Working-Memory Capacity, Proactive Interference, and Divided Attention: Limits on Long-Term Memory Retrieval

Michael J. Kane
Georgia State University

Randall W. Engle
Georgia Institute of Technology

Two experiments examined how individual differences in working-memory capacity (WM) relate to proactive interference (PI) susceptibility. We tested high and low WM-span participants in a PI-buildup task under single-task or dual-task ("load") conditions. In Experiment 1, a finger-tapping task was imposed during encoding *and* retrieval of each list; in Experiment 2, tapping was required during encoding *or* retrieval. In both experiments, low spans showed greater PI than did high spans under no load, but groups showed equivalent PI under divided attention. Load increased PI only for high spans, suggesting they use attention at encoding and retrieval to combat PI. In Experiment 2, only low spans showed a dual-task cost on List 1 memory, before PI built up. Results indicate a role for attentional processing, perhaps inhibitory in nature, at encoding and retrieval, and are discussed with respect to theories of WM and prefrontal cortex function.

The central executive component of working memory is now a focus of theoretical and empirical interest. Although its functions are still under-specified, they are most often characterized as "attentional" (e.g., Baddeley, 1993, 1996; Baddeley & Logie, 1999; Engle, Tuholski, Laughlin, & Conway, 1999). Baddeley (1986), for example, proposed that the central executive may be analogous to the Supervisory Attentional System (SAS) described by Shallice and colleagues (e.g., Norman & Shallice, 1986; Shallice & Burgess, 1993). The SAS has several functions and is theorized to be critical for resolving interference arising when the environment triggers conflicting action schemas. For example, the sight of a colleague with a new unflattering haircut may prompt several reactions, most of which are

socially inappropriate. In such cases, the SAS biases the action-selection process away from the momentarily potent, but contextually unacceptable, response. This attentional bias is accomplished by the SAS providing additional activation to a more appropriate, but less active, schema ("*Love your new hairstyle!*"), and also inhibiting the activation of the inappropriately triggered action.

In a similar vein, Engle and colleagues have argued that individual differences on measures of "working-memory capacity" reflect the capability to use controlled attention to prevent distraction from the environment and interference from events stored in long-term memory (e.g., Conway & Engle, 1994; Engle, 1996; Rosen & Engle, 1997; 1998; for a related view see Dempster, 1991, 1992). Evidence for this view comes primarily from research demonstrating the relationship between working-memory capacity and interference susceptibility, and this is the focus of the present investigation. The notion that limits in immediate memory are associated with interitem interference effects is not new (e.g., Foucalt, 1928; Gibson & Raffel, 1936; for a review see Dempster, 1981). For example, Young and Supa (1941) demonstrated that recall of short lists of one item type, either all digits or all categorized words, was worse than for lists in which the item type switched halfway through (e.g., *2-4-1-9-3-hen-cow-pig*). Young and Supa argued that switching categories within the list released the intralist interference from the prior category of items, and thereby increased span.

Subsequent efforts to correlate short-term memory span with independent measures of interference have had limited success, however. The correlations tend to be in the right direction, with higher span predicting less interference susceptibility, but they are statistically weak. Underwood, Boruch, and Malmi (1978) presented 200 participants with three different span tests and three different "interference" tests, among many other tasks. Participants who recalled more words on the interference tasks tended to have higher spans. However, the correlations were not significant (r ranging from .16 to .32), perhaps because two of the

putative interference tests failed to elicit measurable interference. Dempster and Cooney (1982) also found weak correlations between short-term memory span and proactive interference in two experiments with Brown-Peterson-type tasks (Brown, 1958; Peterson & Peterson, 1959). Although the correlations were not statistically significant ($rs = -.34$ and $-.27$), only participants in the lower median of span showed significant interference, suggesting a relationship between the two variables. Small sample sizes and potential ceiling and floor effects likely limited the power to detect significant correlations.

Researchers have been more successful linking the *working* memory construct to interference. Conway and Engle (1994), for example, tested high and low working-memory-span participants with a modification of the Sternberg (1966) probe-recognition task. In one experiment, participants learned associations between digit cues and different-sized letter sets, for example, R and W associated with 2, and B, K, Q, and Z associated with 4. An extensive study session assured that the span groups were matched at high levels of learning. On a subsequent speeded recognition task, high- and low-span participants were equally fast in recognizing whether a letter belonged to its digit set across all set sizes. In another experiment, however, each letter appeared in two memory sets (e.g., R in RW-2; $BKRZ$-4), and so the multiple associative links between a letter, R, and its two respective sets, 2 and 4, should have produced a considerable level of competition at retrieval. Here, low spans' recognition times were significantly longer than in the noninterference condition, but high spans' were not. Low spans thus appeared more vulnerable to interference than did high spans, who may have blocked or inhibited the interfering link between target letters and their momentarily irrelevant cue.

The present study addressed the question of whether individual differences in working-memory capacity correspond to individual differences in susceptibility to proactive interference (PI). Moreover, we examined whether individuals of different working-memory capabilities might differentially use controlled attention to actively resist PI. Relevant to both of these issues, Rosen and Engle (1997) tested high and low spans in a category fluency task. Participants recalled as many animal names as they could, without repetitions, for 10–15 min. Under standard conditions, high spans recalled more animals than did low spans, with small but significant span differences in the first minute that increased dramatically across the recall period. However, high and low spans performed very similarly under divided-attention conditions, because the load task reduced fluency only for high spans.

The reduction of high spans' fluency under divided attention suggests that they normally used controlled processing to achieve superior fluency. In contrast, the fluency of low spans, which was lower than for high spans, was *unaffected* by the secondary task. This finding suggests that low spans did *not* use controlled processing under standard conditions. If low spans did not use controlled processing to retrieve exemplars, then their recall could not be limited by the addition of a secondary task. It appeared, then, that low

spans may have relied on relatively automatic processing under standard fluency conditions (i.e., on automatic spreading activation among category exemplars), which, although resistant to load effects, allowed for only poor fluency relative to that of high spans.

But were low spans truly not using controlled processing at all during the fluency task? In fact, Rosen and Engle (1997) argued that low spans' attentional processing was directed at covertly monitoring for intrusions. Fluency tasks provide a rich occasion for intrusions, with output interference from earlier-recalled words building over a long recall period. That is, participants typically recall most of the high dominance exemplars from a category, such as *cat, dog, cow, horse,* and so on, well within the first minute of recall. Successful fluency performance across longer intervals therefore requires strategically searching for low dominance exemplars while preventing re-retrieval of (or perseveration on) already-recalled exemplars.

In a subsequent experiment, Rosen and Engle (1997) instructed participants that reporting aloud whatever exemplars came to mind, even if they had already been recalled, would enhance performance. Here, the recall for high spans did not differ from standard conditions, indicating that high spans did not actually re-retrieve reported items. In contrast, the output of low spans grew dramatically, and almost all the increased output consisted of already-recalled items. Rosen and Engle therefore concluded that low spans prevented reporting these covert intrusions under standard conditions by using controlled processing to monitor spoken output. Moreover, by using attention in this way they had little capacity left to strategically search for additional exemplars. High spans, however, actually blocked covert intrusions from coming to mind, and so avoided having to monitor recall.

The present study more directly explored the relationships among working-memory capacity, divided attention, and interference effects in memory. To this end, we tested high- and low-span participants in a PI buildup and release task (e.g., Craik & Birtwistle, 1971; Wickens, Born, & Allen, 1963), with half the participants tested under divided attention. PI is said to occur when retrieval of a recent episode (e.g., Word List 2) is impaired because of the *prior* study of a similar episode (e.g., Word List 1). Several memory theories assume, implicitly or explicitly, that resisting PI should require some type of controlled processing either at the encoding or retrieval of List 2 (e.g., M. C. Anderson & Bjork, 1994; M. C. Anderson & Neely, 1996; Hasher & Zacks, 1988). Indirect evidence for this assumption comes from studies showing that PI is maximized when prior-list information gains access to working memory coincidentally with target information. For example, PI increases when Lists 1 and 2 must be recalled in alternation, when some List 1 stimuli are re-presented during List 2 recall, or when participants know that List 1 will have to be recalled after recalling List 2 (Hawkins, Pardo, & Cox, 1972; Postman & Hasher, 1972; Postman, Stark, & Fraser, 1968). Moreover, the typical reduction of PI that is found when participants are explicitly told to "forget" List 1

before studying List 2 disappears if some List 1 words are re-presented before a final test of List 2 (Bjork, Bjork, & Glenberg, 1973; Bjork, Bjork, & White, 1984). If people are somehow prevented from doing the mental work required to block interfering information at retrieval, that interfering information will impair target recall.

The present study explicitly investigated the attention demands of interference resistance by dividing attention during the PI task. Insofar as interference resistance requires controlled attention, a secondary load should disrupt it and increase PI. For our secondary task we adopted a finger-tapping task devised by Moscovitch (1992a, 1994). This tapping task demands attention, it is endogenously driven, it has little surface similarity to the PI task, and its novel motor-sequencing requirements may put a premium on efficient functioning of the prefrontal cortex (see Frith, Friston, Liddle, & Frackowiak, 1991; Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994). Indeed, Mosco-vitch and others (Martin, Wiggs, Lalonde, & Mack, 1994) have used this secondary task with normal healthy adults in order to simulate the cognitive performance of patients with prefrontal cortex damage. The executive functions of work-ing memory have been closely aligned with structures in the prefrontal cortex (e.g., Baddeley, 1996; Duncan, 1995; Moscovitch, 1992a, 1992b; Shallice & Burgess, 1993), and so a secondary task that relies on prefrontal function might be especially likely to limit the processes of interest here (see Moscovitch, 1994).

