiment 1 versus spatial in Experiment 2) interaction. Differences between operation span and rotation span (e.g., they use different types of stimuli and different numbers of stimuli) make it difficult to interpret the main effect of span type. We therefore converted recall accuracy to z scores within each experiment, eliminating the effect of span type. An ANOVA performed on these z scores indicates that the impact of ignoring/not ignoring words did indeed interact with type of span task, $F(1, 84) = 6.00$, $p = .016$, with the effect of ignoring being greater for the word-based operation span than the arrow-based rotation span. This interaction supports the conclusion that stimulus similarity drives the subsequent task effect observed in Experiment 1 and supplements the unambiguous results within each experiment: ignoring words impaired memory for words in Experiment 1 but ignoring words did not impair memory for arrow directions in Experiment 2.

Experiment 3

A clear prediction of the stimulus-similarity account is that while ignoring *words* prior to memorizing arrows had no impact on memory in Experiment 2, ignoring *arrows* before memorizing arrows should impact memory. That is, replacing the words in the video with arrows at various rotation angles should reduce rotation span scores. In contrast, under a depletion account, there is no clear reason to expect that ignoring arrows during a video should be more detrimental to subsequent memory span performance than ignoring words.

Method

Participants. The reported data are from 19 students in the high-demand condition and 18 in the low-demand condition.

Materials and procedure. This experiment was identical to Experiment 2 except for the to-be-ignored information in the video. The video presented the same woman being interviewed in the upper left corner. However, instead of words, a series of arrows was presented in the lower right corner (see Figure 1B for an example frame). The arrows were presented within a circular outline and randomly cycled through eight directions (i.e., 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°). Note that these are the same eight orientations used for the to-be-remembered arrows in the rotation span task. Under a depletion account, the overlap in stimuli should have no effect. The circle outline and arrows were presented in black against a white background at a rate of six arrows per second (the same rate used for words in the original video). The physical size of the arrows on screen was approximately equal to that of the words in the original video. As with the word-distraction video, the arrow-distraction video was 6 min long and was played without sound. Instructions in the high- and low-demand conditions were identical to those given in Experiment 1, except that references to "words" were changed to references to "arrows." After viewing the video, participants completed the mood scale and then rotation span using exactly the same stimuli and procedure as in Experiment 2.

Results

Rotation span showed a clear subsequent-task effect: high-demand participants recalled an average of only 52.3% (SD = 20.3) of the arrows, whereas low-demand participants recalled an average of 65.0% (SD = 14.5), $t(35) = 2.17$, $p < .05$, $d = 0.73$.

Experiment 4

The pattern of data reported so far is consistent with a stimulus-similarity account: ignoring words led to subsequent impairment on a word-based span task (Experiment 1) but not on an arrow-based span task (Experiment 2), whereas ignoring arrows led to impairment on an arrow-based span task (Experiment 3). To complete the picture, we tested the impact of ignoring arrows on the word-based operation span task. The stimulus-similarity prediction is obvious: ignoring arrows should have no impact on operation span. The depletion account makes the opposite prediction: ignoring arrows should impair operation span performance.

Method

Participants. We report data from 27 participants in the high-demand condition and 22 in the low-demand condition.

Materials and procedure. The original interview video with distracting words contains two potential sources of verbal interference: the distracting words and the speaking woman. While there is no sound, the woman's mouth movements, facial expressions, and hand gestures are clearly visible. Together these non-auditory cues may be sufficient for some participants to understand some of what is being said, creating a potential source of verbal interference. To avoid confounding the verbal nature of the to-be-ignored material with the verbal nature of the to-be-attended material, we created a new video. In place of the woman, the new video showed a 6-min time-lapse film of clouds moving across a mountain range (see Figure 1C). At several points in the video, the camera angle changes to give a different perspective on the range. The video shows purely landscape scenes with no people and no objects other than mountains, trees, and clouds. The cloud video occupied the same area in the upper left proportion of the screen as did the interview scene, and the same distracting arrow display used in Experiment 3 occupied the lower right corner. All participants were read the following cover story before watching the video:

> We are interested in how people make nonverbal assessments of visual scenes. You will be watching a 6-min video without sound of a variety of natural landscapes; later you will be asked to answer some questions about the scenes.

High-demand participants were given the following instructions about the arrows:

> Please focus on the landscapes throughout the video and do not look at any arrows that may appear on the screen. If you find yourself looking at the arrows, please redirect your gaze to the landscapes as quickly as possible.

Low-demand participants were given no instructions about the arrows.

The new video captures what, under a depletion account, are the critical features of the original: a relatively boring scene presented along with potentially attention-capturing peripheral stimuli. Indeed, because the cloud video lacks the potentially interesting social aspects of the interview video, it may actually require more cognitive control for participants to sustain their attention on the cloud video than on the interview video. After watching the cloud video, participants completed the mood scale and the same version of operation span used in Experiment 1.