For Experiment 1, we predicted that high spans would show less PI susceptibility than would low spans, and that for high spans, a secondary load task would significantly increase PI susceptibility. The relatively diminished con-trolled-attention capabilities of low spans should leave them more susceptible to PI under normal conditions. Moreover, if low spans do not use controlled attention to counteract PI, as is indirectly suggested by Rosen and Engle (1997), then low spans should be unaffected by a secondary load task. That is, dividing attention will not impair the performance of low spans if they do not use controlled attention in the first place. We note here that our predictions for working-memory span and secondary-load tasks are aimed at their impact on interference, specifically, which is measured as a *change* in recall from some baseline level, such as the recall of List 1. Thus, our primary measure of interest will be a proportional PI measure, which takes List 1 recall into account in assessing the impact PI buildup. We entertained no specific hypotheses about span and load effects on List 1.

To preview, Experiment 1 demonstrated span differences in PI on a buildup task (modeled after Craik & Birtwistle, 1971), as well as span differences in the impact of divided attention on PI. In Experiment 2, we used the dual-task method to further explore the relative attentional demands of *encoding* versus *retrieval*, by manipulating whether the load task occurred during the encoding or retrieval of each list. Although dividing attention had no effect on the PI buildup for low spans, it increased the PI buildup for high spans whether it occurred at encoding or retrieval.

EXPERIMENT 1

Method

Participant Screening for Working-Memory Capacity

We screened participants for working-memory capacity using the operation–word span task (OSPAN), in which they solved series of simple mathematical operations while attempting to remember a list of unrelated words (for details, see La Pointe & Engle, 1990). Participants saw one operation–word string at a time, and each list of operation–word strings ranged from two to six items in length. For example, a list of three strings might be

$$IS \ (9/3) + 2 = 5 \ ? \ drill$$
$$IS \ (5 \times 1) - 4 = 2 \ ? \ beach$$
$$IS \ (2 \times 2) + 3 = 7 \ ? \ job$$

A Micro Experimental Laboratory (MEL) 2.0 program presented the stimuli at the center of an IBM color monitor with a VGA graphics card (set to black and white). Participants were tested individually and sat at the most comfortable viewing distance from the monitor.

The experimenter instructed the participant to begin reading the operation–word pair aloud as soon as it appeared. If the participant paused before reading aloud, the experimenter reexplained that pausing was not allowed. After reading the equation aloud, the participant verified aloud whether or not the provided answer was correct and then read the word aloud. Immediately after the participant read the word aloud, the next operation appeared. The participant then immediately read the next operation aloud, and the sequence continued until three question marks (*???*) cued the recall of the words from that list only. Participants wrote the words on an answer sheet in the same order in which they had been presented.

The OSPAN score was the sum of the words recalled for all lists that were completely recalled in correct order. Participants were tested on three lists of each length (from 2 to 6 operation–word pairs), and so possible scores ranged from 0 to 60. Unlike the Daneman and Carpenter (1980) procedure in which participants begin with the smallest sized lists and progress to the larger sized lists, here the different-sized lists appeared in an unpredictable order. Thus, participants did not know the number of words to be recalled until the presentation of the recall cue.

Participants

One hundred ninety-two undergraduates from the University of South Carolina—Columbia participated in Experiment 1 in return for psychology course credit. These participants were identified from a larger pool who had participated in the OSPAN task: 96 participants were selected from the top quartile of the distribution (hereafter, "high spans"), and 96 were selected from the bottom quartile (hereafter, "low spans"); any participants who had cor-rectly solved fewer than 85% of the OSPAN operations were not invited back (such participants typically represent fewer than 1% of those tested). Between 1 and 90 days intervened between a given individual's participation in the OSPAN task and the PI-task session.

Design

The design was a 2 × 2 × 4 mixed-model factorial, with span group (high, low) and tapping condition (cascade, complex) manipulated between subjects, and List (1–4) manipulated within subjects. The "tapping" variable refers to the two possible finger-tapping sequences that were maintained during the PI task, and the "list" variable refers to the four consecutive word lists that were studied and recalled. Forty-eight participants in each span group were randomly assigned to one of the two between-subject tapping conditions.

Apparatus and Materials

PI Task

A MEL 1.0 program presented the stimuli for the PI task in standard font at the center of an IBM color monitor with a VGA graphics card. Forty exemplars from each of three taxonomic categories were selected from the Battig and Montague (1969) category norms (animals, occupations, and countries). All of the selected exemplars were single words of 10 or fewer letters and were normatively ranked below the 12 strongest associates from the category. Within each category, we divided the 40 exemplars into 4 subsets of 10 words each, and approximately matched the category associative strengths across these subsets.

From these subsets, we constructed 12 complete stimulus sets of 40 words each, with each set used equally often across participants and tapping conditions (see Appendix A for sample lists). A given participant saw only one stimulus set. Each set consisted of three subsets from one category (e.g., animals) and one subset from another category (e.g., occupations). Across stimulus sets and participants, each subset appeared equally often as the stimuli for Lists 1–4. Four sets consisted of animal PI-buildup lists, four consisted of occupation PI-buildup lists, and four consisted of country PI-buildup lists. For each category of PI-buildup set, half of the sets ended with a release list from one category, and half ended with a release list from another category (e.g., two of the four animal PI-buildup lists were followed by occupation release lists, and two were followed by country release lists).

Rehearsal-Prevention Task

Within the PI task, participants performed a rehearsal-prevention task following each list (see Glanzer & Cunitz, 1966). We based this task on an oral version of the Trail-Making Test (Armitage, 1946; Partington & Leiter, 1949; Reitan, 1955) that was devised independently by Ricker and Axelrod (1994) and Baddeley (1996). On each trial of the Trails task, participants saw a letter and a two-digit number. Four trial orders were created for four letter–number combinations (F–61, P–26, L–83, and O–39), with each letter–number combination appearing once in each PI-list position (i.e., in each of the Memory Lists 1–4). Each of these trial orders was used equally often across animals, occupations, and countries categories, across participants, and across tapping conditions. Two letter–number combinations (C–10, J–45) served as practice stimuli.

The Trails task began with a letter and number appearing in yellow against a black screen for 16 s, and participants alternately counted aloud from this letter and number. For example, if F–61 appeared on screen, participants would count, "F–61, G–62, H–63," and so on. Participants were instructed to count aloud quickly and accurately, and that they would be scored for the number of items they counted minus any errors.

Tapping Task

A program executed by a second computer collected the finger-tapping data. Liquid-crystal finger-pads ("Piezo Film Sensors," from AMP, Inc.) were taped to the four fingertips of the nondominant hand, as indicated by self-report. These finger-pads were each wired to a single circuit box, plugged into the parallel port, and sent a signatured impulse when tapped on the table.

Procedure

We tested each participant individually in a sound-attenuated testing room. The experimenter first taped the four finger-pads onto the participant's nondominant hand and then demonstrated the tapping sequence. Participants were instructed to repeatedly tap the sequence at a "comfortable and consistent rate."

We randomly assigned each participant to either the "complex" or "cascade" tapping condition. The complex-tapping condition provided an attentional burden, or "load," by requiring the maintenance of a novel tapping sequence (after Moscovitch, 1994): index finger—ring finger—middle finger—pinkie. The cascade-tapping condition resembled the complex-tapping condition but was not as demanding: pinkie—ring finger—middle finger—index finger. Pilot testing suggested that this often-habitual tapping sequence had no measurable impact on memorizing. Importantly, only complex-tapping participants were given strict instructions about tapping accuracy because we meant for them to focus attention on the tapping sequence. We told cascade-tapping participants to merely keep tapping in a natural way.

All participants began with three 30-s practice trials of tapping. Complex-tapping participants received on-line accuracy feedback (a 500-ms tone followed every error) and were told that hearing many tones would indicate they should slow down. In a final, 60-s practice trial, no participants received accuracy feedback (although complex-tapping subjects were warned that accuracy was being recorded). Instead, all participants received response-time feedback on-line. The computer determined the feedback cutoff times for each participant individually: During the previous 30-s practice trial, the computer calculated the mean intertap interval and added 150 ms to it. This mean + 150 ms interval then became the feedback cutoff for the 60-s practice trial. That is, if any one intertap interval was more than 150 ms slower than the established cutoff from the prior practice trial, the computer immediately emitted a 500-ms tone.

After hearing the PI-task instructions, participants attempted two 30-s practice trials with the rehearsal-preventing Trails task. Finally, just before beginning the PI task, participants again practiced finger tapping with response-time feedback for 30 s. The experimenter then explained that the primary job in the upcoming PI task was to maintain practice tapping speeds throughout and that tapping should not be compromised to improve memory performance.

The PI task began with a green, "begin tapping," instruction screen. This "baseline tapping" signal remained onscreen for 20 s, during which the participant tapped with response-time feedback; here, the feedback cutoff time was calculated for each participant as the mean intertap interval from the immediately preceding 30-s trial, plus 150 ms. From this point onward, through all of the remaining stages of the PI task, all participants repeatedly tapped with this response-time feedback.

Following the 20-s tapping baseline, a light blue warning signal (!! Get Ready !!) flashed intermittently against a black background for 1.5 s, after which the 10 words from the first memory list appeared one at a time, centered onscreen at a 2,000-ms rate

(1,750-ms stimulus duration + 250-ms interstimulus interval). The words appeared in white capitalized letters against a black background. Participants read each word aloud as it appeared. Immediately after the final word disappeared, participants performed the Trails task. After the participants counted for 16 s, a green screen cued them to orally recall words from the memory list. Participants had 20 s to recall the words in any order and were encouraged to keep working to recall throughout the entire period. At the end of the recall period, a red screen appeared for 2 s, which instructed the participant to stop recalling. This sequence, beginning with the light blue warning signal, then immediately repeated for Lists 2, 3, and 4.

Note that the PI task and the tapping task did not correspond to typical assignments to "primary" and "secondary" task categories. That is, we were primarily interested in the effect of dividing attention on PI susceptibility, and so we structured the tasks to push any performance variability into the PI task, as opposed to the tapping task. We did this by giving extended practice on the tapping task alone with feedback and then by providing on-line feedback on the tapping task during the dual-task procedure. In attempting to keep tapping performance stable, and in emphasizing tapping performance in the dual-task instructions, we led participants to consider tapping as the *primary* task, and memorizing words as the *secondary* task.

Results

Participants

We replaced the data from three low spans and two high spans because of equipment malfunctions. The mean OSPAN scores for the final set of 96 high- and 96 low-span participants, respectively, were 25.2 (*SD* = 6.2, range 19–54) and 6.7 (*SD* = 2.1, range 0–10). Scholastic Aptitude Test (SAT) data from these participants are presented in Appendix B.

Memory Task

We have organized the memory-task results around two primary questions and two secondary questions. Our two primary questions were (a) Under no-load conditions, do high and low spans differ in their susceptibility to PI? and (b) How does an attentional load affect any span differences seen in PI susceptibility? Our secondary questions were (a) Under no-load conditions, do high and low spans differ in PI release? and (b) How does an attentional load affect any span differences seen in PI release? For all of the analyses reported hereinafter, the alpha level was set at .05.

Note that in order to address our primary questions, we conducted four planned comparisons that logically follow from the span and load findings from Rosen and Engle (1997). First, we tested for span differences in PI under no load, with the expectation that low spans would show larger PI effects than would high spans. Second, we tested for span differences in PI under load, with the expectation that the span groups would show equivalent PI effects. Third, we tested for a significant load effect on PI for high-span participants, with the expectation that load would increase PI susceptibility. Fourth, we tested for a significant load effect on PI for low-span participants, with the expectation that load would have no effect on their PI susceptibility.

PI Effects

The mean number of words recalled per list by span groups is presented in Figure 1 for no-load participants and Figure 2 for load participants. We restricted our analyses here and in Experiment 2 to correct recall, as opposed to intrusions, because intrusion rates were extremely low. A vast majority of participants in both experiments committed no intrusions on any list (which is not unusual for interference studies; see Melton & Irwin, 1940).

Span group and tapping load both influenced List 1 recall, with high spans recalling more words than did low spans, and cascade (no-load) tappers recalling more words than did complex (load) tappers. An analysis of variance (ANOVA) of List 1 recall scores confirmed these observations, indicating a main effect of span, $F(1, 188) = 8.54$, $MSE = 1.98$, and of load, $F(1, 188) = 22.74$, $MSE = 1.98$; the Span × Load interaction was not significant, $F(1, 188) < 1$. Because of these span and load effects on List 1, we calculated PI as a "proportion loss from List 1" measure.

Note that our primary interest here was in assessing span and load effects on PI, which reflects the *loss of recall from original, noninterference levels* that results from the progressive encoding and retrieval of similar study episodes. Thus, we calculated proportional PI effects for each participant, treating List 1 recall as baseline in the assessment of the effect of PI buildup. By analogy, if the stock market falls 500 points, one should be less compelled to leap from a high place if the previous baseline was 10,000 points (a 5% drop) than if it was 2,000 points (a 25% drop). We therefore should consider baseline List 1 recall in assessing the practical (as well as statistical) significance of PI buildup on the recall of subsequent lists. The proportional PI effect for List 2, for example, was calculated by subtracting the number of words recalled on List 1 from the number of words recalled on List 2, and dividing that by List 1 recall.

Examining the proportional PI effects depicted for no load in Figure 3, and for load in Figure 4, we can see that interference increased across Lists 2 and 3. Most important,
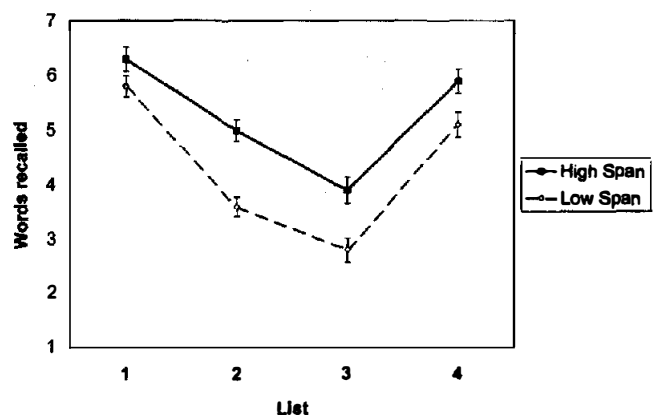

Figure 1. Number of words recalled across lists for high- and low-span participants under no-load ("cascade" tapping) conditions in Experiment 1. Vertical lines depict standard errors of the means.
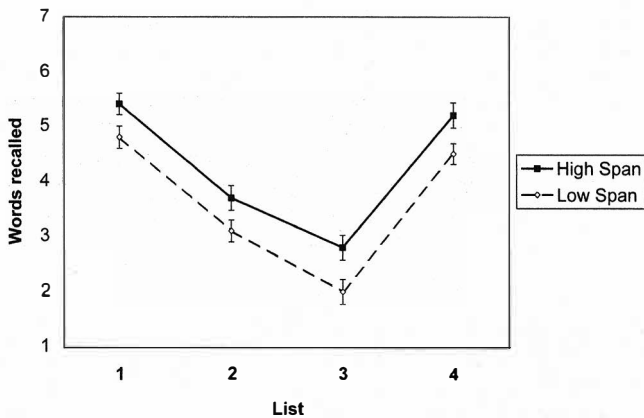
Figure 2. Number of words recalled across lists for high- and low-span participants under load ("complex" tapping) conditions in Experiment 1. Vertical lines depict standard errors of the means.

Figure 4. Proportional proactive interference (PI) effects on Lists 2–4 for high- and low-span participants under load in Experiment 1. Vertical lines depict standard errors of the means.

under no load, low spans showed larger proportional PI effects than did high spans. That is, low spans were more vulnerable to interference. However, under load, high and low spans showed similar PI effects, with both approximating those shown by low spans under no load. Thus, load increased high spans' PI susceptibility, making them as susceptible to PI as were low spans. In contrast, low spans' PI effects were quite similar under load and no-load conditions.

These impressions were confirmed by a 2 (span groups) × 2 (tapping loads) × 2 (lists) mixed-model ANOVA on proportional PI effects, with list as a repeated measures variable. We included only Lists 2 and 3 in this analysis because List 4 was the PI *release* trial. The span effect, $F(1, 188) = 8.91$, $MSE = 0.11$, and list effect, $F(1, 188) = 45.00$, $MSE = 0.08$, were significant. The load effect did not reach conventional significance, $F(1, 188) = 3.25$, $MSE = 0.11$,
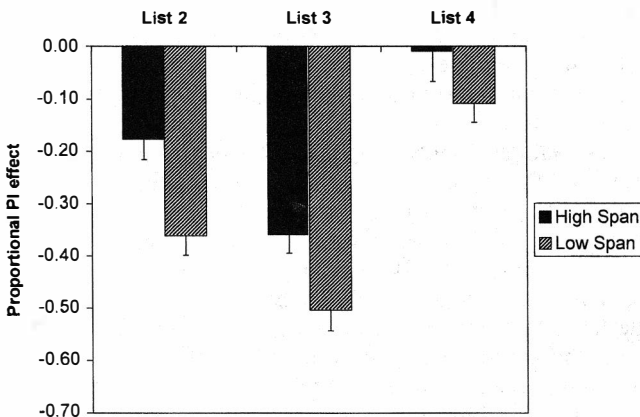

Figure 3. Proportional proactive interference (PI) effects on Lists 2–4 for high- and low-span participants under no load in Experiment 1. Proportional PI on each trial was calculated for each participant by subtracting the number of words recalled on List 1 from the number of words recalled on that trial's list, and then dividing by the number of words recalled on List 1. Vertical lines depict standard errors of the means.
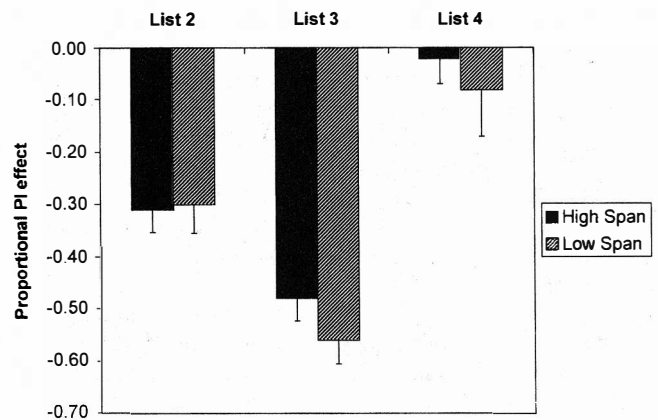
$p = .07$. Most important, however, span interacted with load, $F(1, 188) = 3.87$, $MSE = 0.11$, suggesting that tapping differentially affected the PI of high and low spans. Span did not interact with list, $F(1, 188) < 1$, nor did load, $F(1, 188) < 1$, and the three-way interaction of span, load, and list was not significant, $F(1, 188) = 1.41$, $MSE = 0.08$, $p = .24$.