Results and Discussion

Instructions to ignore the arrows produced no deficit in operation span scores, $t(47) = -0.43$, $d = -0.13$; indeed the high-demand participants (M = 57.3%, SD = 11.7) numerically outperformed the low-demand participants (M = 55.6%, SD = 14.0). Experiments 1–4 suggest that subsequent-task effects impact span tasks only when there is a match between the stimuli ignored in Task 1 and those remembered in Task 2. This pattern is predicted by a stimulus-similarity account but is inconsistent with depletion theory.

Comparison of Experiments 3 and 4

As with Experiments 1 and 2, we tested for a demand condition (low versus high) by span material (verbal vs. spatial) interaction by combining the results of Experiments 3 and 4. We again converted recall accuracy to z scores within each experiment, eliminating the effect of span type. The interaction fell just short of traditional significance levels, $F(1, 82) = 3.49$, $p = .065$. We stress that because this test depends on a cross-task and cross-experiment comparison, it should be considered as a supplement to the main finding within each experiment: ignoring arrows impaired memory for arrows (Experiment 3) but ignoring arrows did not impair memory for words (Experiment 4), a perfect reversal of the pattern found with ignoring words in Experiments 1 and 2. This pattern of results is exactly what one would predict if stimulus similarity produces subsequent task effects but makes little sense under a depletion account.

Discussion of Experiments 1–4

Experiments 1–4 showed that with the video task as an initial task, subsequent-task effects emerged only when the stimuli participants ignored in the initial task were similar to the stimuli they had to remember in the span task, even though all versions of the video task should have tapped the same control processes implicated in complex span. While this pattern rules out resource depletion as an explanation, it does not point clearly to an alternative account. Our main aim in this article is to test depletion theory. However, before moving on to a more direct test of the depletion hypothesis in Experiments 5 and 6, we offer some speculation on how ignoring stimuli at one point in time could impair later memory for similar stimuli.

Considerable evidence shows that young university students (but not older adults) are quite adept at dealing with distraction from both external and internal sources (e.g., memory interference). However, this ability to ignore distraction generally comes at the cost of reduced access to the ignored information at a later point (Campbell, Hasher, & Thomas, 2010; Rowe, Valderrama, Hasher, & Lenartowicz, 2006; see Healey, Campbell, & Hasher, 2008, for a review). For example, when participants in a recent study resolved competition between two similar words (e.g., *allergy* and *analogy*) to solve a word fragment (e.g., *a _ l _ _ gy*), the rejected word was subsequently less accessible in memory (Healey et al., 2010; see Anderson, Bjork, & Bjork, 1994; Anderson & Spellman, 1995, for related findings). Similar effects have been found when participants ignore externally presented distraction (e.g., the negative priming effect; see May, Kane, & Hasher, 1995, for a critical review).

Could a similar reduction in accessibility account for the findings in the present study? Could ignoring words (or arrows) during the video task reduce accessibility of other words (or other arrows) during the span task? If access to an entire class of stimuli, such as words, were impaired in this manner, it could lead to difficulty in encoding and thus difficulty in remembering any words on a subsequent span task. One complication is that in previous studies showing reduced access, usually only the specific words that were previously rejected or ignored were subject to reduced accessibility, not words in general. However, as these earlier studies used a mix of relevant and irrelevant words, it would be counterproductive to reduce access to all words. If, however, all words are declared irrelevant, as they are in the video task, it may be possible and efficient for the system to reduce access to words (or arrows in the arrow version of the task) in general. Of course, much work would be needed to determine if access to an entire class of stimuli could be reduced in this manner (see Postman, Stark, & Fraser, 1968, for evidence that reduced availability of entire classes of responses contributes to retroactive interference in paired associate learning).

Neuroimaging evidence suggests a subtly different way in which ignoring stimuli at one point could impair later processing of similar stimuli. Gazzeley, Cooney, Rissman, and D'Esposito (2005) showed participants a series of pictures; some were of faces, others of scenes (e.g., a sunset). When told to ignore the scenes and memorize the faces, young adults showed decreased activity in the parahippocampal place area, a region involved in processing scenes. That is, instructions to ignore a particular type of stimuli led to reduced activity in brain areas responsible for processing that type of stimuli. In such studies, reduced processing is an appropriate response to the task demands. If, however, such reduced processing continues even when the task changes, it could contribute to stimulus similarity dependent subsequent task effects. Consider the current Experiments 1 and 2 as examples: if participants suppress processing of words during the video task and inappropriately continue to do so once the task ends, it would likely disrupt encoding of words during operation span in Experiment 1, but to the extent that the suppression of processing is stimulus specific, it should not disrupt encoding of arrows during rotation span in Experiment 2. Supporting the possibility that reduced processing could continue across task boundaries, recent studies have shown that brain activity (Barnes, Bullmore, & Suckling, 2009) and patterns of functional connectivity (Grigg & Grady, 2010) related to one task can take considerable time to dissipate after the task ends. Note that under this interpretation, subsequent task effects occur because inappropriately strong cognitive control, in the form of suppressing processing of a category of stimuli, is exerted during the subsequent task. This suggestion is almost the opposite of the depletion theory claim that control is weakened on the subsequent task.