We first explored the Span × Load interaction by comparing PI effects between high- and low-span participants under no load (cascade tapping). This analysis addressed our primary question of whether we would find significant span differences in PI under no-load conditions. These data are displayed in Figure 3. Low spans clearly showed larger PI effects than did high spans, $F(1, 94) = 14.98$, $MSE = 0.09$. Although PI effects varied significantly across lists, $F(1, 94) = 21.70$, $MSE = 0.06$, the Span × List interaction was not significant, $F(1, 94) < 1$.

The next analysis addressed whether high and low spans would differ in PI buildup under attentional load conditions. Here we expected the concurrent task to equalize the span groups: Load should impair attentional control for high spans and so increase PI susceptibility, but it should have little effect on low spans because they had (or used) little attentional control from the start. Thus, support for our hypothesis came from a null effect: As is evident in Figure 4, high and low spans demonstrated equivalent PI on both Lists 2 and 3. Neither the span effect, $F(1, 94) < 1$, nor the Span × List interaction, $F(1, 94) = 1.17$, $MSE = 0.10$, $p = .29$, reached significance.

Our final set of PI analyses examined the effect of divided attention in another way, by determining whether load significantly increased PI compared with no load for either high or low spans. Again, we hypothesized that load would impair high spans' control capabilities and thus increase their PI effects. However, load should have little effect on low spans' PI because even under no-load conditions they were not using controlled attention to prevent PI as effectively. Both hypotheses were confirmed. When comparing proportional PI between high spans under no load (cascade tapping) and under load (complex tapping), it is evident

from Figures 3 and 4 that load increased PI for high spans, $F(1, 94) = 9.29, MSE = 0.08$.

For low spans, we hypothesized that PI would be unaffected by load, similar to the findings of Rosen and Engle (1997). This was indeed the case; low spans showed equivalent PI under load and under no-load conditions, $F(1, 94) < 1$. Thus, whereas PI increased under load for high spans, PI was relatively immune to load for low spans. Low spans showed equivalent proportional PI effects under load and no-load conditions.

Although we would argue that our proportional analyses are justified and most appropriate for the questions of interest, for completeness we also present the corresponding PI analyses on the raw recall data in Appendix C. These analyses, and the means displayed in Figures 1 and 2, are generally in line with the conclusions drawn from the proportional PI data. However, the data are noisier, and not all comparisons of interest are statistically significant.

PI Release Effects

Neither span nor load affected PI release, even though we assessed release effects with two different comparisons: (a) the "proportion loss" for List 4 versus List 1 (smaller loss score indicated more release) and (b) the rebound of raw recall from List 3 to List 4 (larger rebound indicated more release). Under no load, high and low spans showed equivalent release. ANOVAs indicated no span differences in the List 1 proportion-loss measure ($Ms = -0.008$ and -0.108, respectively), $F(1, 94) = 2.02, MSE = 0.18, p = .16$, or in the Lists 3–4 rebound measure ($Ms = +1.96$ and $+2.25$, respectively), $F(1, 94) < 1$. Under load, high and low spans did not differ in PI release in either the proportion-loss measure ($Ms = -0.017$ and -0.081, respectively), $F(1, 94) < 1$, or the rebound measure ($Ms = +2.40$ and $+2.48$, respectively), $F(1, 94) < 1$. There were also no significant effects of load on PI release for either of the span groups considered individually.

Tapping Task

Unfortunately, for a significant number of participants the tapping data were unreliable. Gaps of seconds (or tens of seconds) were scattered throughout the data sets. This was equally true for high and low spans, and for cascade and complex tappers. Apparently the finger-pads were not equivalently sensitive to all participants' tapping. Because of this significant data loss, we were unable to analyze tapping performance. The problem was remedied in Experiment 2.

Trail-Making Task

Each participant performed four trials of the Trails task. For each trial, we summed the number of letters and numbers counted, and subtracted the number of errors. In order to establish a most stable measure of Trails performance, we then averaged the scores across the four trials to yield a single Trails score. Table 1 presents the mean Trails scores by span group and tapping condition, where it is clear

Table 1
Mean Trails Scores and Standard Deviations by Span Group and by Tapping Condition in Experiment 1

Span	Cascade (no load)		Complex (load)	
	M	SD	M	SD
High	15.77	3.32	12.34	4.08
Low	13.92	2.91	10.01	2.71

that high spans counted more items than did low spans, $F(1, 188) = 19.34, MSE = 10.88$. In addition, complex (load) tapping impaired counting compared with cascade (no-load) tapping, $F(1, 188) = 59.49, MSE = 10.88$. In contrast to the recall data reported above, however, the secondary load task appeared to have equivalent debilitating effects on high and low spans; the Span × Load interaction was not significant, $F(1, 188) < 1$.

Discussion

High- and low-working-memory span participants differed in PI susceptibility and in reaction to an attention-demanding concurrent task. High-span participants were less susceptible to PI than were low-span participants under no-load conditions. Under an attentional load, however, the PI effects for the two groups were equivalent. Divided attention increased PI for high spans, but it had no measurable effect on PI for low spans. These findings suggest that high spans use controlled attention to resist PI under normal conditions, but the attention-demanding tapping task prevents its use. The findings further suggest that the greater susceptibility of low spans to PI stems from an inability (or unwillingness) to use controlled attention to counteract the effect of PI on recall from long-term memory.

In contrast to the PI-task data, the Trails data showed that dividing attention hurt high and low spans equally, perhaps because the Trails task can be performed only through controlled processing. In the PI task, automatic spreading activation among related words may have allowed for some recall without effort or intention, particularly for low spans. In the Trails task, however, such automatic activation among representations should hurt performance. That is, participants in the Trails task must *avoid* counting along an automatic, practiced path of reciting two letters or two numbers in a row. We therefore hypothesized that low spans would count relatively few items, because of reduced capability to maintain active memory representations (i.e., the last letter counted) in the face of attention shifts away from those representations (i.e., to numbers in order to determine the next number to be counted). Moreover, it follows in an admittedly post hoc fashion that, in this task, a load should hurt even low spans' counting. In order to perform the task at all, any controlled attention capabilities that are available must check the more automatic response of continuing to count along a single stimulus category.

With respect to our questions of primary interest, however, some problems with Experiment 1 may limit our conclusions about working-memory capacity and PI. First,

by not using a true *no*-load condition as a baseline, we may have underestimated span differences in PI. High spans' PI effects were more sensitive to load than were low spans' effects. Therefore, the small "load" of maintaining cascade tapping may have affected high spans' control capabilities and so may have increased their apparent "no-load" PI effects.

Second, without tapping data it is difficult to interpret the noneffect of load on low spans' PI effects. Low spans may have shown no increase in PI under load because they traded off effort between memorizing and tapping. Low spans may have been increasingly vulnerable to PI under load, but instead of demonstrating this by recalling fewer words across lists, they may have tapped more slowly or less accurately (but parallel recall findings from Rosen and Engle, 1997, suggest that this is not the case). Both limitations were remedied in Experiment 2.

Third, we failed to replicate findings from Moscovitch (1994). Moscovitch found that a complex-tapping load had no effect on PI buildup but it virtually eliminated PI release, and we found just the opposite. A critical difference between these two experiments may have been the number of lists that each participant learned. Our participants saw only *one* series of four lists, and Moscovitch's participants saw *two* series of five lists (after studying and recalling a practice list). PI builds and asymptotes rapidly (e.g., Keppel & Underwood, 1962), and, anecdotally, we have found that PI is not completely released even after a category switch, a subsequent task switch, and 5 min of rest and listening to instructions (n.b., Underwood, 1957; but see Loess & Waugh, 1967).[1] Moscovitch's participants may have shown exaggerated PI on their second set of lists, regardless of load condition, and so his procedure may have been insensitive to load effects on PI.

EXPERIMENT 2

Experiment 2 was motivated in part by methodological concerns and the replicability of the PI findings from Experiment 1. Here we compared attentional load conditions with a no-load condition, in contrast to the "cascade" condition from Experiment 1. Participants also tapped their fingers on a computer keyboard so that we could collect reliable tapping data (see Martin et al., 1994). However, Experiment 2 also addressed two theoretical issues. The first was to examine the effects of dividing attention on encoding versus retrieval. The second was to determine whether the divided-attention cost in PI for high spans resulted from impairing control at encoding, at retrieval, or both.

Craik, Govoni, Naveh-Benjamin, and Anderson (1996) recently demonstrated that dividing attention during encoding of a word list impaired subsequent recall much more than did dividing attention during retrieval of a list (see also Baddeley, Lewis, Eldridge, & Thomson, 1984; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Park, Smith, Dudley, & Lafronza, 1989). Costs to memory performance due to divided attention at encoding ranged from 22% (in recognition) to 46% (in free recall). With divided attention at retrieval, costs ranged only from 1% (in recognition) to 13%

(in free recall). Clearly, encoding processes appear to be attention demanding and under conscious control, but what of retrieval?

The most compelling evidence that Craik et al. (1996) found for attentional processing during retrieval was that performance on the *secondary* task (a serial choice response-time task) suffered as much during memory retrieval as it did during encoding, if not more (see also Baddeley et al., 1984; Murdock, 1965). Moreover, secondary-task performance during retrieval was sensitive to emphasis instructions, with better performance under choice-task-emphasis instructions and worse performance under memory-emphasis instructions. Craik et al. thus concluded that when self-initiated retrieval demands were high, such as in free recall, some aspects of the retrieval process were controlled (although much less so than encoding processes). This was evident in the free-recall costs under load at retrieval. And, even when memory costs were absent during retrieval tasks such as recognition, the considerable secondary-task costs obtained indicate that retrieval cannot be fully automatic. Craik et al. speculated that retrieval is obligatory, in that it can proceed without much strategic control, and it is not easily disrupted by a secondary task. However, it is also attention demanding, in that the initiation of the retrieval process, or a "retrieval mode" (Tulving, 1983), requires significant resources and so limits performance on the secondary task. Retrieval "captures" attention.