Clearly, many issues need to be resolved, and many other possible explanations explored, before these speculations can be developed into a coherent account of the subsequent task effects observed here. For now, we simply make the empirical observation that ignoring stimuli in the video task impacts subsequent span performance only when the ignored material is similar to the memoranda in the span task, an observation that is inconsistent with depletion theory. The last two experiments provide a final, direct test of depletion theory, as applied to working memory.

Experiment 5

In Experiments 1–4 we tested depletion theory by varying the type of stimuli participants ignored in the video task and concluded that the observed subsequent task effects were due to stimulus-similarity, not resource depletion. However, apart from demonstrations of subsequent task effects, there is actually very little research establishing that the video task places demands on the attentional control abilities that subserve complex span performance (Bunting, 2006; Hasher et al., 1999, 2007; Kane et al., 2001, 2007; Lustig, May, & Hasher, 2001; May, Hasher, & Kane, 1999; Rowe, Hasher, & Turcotte, 2010). Thus, it is possible that we failed to find true depletion effects not because depletion theory is flawed, but because the video task does not actually tap attentional control. Therefore, in Experiments 5 and 6, we replaced the video task with the antisaccade task, a widely used attention control task (Butler et al., 1999; Roberts et al., 1994) known to be related to complex span (Kane et al., 2001; Unsworth, Schrock, & Engle, 2004). In Experiment 5, antisaccade performance is followed by operation span; in Experiment 6, it is followed by a spatial complex span task. Using a well-validated attention control task as the initial task provides a fair and direct test of depletion theory.

In the antisaccade task, participants must identify a target that is presented very briefly on one side of the screen. However, before the target is presented, a misleading cue is presented on the opposite side of the screen. Given the rapid target presentation rate, participants are likely to miss the target if the misleading cue captures their attention. Therefore, rapid and accurate target identification requires the participants to exert attentional control to avoid looking at the distracting cue. The task is well validated as a measure of attentional control. Antisaccade is also conceptually similar to the video task used in the preceding experiments as both involve ignoring salient peripheral stimuli. Moreover, while antisaccade shares deep similarities with complex span (i.e., in the need to control attention), at a superficial level the two tasks are very dissimilar both in the types of stimuli they use and in the sorts of actions participants perform on them, so there is no reason to expect subsequent task effects due to stimulus similarity.[3] We tested for subsequent-task effects by having participants complete the high-demand antisaccade task followed by operation span. For a low-demand condition, a separate group of participants completed a prosaccade task (in which the cue accurately predicts the target location and thereby reduces the participants' need to control attention), followed by operation span.

Method

Participants. Twenty-five undergraduate students were tested in the high-demand condition. and 22 were tested in the low-demand condition.

Materials and procedure. The antisaccade task was based on the version used by Kane et al. (2001). On each trial, two boxes were displayed, one on either side of the screen. A flashing cue (an equal sign) appeared briefly in one box, followed by the brief, masked presentation of a target letter (*B*, *P*, or *R*). In the high-demand antisaccade task, the cue and target appeared in different boxes. Thus, performing the task well requires participants to restrain the tendency to look toward the cue. In the low-demand control condition, the cue and target appeared in the same box (prosaccade), greatly reducing the participants' need to control attention. When the goal is to measure individual differences in attention control, participants complete both the pro- and antisaccade versions; however, given that our goal was not to measure but to tax (or not tax) attention control, participants completed only one version.

The sequence of events within a trial was as follows: the word *READY?* was presented at the center of the screen, and participants pressed the space bar to begin the trial. The screen then went blank for 400 ms, after which a central fixation cross was displayed along with two empty boxes (a white frame against the black background), one on either side of fixation. The fixation remained onscreen for an unpredictable interval (randomly selected from 200, 600, 1,000, 1,400, 1,800, or 2,200 ms), after which it disappeared leaving only the two boxes. Fifty milliseconds after fixation offset, a cue appeared in one of the boxes for 100 ms, disappeared for 50 ms, and then reappeared for 100 ms, producing an attention-capturing, flashing effect (Roberts et al., 1994). Then 50 ms after the second cue offset, the target appeared in one of the boxes for 100 ms followed immediately by a mask which consisted of the letter *H* displayed for 50 ms followed by the number 8 displayed until a response was made. In the high-demand condition, the cue and target appeared in different boxes (antisaccade). Participants had to identify the target letter using the 1, 2, and 3 keys of the number pad, which were relabeled *B*, *P*, and *R*.