Craik et al. (1996) also noted that during retrieval, "other components [of attention] may also be operating; for example, voluntary, strategic operations may elaborate and augment the retrieval information provided, especially when this information is minimal, as in free recall" (p. 175). These voluntary and strategic operations may be particularly important to the memory tasks at the focus of our present investigation. Indeed, retrieval during fluency tasks—which we argue requires interference resistance as well as strategic search—is highly sensitive to divided-attention requirements (Baddeley et al., 1984; Martin et al., 1994; Moscovitch, 1994; Phillips, 1997; Rosen & Engle, 1997). Anyone who has ever forgotten the last place they left their car keys (or car!) can certainly attest to the phenomenological "effort" required to retrieve information that is blocked by related, interfering episodes. If resisting interference is part of what demands attention during fluency tasks, then we should see increased PI when attention is divided during the retrieval of later lists in a PI task. At least, this should be true for high spans, who appear to use controlled attention to counteract interference.

---

[1] In fact, our Experiment 2 was initially conceived of as a within-subjects design, with all participants completing a trial under no load, a trial under encoding load, and a trial under retrieval load. However, even after using a different taxonomic category for each trial, and allowing a five-minute task-filled "break" between each trial, we found dramatic carry-over effects from trial to trial. We therefore manipulated load between subjects, giving each participant only one trial of three lists. We have found similar carry-over effects across category changes and filled breaks in the context of an unpublished directed-forgetting experiment.

Converging evidence that attentional processes at retrieval are important to PI resistance comes from recent research on cognitive inhibitory mechanisms (for reviews, see M. C. Anderson & Bjork, 1994; Dagenbach & Carr, 1994; Dempster, 1991, 1992; Zacks & Hasher, 1994). Such work has suggested that an integral aspect of selectively attending to objects in the environment is the active suppression of competing distractor objects (e.g., Houghton & Tipper, 1994; Neill, 1977). Moreover, these inhibitory processes are effortful and may be thwarted by an attentional load (e.g., Engle, Conway, Tuholski, & Shisler, 1995; Roberts, Hager, & Heron, 1994). M. C. Anderson and Neely (1996) further proposed that retrieval is an attentional act, that is, an *internally* focused selective-attention process. By this view, retrieval of target information is accompanied by the suppression of related, potentially competing, items stored in memory, just as is hypothesized to occur in externally focused attention. This suppression process reduces competition from interfering items in order that target information may be retrieved quickly and accurately (e.g., Dagenbach & Carr, 1994; Hasher & Zacks, 1988).

However, M. C. Anderson and Neely (1996) also suggested that, paradoxically, the retrieval process is responsible for producing interference. Interference is a pathological byproduct of a normally beneficial cognitive mechanism. That is, interference arises from inadvertently *retrieving* competing information that is similar to the target, rather than from merely *encoding* that information. PI susceptibility therefore results from retrieving prior episodes in an attempt to recover a more recent episode: When nontarget information is accidentally retrieved, it may inhibit the target's memory representation. Resistance to PI therefore comes from overriding this accidental inhibition from competitors by suppressing them.

Thus, on one hand, dual-task studies of encoding and retrieval suggest that encoding processes are the more controlled, or attention demanding. If this is so, then dividing attention during the encoding of each list in a PI task should significantly decrease recall and increase interference. On the other hand, studies from an inhibitory perspective suggest that retrieval is an attentional act, particularly in the face of interference, and this attentional inhibition can be derailed by a dual-task imposition. If this is so, then a concurrent load task at retrieval should impair recall, but perhaps only as interference builds across lists. In order to address these hypotheses, Experiment 2 tested high- and low-span participants in the PI buildup task from Experiment 1, but with a tapping task required either during encoding, retrieval, or not at all.

Method

Participants

Two hundred sixteen undergraduates from the University of South Carolina—Columbia (USCC) and Georgia State University (GSU) participated in Experiment 2 in return for psychology course credit, or in exchange for $10. These participants were identified from a larger pool who had participated in the OSPAN task (and a different pool than had participated in Experiment 1).

Between 1 and 90 days may have intervened between participation in the OSPAN task and the PI task.

Design

The design was a 2 × 3 × 3 mixed-model factorial, with span group (high, low) and tapping load (no load, encoding load, retrieval load) manipulated between subjects, and List (1–3) manipulated within subjects. The "tapping load" variable refers to whether participants tapped the complex sequence while encoding each list, while retrieving each list, or not at all. The "list" variable refers to the three word lists that were studied and recalled. Thirty-six participants in each span group were randomly assigned to one of the tapping conditions.

Apparatus and Materials

PI Task

USCC participants used the same computer apparatus from Experiment 1. GSU participants saw the stimuli presented via the same MEL 1.0 program, but on a Dell color monitor with a VGA graphics card.

The materials were very similar to those used in Experiment 1, but instead of using all four "subsets" from the *animals, occupations,* and *countries* categories, we used only three (corresponding to three buildup lists). We constructed nine complete stimulus sets of 30 words each (10 words for each of three lists) and used each set equally often across participants and tapping conditions. Each subset appeared equally often in each list position (List 1, 2, or 3) across participants and tapping conditions. As in Experiment 1, a given participant saw a single stimulus set of 30 items. Each set consisted of three lists from the same category.

Rehearsal Prevention Task

In Experiment 2 we used only three letter–number combinations (P–29, G–77, and U–45) because participants received only three memory lists. We created three trial orders from these combinations, with each letter–number combination appearing once in each PI-list position, and with each order used equally often across categories and across tapping conditions. The same letter–number combinations from Experiment 1 served as practice.

Tapping Task

All participants tapped on the "V," "B," "N," and "M" keys on the keyboard of a second computer, and taps were recorded by the same program used in Experiment 1.

Procedure

The procedure was identical to that in Experiment 1, with the following few exceptions.

The session began with the same sequence of tapping practice as did Experiment 1, except all participants practiced tapping the "complex" sequence (index finger—ring finger—middle finger—pinkie). All participants received on-line, auditory feedback on accuracy for the first three 30-s practice trials. In the fourth, 60-s practice trial, all participants received on-line response-time feedback, calculated individually as in Experiment 1. After hearing PI task instructions and practicing the Trails task, all participants practiced tapping for 30 s with response-time feedback. And, for all participants, the PI task then began with a green screen instructing

them to begin tapping. Participants tapped for 20 s with response-time feedback, calculated from the immediately preceding 30-s practice trial.

From this point the procedure differed in the three different tapping conditions: "No-load" participants stopped tapping after the green screen and never tapped again. "Encoding-load" participants tapped through the presentation of every word list but did *not* tap while recalling the lists or during the Trails task. "Retrieval-load" participants tapped only while recalling each list but *not* while studying the lists or during the Trails task. Both encoding-load and retrieval-load participants received on-line response-time feedback based on tapping speed during the final 30-s practice trial.

Practice, feedback, and instructions encouraged participants to consider the tapping task as the "primary" task and the memory task as the "secondary" task. In this way we hoped to focus any dual-task variability into performance on the memory task.

Results

Participants

We replaced the data from one low span and two high spans because of equipment malfunctions, and two low spans and one high span who were not native English speakers. The mean OSPAN scores for the final set of 108 high- and 108 low-span participants, respectively, were 26.0 ($SD = 6.3$, range 18–44) and 6.5 ($SD = 2.3$, range 0–9). Means and analyses of the SAT scores for these participants are presented in Appendix B.

Memory Task

We have organized the memory-task results around four questions: (a) Under a true no-load condition, do high and low spans differ in PI susceptibility? (b) How does an attentional load at encoding affect span differences in PI? (c) How does an attentional load at retrieval affect span differences in PI? and (d) How does an attentional load at encoding or retrieval affect either span group's recall on List 1, before any PI has built up?

Again, our questions required that we perform a number of planned contrasts on PI effects, all following from our Experiment 1 findings and from Rosen and Engle (1997). We first tested for span differences in PI under no load, load at encoding, and load at retrieval. We expected that low spans would show more PI than would high spans under no load but that the groups would show equivalent PI under load conditions. We then tested whether PI effects increased significantly under load at encoding and retrieval (vs. no load). As in Experiment 1, we expected that load would increase PI for high spans—to the level of low spans—but it would not increase PI any further for low spans. Finally, we tested whether load at encoding and retrieval had equivalent effects on PI.

The mean number of words recalled per list, by span group, is presented in Figure 5 for no load, Figure 6 for load at encoding, and Figure 7 for load at retrieval. As in Experiment 1, span and load influenced List 1 recall. ANOVA indicated that, overall, high spans recalled more words than did low spans on List 1, $F(1, 210) = 20.50$, $MSE = 1.71$, and no-load participants recalled more words
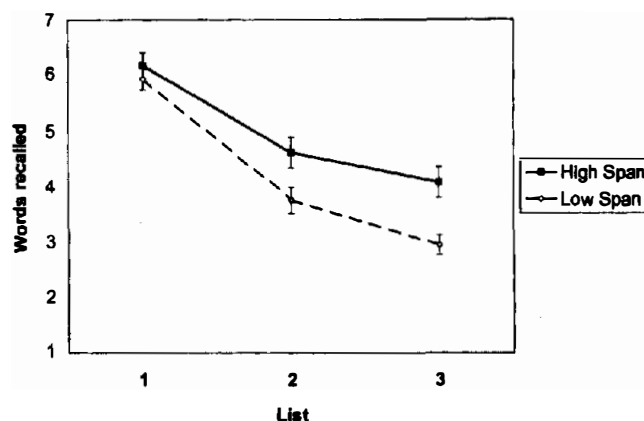

Figure 5. Number of words recalled across lists for high- and low-span participants under no load in Experiment 2. Vertical lines depict standard errors of the means.

than did participants under load, $F(2, 210) = 7.49$, $MSE = 1.71$. A "marginal" Span × Load interaction fell short of conventional significance levels, $F(2, 210) = 2.36$, $MSE = 1.71$, $p < .10$. Note, however, that under no load (Figure 5), high and low spans did not differ in List 1 recall ($Ms = 6.19$ and 5.94, respectively), $F(1, 70) < 1$; we expand on this finding later. Again, we were interested in the effects of span group and load conditions on PI, which represent changes from original recall, and so proportion-loss scores were obtained using individual participants' List 1 recall as a baseline. These proportional PI effects are presented in Figures 8, 9, and 10.