To familiarize participants with the response mapping and pacing of the task, we began the experiment with all participants completing 36 practice trials on which targets appeared at fixation (without cues). After the initial practice, participants completed 54 trials of either the high-demand (antisaccade) or low-demand (prosaccade) version of the task, depending on condition. Following the saccade task, participants completed the mood scale followed by the same version of operation span used in previous experiments.

Results and Discussion

The demand level of the initial task had no impact on operation span, $t(48) = 1.26$, $p = .21$, $d = 0.37$ ($M = 52.9\%$, $SD = 16.0$, for the antisaccade condition; $M = 58.7\%$, $SD = 16.3$, for the prosaccade condition), indicating that the demanding antisaccade task did not measurably deplete resources participants needed for operation span. To ensure the antisaccade task was indeed more demanding than the prosaccade task, we examined performance on the two saccade tasks. Participants in the antisaccade condition correctly identified only 71.38% of the targets ($SEM = 3.30\%$), whereas participants in the prosaccade condition identified 94.1% ($SEM = 0.93\%$), $t(48) = 6.17$, $p < .001$.

These results provide a strong test of the claim that working memory span is vulnerable to resource depletion. Here, participants performed a well-validated, highly demanding task (antisaccade) known to tap the same cognitive resources as complex span tasks (Kane et al., 2001; Unsworth et al., 2004), yet suffered no deficit on a subsequent span task.

Experiment 6

Depletion theory postulates a limited domain-general resource that is responsible for tasks as dissimilar as forcing oneself to eat radishes (Baumeister et al., 1998) and performing a complex span task (Schmeichel, 2007). Experiments 1–4 showed that with the video task, subsequent task effects emerge only when the distracting material in the video is similar to the memoranda in the span task. Experiment 5 showed that antisaccade, a validated measure of attentional control abilities, does not produce subsequent task effects on operation span. Taken together these results cannot be explained by an unembellished version of depletion theory. However, a modified version of the theory may be consistent with the results. Specifically, it is possible that there are separate resources for verbal and spatial working memory tasks (e.g., Baddeley & Logie, 1999; Logie, 1996; Todd & Marois, 2004; but see also Kane et al., 2004; Miyake et al., 2001) and that the similarity-dependent effects observed in Experiments 1–4 reflect depletion of those domain-specific resources rather than a domain-general resource. Similarly, antisaccade tasks, which require participants to control allocation of attention to spatial locations, may deplete only spatial resources, leaving verbal resources unimpaired and able to deal with the demands of the verbal operation span. In the final experiment, we tested of the idea that domain-specific working memory

[3] Determining if two tasks tap the same construct, such as attention control, is complicated by the fact that all tasks measure multiple constructs. Therefore, even if two tasks tap exactly the same construct, the correlation between them may nonetheless be low because they also measure many nonoverlapping constructs (e.g., in addition to cognitive control, antisaccade likely measures differences in visual acuity, perceptual processing speed, experience with rapidly presented visual material, and a host of other constructs). Some studies have shown somewhat modest correlations between antisaccade performance and complex span tasks (Miyake et al., 2000). However, when more sensitive techniques such as latent variable analyses are used, the evidence points toward a strong connection between complex span and antisaccade (Unsworth & Spillers, 2010). Indeed few, if any, tasks have a better established link to complex span performance, while also sharing a conceptual similarity to the video task and while using stimuli that are dissimilar to those in the span tasks.

resources are depletable by administering antisaccade, a task that requires participants to control the spatial allocation of attention, followed by the administration of a clearly spatial span task.

Method

Participants. Twenty-three undergraduate students were tested in the high-demand condition, and 25 were tested in the low-demand condition.

Materials and procedure. Participants completed the same antisaccade task (or prosaccade task in the low-demand condition) used in Experiment 5 followed by the highly spatial symmetry span task.

Like rotation span, symmetry span is a spatial analogue of verbal complex span tasks and correlates well with other complex span tasks as well as with fluid intelligence measures (Kane et al., 2004; Unsworth, Brewer, & Spillers, 2009). We used symmetry span rather than rotation span to increase the reliance on spatial abilities. Whereas the relatively simple stimuli in rotation span could conceivably be encoded in a verbal manner (e.g., the directions of arrows could be translated into hours on a clock face), the complexity of symmetry span makes verbal coding unlikely.