PI Effects

It is evident from Figures 8, 9, and 10 that under no-load conditions, at least, low spans showed larger proportional PI effects than did high spans, as in Experiment 1. However, under load at either encoding or retrieval, high and low spans showed equivalent PI. Also similar to Experiment 1, then, dividing attention equalized the performance of high-
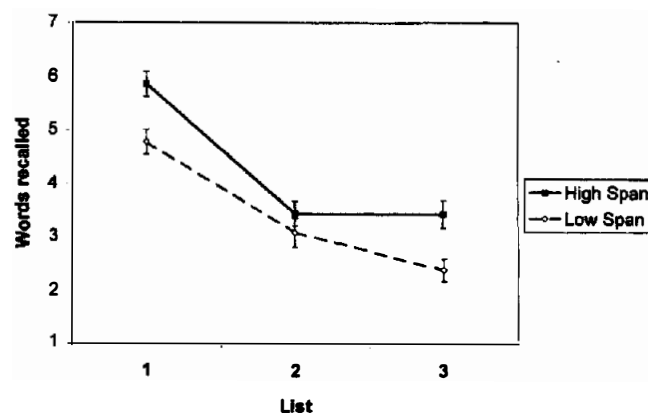

Figure 6. Number of words recalled across lists for high- and low-span participants under an encoding load in Experiment 2. Vertical lines depcit standard errors of the means.
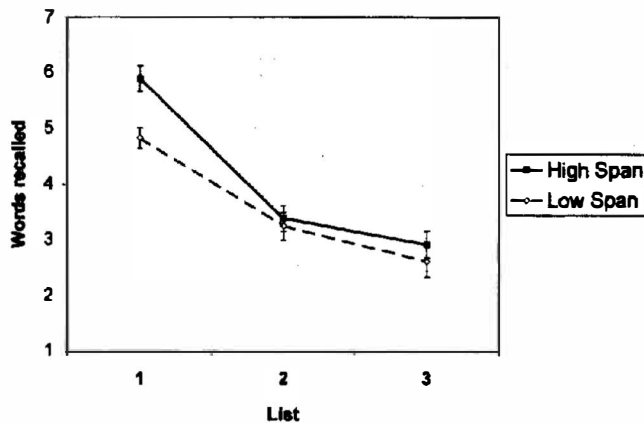
Figure 7. Number of words recalled across lists for high- and low-span participants under a retrieval load in Experiment 2. Vertical lines depict standard errors of the means.

Figure 9. Proportional proactive interference (PI) effects on Lists 2–3 for high- and low-span participants under an encoding load in Experiment 2. Vertical lines depict standard errors of the means.

and low-span participants under interference. Moreover, this group equivalence occurred because load at either encoding or retrieval increased PI for high spans. However, load at neither encoding nor retrieval affected PI for low spans, which remained relatively constant across load conditions.

A 2 (span) × 3 (load) × 2 (lists) mixed-model ANOVA confirmed these observations. Although neither the span effect, $F(1, 210) < 1$, nor the load effect, $F(1, 210) = 1.30$, $MSE = 0.13, p > .25$, was significant, there was a significant Span × Load interaction, $F(2, 210) = 4.44$, $MSE = 0.13$. Thus, load increased PI more for high spans than for low spans. The list effect was also significant, $F(1, 210) = 16.39$, $MSE = 0.06$, but the Span × List interaction, $F(1, 210) = 2.87, MSE = 0.06, p = .09$, the Load × List interaction, $F(2, 210) < 1$, and the Span × Load × List interaction, $F(2, 210) < 1$, were not.

Span differences under each load condition. We first explored the Span × Load interaction by testing for span differences in PI at each level of load. First, as can be seen in Figure 8, proportional PI effects under no load differed
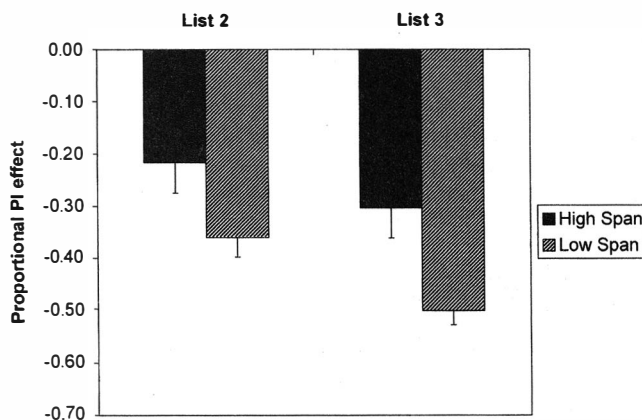
between span groups, as in Experiment 1. Low spans showed significantly more PI than did high spans. Under load at either encoding or retrieval, however, the span groups showed equivalent PI, as can be seen in Figures 9 and 10. Again, as in Experiment 1, load equalized the span groups in PI susceptibility, whether during the encoding or the retrieval of each list.

ANOVAs confirmed that under no load, the span effect was significant, $F(1, 70) = 9.65$, $MSE = 0.11$. In contrast, under load at encoding, the span effect was not significant, $F(1, 70) < 1$, nor was it significant under load at retrieval $F(1, 70) = 1.73$, $MSE = 0.11$, $p = .19$ (if anything, high spans showed slightly *larger* proportional PI effects than did low spans). With respect to list effects, under no load it was significant, $F(1, 70) = 8.64$, $MSE = 0.05$, but it did not interact with span, $F(1, 70) < 1$. Under load at encoding, the list effect fell just short of conventional significance levels, $F(1, 70) = 3.20$, $MSE = 0.05$, $p = .08$, as did the Span × List interaction, $F(1, 70) = 3.18$, $MSE = 0.05$, $p = .08$. Under load at retrieval, the list effect was significant, $F(1,$

Figure 8. Proportional proactive interference (PI) effects on Lists 2–3 for high- and low-span participants under no load in Experiment 2. Vertical lines depict standard errors of the means.
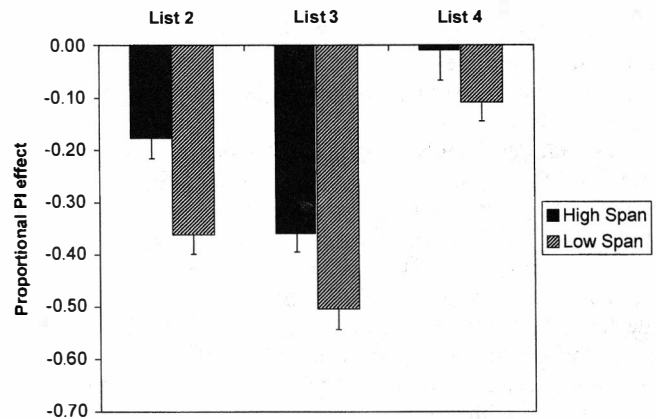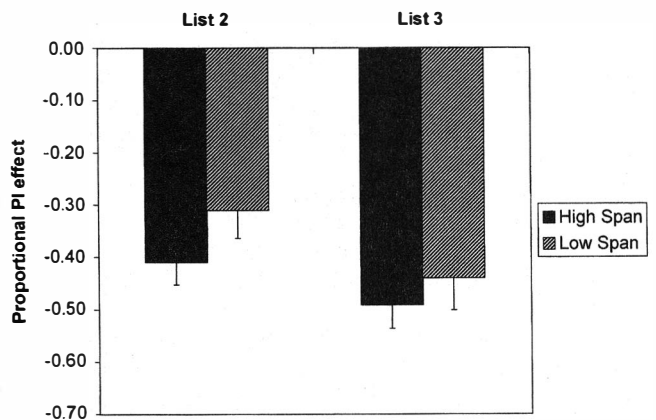

Figure 10. Proportional proactive interference (PI) effects on Lists 2–3 for high- and low-span participants under a retrieval load in Experiment 2. Vertical lines depict standard errors of the means.

70) = 5.38, MSE = 0.08, but it did not interact with span, $F(1, 70) < 1$.

Load effects for each span group. In parallel to Experiment 1, we also examined the divided-attention effect on PI for each span group separately by comparing each no-load group to its respective encoding- and retrieval-load groups. First, we consider high spans' load effects, as depicted in Figures 11 and 12 (these data were presented in the prior figures, but the comparisons of interest are more easily seen here). Load at either encoding or retrieval exacerbated PI for high spans. However, as can be seen in Figures 13 and 14 (again, presented for ease of comparison), low spans' PI effects were unaffected by load. Just as in Experiment 1, then, dividing attention made high spans, but not low spans, more vulnerable to interference.

ANOVAs indicated that for high spans, load at encoding increased proportional PI, $F(1, 70) = 4.28$, $MSE = 0.14$; list did not, $F(1, 70) = 1.05$, $MSE = 0.07$, $p > .30$, nor did list interact with load, $F(1, 70) = 1.04$, $MSE = 0.07$, $p > .30$. Load at retrieval also had a significant effect on PI, $F(1, 70) = 10.12$, $MSE = 0.13$. The list effect fell just short of conventional significance, $F(1, 70) = 3.72$, $MSE = 0.07$, $p = .06$, and the List × Load interaction was clearly not significant, $F(1, 70) < 1$. For low spans, load at encoding did not increase PI, $F(1, 70) < 1$. List did, $F(1, 70) = 16.88$, $MSE = 0.04$, but it did not interact with load, $F(1, 70) < 1$. Load at retrieval also did not increase PI for low spans, $F(1, 70) = 1.19$, $MSE = 0.10$, $p > .25$. List did increase PI for low spans, $F(1, 70) = 10.65$, $MSE = 0.07$, but list did not interact with load, $F(1, 70) < 1$.

We also contrasted the effects on PI of load at encoding or load at retrieval. As is evident in Figure 12, load at encoding and load at retrieval produced equivalent PI for high spans, $F(1, 70) = 1.24$, $MSE = 0.10$, $p > .25$ (the effect of list, and the interaction of list with load, were both nonsignificant with $ps > .30$). Similarly, for low spans (Figure 14), load at encoding and load at retrieval elicited showed equivalent PI, $F(1, 70) < 1$. The list effect was significant, $F(1, 70) = 9.12$, $MSE = 0.07$, but the List × Load interaction was not, $F(1,$

Figure 12. Proportional proactive interference (PI) effects on Lists 2–3 for high-span participants under no load, encoding (Enc) load, and retrieval (Ret) load in Experiment 2. Vertical lines depict standard errors of the means.

70) < 1. Dividing attention at encoding or retrieval therefore had equivalent effects (or for low spans, equivalent *noneffects*) on PI susceptibility.

Finally, for archival purposes, we again present the PI analyses on the raw recall data in Appendix D. As in Experiment 1, these analyses, and the means presented in Figures 5–7, 11, and 13, are generally consistent with the conclusions we have drawn from the proportion analyses. However, we note again that the raw recall data are noisy and that some comparisons of interest, although very close, do not quite reach conventional significance.

List 1 Recall

As already indicated above, the Span × Load interaction for List 1 recall approached, but did not quite reach, conventional significance ($p < .10$). However, we conducted further List 1 analyses on the basis of two prior

Figure 11. Number of words recalled across lists for high-span participants under no load, encoding (Enc) load, and retrieval (Ret) load in Experiment 2. Vertical lines depict standard errors of the means.

Figure 13. Number of words recalled across lists for low-span participants under no load, encoding (Enc) load, and retrieval (Ret) load in Experiment 2. Vertical lines depict standard errors of the means.
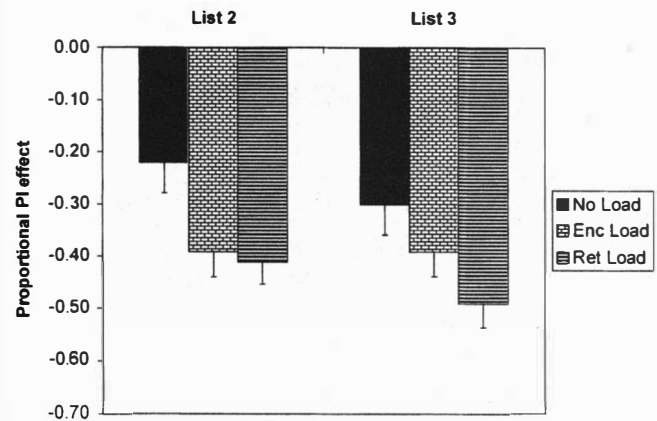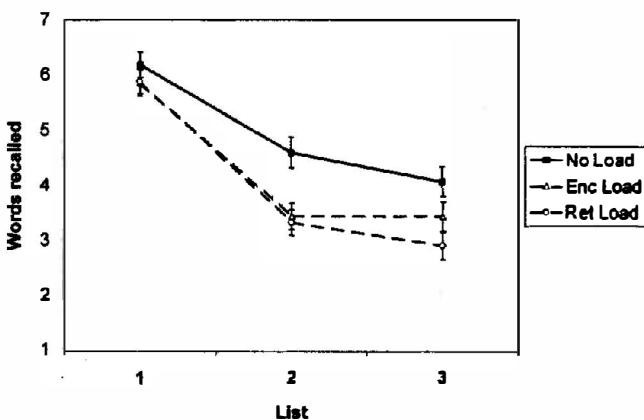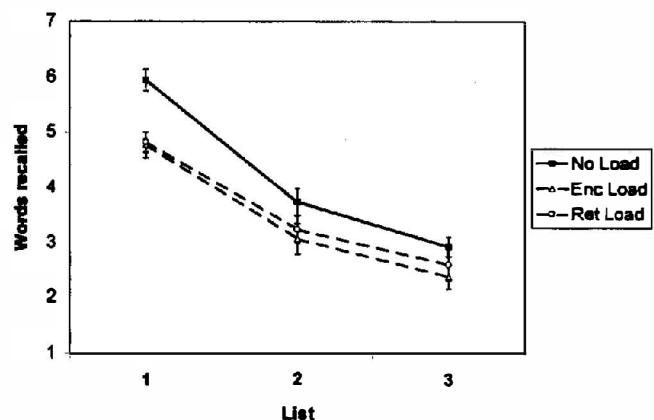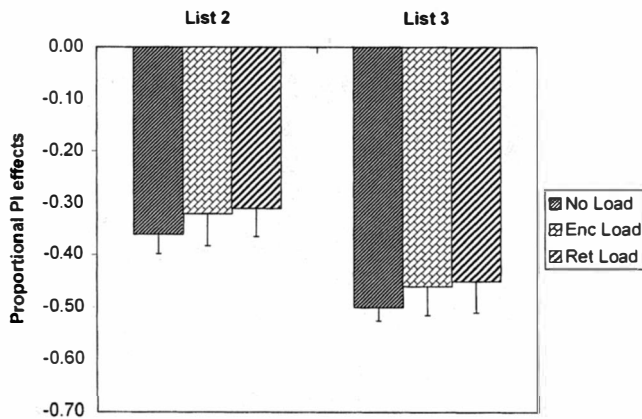
Figure 14. Proportional proactive interference (PI) effects on Lists 2–3 for low-span participants under no load, encoding (Enc) load, and retrieval (Ret) load in Experiment 2. Vertical lines depict standard errors of the means.

findings. First, a load at encoding typically impairs recall more than does a load at retrieval (e.g., Baddeley et al., 1984; Craik et al., 1996). Second, only high spans' PI effects appear to be affected by a secondary task (see also Rosen & Engle, 1997).

Figure 11 shows that under no load, high and low spans did *not* differ in List 1 recall. Thus, in contrast to the PI findings discussed above, high spans were relatively immune to load effects as long as the interference potential was low. However, the different load tasks affected List 1 recall differently for the two span groups. For high spans, neither load at encoding nor load at retrieval significantly reduced List 1 recall. Under no load, high spans recalled 6.19 words from List 1, whereas under load at encoding and retrieval, they recalled 5.86 and 5.89 words, respectively. These load effects were not statistically significant. For load at encoding, $F(1, 70) = 1.05$, $MSE = 1.91$, $p > .30$; for load at retrieval, $F(1, 70) < 1$.

For low spans, however, load at either encoding or retrieval significantly, and equivalently, reduced recall of List 1 (see Figure 13). Under no load, low spans recalled 5.94 words on List 1, but under load at encoding and retrieval, they recalled only 4.78 and 4.83 words, respectively, an approximate 20% loss. These load effects on List 1 were significant. For load at encoding, $F(1, 70) = 14.77$, $MSE = 1.66$; for load at retrieval, $F(1, 70) = 17.12$, $MSE = 1.30$. Moreover, List 1 recall under load at encoding was equivalent to that under load at retrieval, $F(1, 70) < 1$, suggesting that low spans used attention equally for encoding and retrieving List 1. Thus, whereas load at neither encoding nor retrieval affected the PI for low spans, both load conditions impaired recall when interference was relatively low.

Tapping Task

For the tapping task, we analyzed encoding-load and retrieval-load participants' mean number of correct taps per second across each tapping period (i.e., across the relevant

portion of each of the three recall trials and the two tapping-practice trials). Thus, higher means reflect better performance. Table 2 presents these data.

Tapping Practice

First, we examined tapping practice effects on the two trials for which we had data—the 60-s practice trial with response-time feedback, and the 20-s tapping trial that immediately preceded the presentation of the first memory list.[2] From Table 2 it is fairly clear that high and low spans tapped equally well, and that tapping improved, during practice. Moreover, tapping was equivalent between encoding-load and retrieval-load participants (at this point in the experiment all participants were treated identically). Finally, low spans benefited more from practice than did high spans. A 2 (spans) × 2 (tapping loads) × 2 (trials) mixed-model ANOVA supported these conclusions. Overall tapping for high and low spans was statistically equivalent, $F(1, 140) < 1$, and locus of load had no effect, $F(1, 140) < 1$. Tapping did, in fact, improve significantly from the 60-s practice trial to the subsequent 20-s trial, $F(1, 140) = 74.29$, $MSE = 0.10$. Furthermore, the Span × Trial interaction was significant, $F(1, 140) = 5.50$, $MSE = 0.10$, indicating that low spans improved more with practice than did high spans (M change = +0.42 and +0.24 correct taps per second, respectively).