We used the automated version (Unsworth, Heitz, Schrock, & Engle, 2005) of symmetry span (Kane et al., 2004). The task interleaves a series of symmetry judgments with presentation of spatial locations within a matrix for later recall. Each symmetry judgment presents a pattern formed by shading some cells in 8 × 8 matrix of white squares. Participants must indicate if the pattern is vertically symmetrical by pressing a key. After participants make the symmetry judgment, a blank 4 × 4 matrix is displayed, and one of the cells turns red for 650 ms. The participants must remember the location of the red cell. After the participants have made between two and five symmetry judgments and have seen a matching number of to-be-remembered red cells, a blank 4 × 4 matrix appears, and participants must reproduce the pattern of red cells from that trial by clicking on the cells in the order they were presented. There were three trials at each set size (between two and five). As before, partial credit scoring was used. Participants were given practice with the symmetry judgment and the memory component separately before beginning the scored trials.

Results

Whether participants completed the high-demand antisaccade task or the low-demand prosaccade task had no impact on subsequent symmetry span scores. Indeed, antisaccade participants were actually slightly more accurate on the span task ($M = 74.3\%$, $SD = 12.1$) than were prosaccade participants ($M = 69.2\%$, $SD = 17.3$), though the difference was not significant, $t(46) = -1.17$, $p = .25$, $d = -0.35$. As in Experiment 5, participants had much more difficulty with the antisaccade task than the prosaccade task: Antisaccade participants correctly identified only 64.19% of the targets ($SEM = 3.60\%$), while prosaccade participants identified 92.19% ($SEM = 1.90\%$), $t(43) = 6.60$, $p < .001$ (note that saccade data for three participants were lost due to a computer backup error).

The fact that the highly spatial antisaccade task had no impact on the spatial symmetry span task speaks against a modified version of depletion theory wherein domain-specific verbal and spatial resources are vulnerable to depletion. That is, the processes underlying complex span, whether they are domain general or domain specific, do not suffer from resource depletion under the conditions that we tested.

General Discussion

In this article, we sought to assess whether subsequent-task effects on complex span occur because the initial task depletes the cognitive control resources needed for complex span. A key feature of the depletion hypothesis is that subsequent-task effects should emerge whenever the two tasks require the same resources. To test the depletion account, we maintained the need for cognitive control in all tasks but varied the similarity among the stimuli that were to be ignored in the initial task and those that were to be remembered in the span task. Subsequent-task effects appeared only when there was a match between the to-be-ignored stimuli in the first task and the to-be-remembered stimuli in the span task: Ignoring words impaired memory for words (Experiment 1) but not memory for arrows (Experiment 2), while ignoring arrows impaired memory for arrows (Experiment 3) but not memory for words (Experiment 4). That is, subsequent-task effects depended not on deep similarities among the resources required for the tasks, as predicted by depletion theory, but rather on similarities among the stimuli. Perhaps more damaging to the depletion account, antisaccade, a task known to tap the domain-general executive control components of span (Kane et al., 2001; Unsworth et al., 2004), did not produce depletion effects on either verbal (Experiment 5) or spatial (Experiment 6) span. Together these results suggest that subsequent-task effects on complex span are not due to depletion of resources and do not represent the consequences of intrinsic changes to the working memory system.

It is important to note that while the current results strongly suggest that working memory span is relatively immune to true resource depletion, they do not rule out the possibility that other cognitive processes are susceptible to depletion. There is evidence that demanding self-control tasks can produce subsequent-task effects on other common cognitive tasks. For example, suppressing emotional reactions while watching provocative videos has been shown to increase the incongruency effect on a subsequent Stroop task (Inzlicht & Gutsell, 2007). Similarly, it has been found that performing a high-interference, but not a low-interference, version of the recent-probes task, in which probe items must be matched to a current memory set and in which experimenters manipulate interference by varying whether the probe matches items in earlier memory sets, reduced subsequent performance on both a verb-generation task and a cued recall paired-associates task (Persson, Welsh, Jonides, & Reuter-Lorenz, 2007). There is also evidence that engaging in a self-control task can lower performance on some standardized tests (Schmeichel, Vohs, & Baumeister, 2003). Such results may reflect true depletion.

Moreover, while complex span may not be vulnerable to true depletion effects, the present findings and those of Schmeichel (2007) clearly show that prior tasks can influence span scores in unexpected ways. These subsequent task effects have important practical and methodological implications. In most individual difference studies, complex span tasks are embedded within a long series of other tasks (e.g., Kane et al., 2004; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010; Miyake et al., 2001; Unsworth et al., 2009); the susceptibility of complex span to subsequent-task effects is therefore a critical concern. Designing task sequences to avoid contaminating results with subsequent-task effects will require understanding the factors responsible for such effects. Another concern is whether, in addition to lowering mean performance on span tasks, subsequent-task effects change the pattern of correlations between span and other aspects of cognition such as fluid intelligence.

Extrinsic Versus Intrinsic Changes

Previously, we introduced the distinction between differences in memory performance that result from intrinsic changes to the memory system itself (e.g., the efficiency of interference control mechanisms) versus from factors extrinsic to the memory system (e.g., the amount of interference that must be resolved). This distinction is critical because only intrinsic changes are likely to impact other aspects of cognition that require memory, such as reasoning, whereas extrinsic changes are likely to be limited to particular test tasks. The present data suggest that resource depletion manipulations do not produce true intrinsic reductions in working memory and instead depend on extrinsic factors such as the similarity of the materials used in the initial task and the subsequent task.

Are there any factors that do represent intrinsic changes? We suspect that developmental changes in span (i.e., increases in childhood, decreases in late adulthood) are likely to reflect intrinsic alterations to the underlying mechanisms. For example, elsewhere we have argued that the inhibitory mechanisms that regulate the flow of information into and out of memory become less efficient with age and that this reduced inhibitory efficiency directly translates to reduced working memory span (i.e., the system is intrinsically less able to deal with interference; Hasher et al., 1999, 2007). On the opposite end of the development spectrum, Cowan (e.g., 2010) has suggested that the focus of attention increases in size during maturation, which would also constitute an intrinsic alteration of the system.

Time-of-day effects are another potential instance of intrinsic change. Individuals often perform best on cognitive tasks, including working memory tasks, when the time of testing coincides with the peak of their circadian arousal pattern (see Hasher, Goldstein, & May, 2005, and Winocur & Hasher, 2002, for reviews). In these studies, only the synchrony/asynchrony between the time of testing and arousal patterns is manipulated, suggesting that the demands placed on working memory do not differ but rather that the ability to deal with those demands varies across the day. There is also evidence that measures of fluid intelligence vary with circadian arousal (Goldstein, Hahn, Hasher, Wiprzycka, & Zelazo, 2007).

The distinction between intrinsic and extrinsic influences on memory performance is clear in most cases. For example, depletion of attentional resources would clearly be an intrinsic factor, whereas similarity of stimuli is clearly an extrinsic factor. However, other factors that influence memory performance are more difficult to classify as intrinsic or extrinsic to the memory system. For example, level of motivation can influence performance on a memory task (Heitz, Schrock, Payne, & Engle, 2008). On the one hand, motivation does not change the external demands placed on the memory system, suggesting changes in performance are due to intrinsic factors. On the other hand, it would be odd to say that individuals who are highly motivated to perform a given task have intrinsically different memory systems than individuals with low motivation. That is, motivation is intrinsic in the sense of being an internal psychological factor, rather than an external environmental factor, but extrinsic in the sense that it can be thought of as external to the memory system. This ambiguity raises the issue of interactions between different cognitive systems. For example, a motivational system may modulate the degree to which the memory system is engaged during a particular task. Indeed, Robinson et al. (2010) proposed that depletion effects arise from the interaction of three systems: a control system that prevents contextually inappropriate thoughts or actions (e.g., reading words in a Stroop task); a monitoring system that detects potential failures of control; and a motivational system that modulates the control and monitoring systems, so that when motivation is low, detection and correction of inappropriate responses are impaired. They suggest that many depletion effects may be due to reduced motivation rather than to intrinsic alterations to the monitoring and control systems.

We would be inclined to classify changes in memory performance due to changes in motivation as extrinsic because the locus of the change is outside the memory system; however, at this level of subtlety, the intrinsic/extrinsic distinction may be less useful. Rather, the strength of the intrinsic/extrinsic distinction is as a heuristic for discriminating between manipulations that alter the way that the memory system operates (an intrinsic change) and are therefore likely to lead to performance differences on a range of tasks and extrinsic manipulations that simply increase the demand that a given task places on the memory system and are therefore likely to be task specific.

Summary and Conclusion

Considerable progress has been made in illuminating the mechanisms that limit working memory and how they differ between individuals. But do memory limits also vary within individuals? Depletion theory claims that engaging working memory resources on one task reduces their efficiency, suggesting that limits do indeed vary within individuals (see also Kahneman, 1973). Consistent with this claim, Schmeichel (2007) found that performing a difficult task, but not an easy task, prior to complex span led to reduced span scores. It is critical to determine if the reported reductions in span scores reflect intrinsic changes to the working memory system, which given the strong and possibly causal link between working memory and intelligence might be expected to translate to overall cognitive impairment. In a series of experiments, we found that while initial tasks can reduce subsequent span scores, the effect does not depend on deep similarity between the cognitive processes engaged by the two tasks, as predicted by depletion theory, but rather depends on similarity between the stimuli used in the two tasks. These results suggest that subsequent-task effects do not reflect intrinsic changes to working memory limits.

References

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087. doi:10.1037/0278-7393.20.5.1063

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100. doi:10.1037/0033-295X.102.1.68

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and structure of short-term memory. *Journal Verbal Learning & Verbal Behavior, 14,* 575–589. doi:10.1016/S0022-5371(75)80045-4

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 28-61). New York, NY: Cambridge University Press.

Barnes, A., Bullmore, E. T., & Suckling, J. (2009). Endogenous human brain dynamics recover slowly following cognitive effort. *PLoS ONE, 4*(8), e6626. doi:10.1371/journal.pone.0006626

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74,* 1252–1265. doi:10.1037/0022-3514.74.5.1252

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16*, 351–355. doi:10.1111/j.1467-8721.2007.00534.x

Bunting, M. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 183–196. doi:10.1037/0278-7393.32.2.183

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 & Cognition, 27,* 584–591. doi:10.3758/BF03211552

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science, 21,* 399–405. doi:10.1177/0956797609359910

Chasteen, A. L., Bhattacharyya, S., Horhota, M., Tam, R., & Hasher, L. (2005). How feelings of stereotype threat influence older adults' memory performance. *Experimental Aging Research, 31,* 235–260. doi:10.1080/03610730590948177

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds.). (2007). *Variation in working memory.* New York, NY: 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. doi:10.3758/BF03196772

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. doi:10.1017/S0140525X01003922

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19,* 51–57. doi:10.1177/0963721409359277

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19,* 450–466. doi:10.1016/S0022-5371(80)90312-6

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433. doi:10.3758/BF03214546

De Beni, R., Borella, E., & Carretti, B. (2007). Reading comprehension in aging: The role of working memory and metacomprehension. *Neuropsychology, Development, and Cognition, Section B: Aging, Neuropsychology, and Cognition, 14,* 189–212. doi:10.1080/13825580500229213

Ericsson, K. A., Delaney, P. F., Weaver, G., & Mahadevan, R. (2004). Uncovering the structure of a memorist's superior "basic" memory capacity. *Cognitive Psychology, 49,* 191–237. doi:10.1016/j.cogpsych.2004.02.001

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience, 8,* 1298–1300. doi:10.1038/nn1543

Goldstein, D., Hahn, C. S., Hasher, L., Wiprzycka, U. J., & Zelazo, P. D. (2007). Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a synchrony effect? *Personality and Individual Differences, 42,* 431–440. doi:10.1016/j.paid.2006.07.008

Grigg, O., & Grady, C. L. (2010). Task-related effects on the temporal and spatial dynamics of resting-state functional connectivity in the default network. *PLoS ONE, 5*(10), e13311. doi:10.1371/journal.pone.0013311

Hasher, L., Goldstein, D., & May, C. P. (2005). It's about time: Circadian rhythms, memory, and aging. In C. Izawa & N. Ohta (Eds.), *Human learning and memory: Advances in theory and application. The 4th Tsukuba International Conference on Memory* (pp. 199–217). Mahwah, NJ: Erlbaum.

Hasher, L., Lustig, C., & Zacks, R. T. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, A., & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York, NY: Oxford University Press.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. In W. S. Sossin, J. C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Progress in brain research, Vol. 169*: *Essence of memory* (pp. 353–363). Amsterdam, the Netherlands: Elsevier.

Healey, M. K., Campbell, K. L., Hasher, L., & Ossher, L. (2010). Direct evidence for the role of inhibition in resolving interference. *Psychological Science, 21,* 1464–1470. doi:10.1177/0956797610382120

Healey, M. K., Zacks, R. T., Hasher, L., & Helder, E. (2011). *Working memory, long-term memory, and fluid intelligence: The role of interference control.* Manuscript submitted for publication.

Heitz, R. P., Schrock, J., Payne, T., & Engle, R. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45,* 119–129.

Hess, T. M., Hinson, J. T., & Hodges, E. A. (2009). Moderators of and mechanisms underlying stereotype threat effects on older adults' memory performance. *Experimental Aging Research, 35,* 153–177. doi:10.1080/03610730802716413

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11,* 365–371. doi:10.1111/1467-9280.00272

Inzlicht, M., & Gutsell, J. N. (2007). Running on empty: Neural signals for self-control failure. *Psychological Science, 18*(11), 933–937. doi:10.1111/j.1467-9280.2007.02004.x

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General, 137,* 691–705. doi:10.1037/a0013834

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59,* 193–224. doi:10.1146/annurev.psych.59.103006.093615

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice–Hall.

Kane, M. J., Conway, A. R. A., Bleckley, M. K., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130,* 169–183. doi:10.1037/0096-3445.130.2.169

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse, (Eds.), *Variation in working memory* (pp. 21–48). New York, NY: Oxford University Press.

Kane, M. J., Tuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133,* 189–217. doi:10.1037/0096-3445.133.2.189

Kyllonen, P. C. (1996). Is working memory capacity Spearman's g? In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 49–75). Mahwah, NJ: Erlbaum.

Logie, R. H. (1996). The seven ages of working memory. In J. T. E. Richardson, R. W. Engle, L. Hasher, R. H. Logie, E. R. Stoltzfus, & R. T. Zacks (Eds.), *Working memory and human cognition*. New York, NY: Oxford University Press.

Lustig, C., & Hasher, L. (2006). Interference. In R. Schulz, L. Noelker, K. Rockwood, & R. Sprott (Eds.), *Encyclopedia of aging: A comprehensive resource in gerontology and geriatrics* (4th ed., Vol. 1, pp. 602–607). New York, NY: Springer.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a "new view." In D. S. Gorfein & C. M. MacLeod (Eds.), *The place of inhibition in cognition* (pp. 145–162). Washington, DC: American Psychological Association. doi:10.1037/11587-008

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. doi:10.1037/0096-3445.130.2.199

May, C. P., Hasher, L., & Kane, M. J. (1999). The role of interference in memory span. *Memory & Cognition, 27,* 759–767. doi:10.3758/BF03198529

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin, 118,* 35–54. doi:10.1037/0033-2909.118.1.35

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of Personality and Social Psychology, 55,* 102–111. doi:10.1037/0022-3514.55.1.102

McCabe, D. P., Roediger, H. L., 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. doi:10.1037/a0017619

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. doi:10.1037/h0043158

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. doi:10.1006/cogp.1999.0734

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. doi:10.1037/0096-3445.130.4.621

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421. doi:10.1037/0278-7393.28.3.411

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. doi:10.1037/0096-3445.134.3.368

Persson, J., Welsh, K., Jonides, J., & Reuter-Lorenz, P. (2007). Cognitive fatigue of executive processes: Interaction between interference resolution tasks. *Neuropsychologia, 45,* 1571–1579. doi:10.1016/j.neuropsychologia.2006.12.007

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. *Journal of Verbal Learning & Verbal Behavior, 7,* 672–694. doi:10.1016/S0022-5371(68)80124-0

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1,* 19–40. doi:10.3758/BF03198064

Roberts, R. J., 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. doi:10.1037/0096-3445.123.4.374

Robinson, M., Schmeichel, B., & Inzlicht, M. (2010). A cognitive control perspective of self-control strength and its depletion. *Social and Personality Psychology Compass, 4,* 189–200. doi:10.1111/j.1751-9004.2009.00244.x

Rowe, G., Hasher, L., & Turcotte, J. (2010). Interference, aging and visuospatial working memory: The role of similarity. *Neuropsychology, 24,* 804–807. doi:10.1037/a0020244

Rowe, G., Valderrama, S., Hasher, L., & Lenartowicz, A. (2006). Attention disregulation: A benefit for implicit memory. *Psychology and Aging, 21,* 826–830. doi:10.1037/0882-7974.21.4.826

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440–452. doi:10.1037/0022-3514.85.3.440

Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General, 136*(2), 241–255. doi:10.1037/0096-3445.136.2.241

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology, 85,* 33–46. doi:10.1037/0022-3514.85.1.33

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4–28. doi:10.1006/jesp.1998.1373

Todd, J. J., & Marois, R. (2004, April 15). Capacity limit of visual short-term memory in the human posterior parietal cortex. *Nature, 428,* 751–754. doi:10.1038/nature02466

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language, 39,* 195–217. doi:10.1006/jmla.1998.2574

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison–Wesley.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49–60. doi:10.1037/h0044616

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2009). There's more to the working memory capacity–fluid intelligence relationship than just secondary memory. *Psychonomic Bulletin & Review, 16,* 931–937. doi:10.3758/PBR.16.5.931

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. doi:10.3758/BF03192720

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1302–1321. doi:10.1037/0278-7393.30.6.1302

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal Of Memory And Language, 62,* 392–406. doi:10.1016/j.jml.2010.02.001

Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review, 77,* 1–15. doi:10.1037/h0028569

Winocur, G., & Hasher, L. (2002). Circadian rhythms and memory in aged humans and animals. In L. Squire & D. Schacter (Eds.), *Neuropsychology of memory* (3rd ed., pp. 273–285). New York, NY: Guilford.

Received June 23, 2010
Revision received April 20, 2011
Accepted May 2, 2011 ■