Tapping and Memory

Next, we examined the effects of dividing attention on tapping under *non*-PI conditions by testing for span and load effects on the 20-s tapping trial (immediately before the memory task began) versus Trial 1 of the PI task. This comparison should indicate the effect of encoding or retrieving a word list on concurrent tapping. As can be seen in Table 2, tapping was impaired by the onset of the memory task, and equally so for encoding and retrieval. Moreover, low spans were more impaired by the onset of the memory task than were high spans. A 2 (spans) × 2 (tapping loads) × 2 (trials) mixed-model ANOVA confirmed these observations. Neither span nor the locus of load impacted tapping overall. For span, $F(1, 140) < 1$; for load, $F(1, 140) < 1$. However, tapping deteriorated significantly from the tapping-only trial to the tapping-plus-memory trial, $F(1, 140) = 29.00$, $MSE = 0.17$. The Span × Trial interaction was significant, $F(1, 140) = 4.22$, $MSE = 0.17$, indicating that low spans were more disrupted with the onset of the memory task than were high spans (M change = −0.36 and −0.16 correct taps per second, respectively). However, both groups' tapping was significantly disrupted. For high spans, $F(1, 70) = 7.59$, $MSE = 0.13$; for low spans, $F(1, 70) = 21.82$, $MSE = 0.22$. Thus, the greater effect of load on List 1 recall for low spans than for high spans would only have increased if low spans had maintained the tapping performance of high spans.

---

[2] One 30-second tapping session intervened between these two trials.

Table 2
Mean Numbers and Standard Deviations of Correct Taps Per Second, by Span Group and Trial, for Encoding Load and Retrieval Load Conditions in Experiment 2

| | Trial | | | | | | | | | |
| | Prac60 | | Base20 | | Rec1 | | Rec2 | | Rec3 | |
Load	M	SD	M	SD	M	SD	M	SD	M	SD
Encoding										
High spans	2.45	0.57	2.70	0.81	2.49	0.71	2.88	0.83	2.73	0.78
Low spans	2.26	0.69	2.71	0.72	2.29	0.77	2.53	0.85	2.60	0.86
Retrieval										
High spans	2.49	0.47	2.72	0.52	2.61	0.74	2.70	0.72	2.80	0.64
Low spans	2.41	0.73	2.79	0.69	2.48	0.99	2.52	1.08	2.62	1.11

Note. Prac60 = the 60-s practice trial with response-time feedback; Base20 = the 20-s baseline tapping trial immediately before onset of the PI-task stimuli; Rec1–3 = PI-buildup trials 1–3; PI = proactive interference.

Tapping and PI

We then analyzed tapping across the three trials of the PI task in order to determine the effects of PI on tapping. To the degree that memory under increasing interference was increasingly attention demanding, tapping should have deteriorated across trials. However, tapping actually improved, perhaps indicating that with practice the tapping task became more automatized, both during list encoding and retrieval. Most important, however, high and low spans improved *equivalently* across trials in both load conditions, and so the span differences we observed in the effects of load on PI *cannot* be attributed to high and low spans differentially trading off effort between the memory and tapping tasks. Both groups improved their tapping across trials, but tapping load increased PI effects in memory for only high spans. Compared with high spans, low spans differentially slowed down their tapping at only the *onset* of the memory task (List 1), suggesting a greater accommodation for the mental work of encoding and retrieving a list *in the absence of PI*. However, given that this span difference did not increase across lists, low spans cannot have avoided a load effect on PI by making additional trade-offs to accommodate the *extra* mental work of combating PI.

A 2 (spans) × 2 (tapping loads) × 3 (trials) mixed-model ANOVA indicated that high and low spans tapped equivalently well across word lists, $F(1, 140) = 2.10, MSE = 1.90$, $p = .15$, and tapping did not differ between concurrent encoding and retrieval, $F(1, 140) < 1$. Tapping did improve significantly across Trials 1 to 3, $F(2, 280) = 14.67, MSE = 0.14$. However, as noted above, it is most important that high and low spans improved *equivalently* across trials for both encoding and retrieval conditions: The Span × Trial interaction was not significant, $F(2, 280) < 1$, nor was the Span × Load × Trial interaction, $F(2, 280) < 1$.

As an additional test of whether high and low spans differentially traded off effort between tapping and memory, we examined tapping across Trials 1–3 as a proportional change from the 20-s practice. The significant Span × Trial

interactions discussed above for practice indicated that initial tapping differences should be considered in interpreting change across PI trials, much in the same way that we considered List 1 recall in evaluating PI effects in memory. Thus, we transformed each participant's tapping score on Trials 1–3 to a proportion-loss measure from their 20-s baseline tapping score. A 2 (spans) × 2 (tapping loads) × 3 (trials) mixed-model ANOVA indicated that high spans showed less of an overall disruption in tapping than did low spans, $F(1, 140) = 6.80, MSE = 0.17$, and that tapping disruptions decreased as trials proceeded, $F(2, 280) = 12.75, MSE = 0.02$. Most important, again, the decrease in dual-task disruption on tapping was equivalent for high and low spans. The Span × Trial interaction was not significant, $F(2, 280) < 1$, nor was the three-way interaction of Span × Load × Trial, $F(2, 280) < 1$. Thus, both the raw tapping data and the proportion-change data suggested no evidence that high and low spans differentially traded off their tapping and memory performance as PI built up across memory trials.

Trail-Making Task

Trails scores were calculated as in Experiment 1. No participants tapped during the Trails task, so we collapsed over load condition. A one-way ANOVA indicated that high spans ($M = 15.29$) counted more items than did low spans ($M = 13.03$), $F(1, 214) = 28.78, MSE = 9.58$. Thus, we again found that high and low spans differed in a task that required maintaining some information in memory while only intermittently attending to that information.

Discussion

As in Experiment 1, high and low spans varied in PI susceptibility, and in response to a concurrent load task. Under single-task conditions, low spans demonstrated significantly larger PI effects than did high spans. Low spans' increased susceptibility to PI in Experiment 1 was therefore not an artifact of our having used a "cascade-tapping" condition as a baseline. Under an attentional load, moreover, the span groups showed equivalent PI, and this was true whether attention was divided during list encoding or retrieval. High spans tested under a load at either encoding or retrieval demonstrated increased PI relative to those tested under no load. In contrast, low spans tested under load and no load showed equivalent PI.

Our findings thus suggest that, at least for high-span participants, encoding and retrieving under conditions of interference are controlled and attention-demanding processes. That is, high spans use controlled processing when memory is taxed by interference. For low-span participants, however, encoding and retrieval may be no more attention demanding in the presence than in the absence of interference. Indeed, the lack of controlled processing by low spans may be responsible for their increased PI susceptibility under normal conditions.

The recall data from List 1 of the memory task provided a surprising counterpoint to the PI data and suggest that, in the

absence of significant interference, high spans encode and recall associatively related word lists relatively automatically. High spans showed no effect of load on List 1 recall, either for encoding or retrieval. However, neither their encoding nor retrieval was completely automatic, because they demonstrated a significant tapping decrement at the onset of List 1 in the memory task. Thus, although high spans could encode and retrieve List 1 unencumbered by the load task, these processes were attention demanding, in that their initiation and maintenance disrupted the tapping task.

Unlike high spans, low spans' dual-task costs on List 1 were reflected in both memory and tapping performance. They could not maintain their encoding or retrieval of List 1 while tapping, nor could they maintain their tapping while encoding or retrieving List 1. Thus, low-span participants showed the opposite response to load on List 1 than they had under PI, as did high spans (but in the opposite direction). On List 1, a load at either encoding or retrieval significantly reduced recall, but load did not affect low spans' subsequent interference susceptibility.

GENERAL DISCUSSION

We have argued elsewhere that working-memory span tasks measure an individual's ability to maintain information in an active state while only intermittently attending to it. Moreover, this general "executive" capability drives the correlations between span measures and a broad variety of cognitive tasks (see Engle, 1996; Engle, Kane, & Tuholski, 1999). Working-memory span should therefore predict the ability to keep information accessible for use in situations where focused attention is taxed, such as under conditions of distraction or interference. We hypothesized that working-memory capacity would be related to interference susceptibility in a PI task. Furthermore, if attentional control is important to recall under interference, then dividing attention should impair the ability to overcome interference. We therefore tested some participants under dual-task conditions.

In two experiments, working-memory capacity did, indeed, predict PI susceptibility. Under standard conditions, low spans demonstrated larger interference effects than did high spans, whether the "standard" context represented a single-task, no-load condition (Experiment 2), or a dual-task condition with minimal attentional demands (Experiment 1). These findings are consistent with prior work using both a Sternberg (1966) probe-recognition procedure (Conway & Engle, 1994) and a verbal fluency procedure (Rosen & Engle, 1997). The evidence is now quite convincing that compared with high spans, low spans are more vulnerable to memory interference in several of its manifestations (i.e., fan-effect/cue-overload interference, output interference, and proactive interference).

In addition to establishing span differences in PI susceptibility, the present experiments suggest that high and low spans engage attention differently when encoding and retrieving information under interfering versus noninterfering conditions. The fact that high spans were more vulnerable to PI when attention was divided suggests that they normally

use controlled attention to combat interference (see also Rosen & Engle, 1997). Moreover, Experiment 2 showed that a concurrent-load task at either the encoding or retrieval of each list increased PI for high spans. Thus, for high spans, attention may operate at either encoding or retrieval to resist PI. Low spans, in contrast, showed equivalent PI regardless of the concurrent-task requirements, suggesting that they normally do not engage in any more controlled processing under interference than they do under noninterference conditions. Furthermore, Experiment 2 indicated that low spans used attention at neither encoding nor retrieval to combat PI. This lack of attentional deployment may account for their normally higher PI susceptibility.[3]

We have based the discussion above on our analyses of proportional PI effects, and we believe these are justified. We also note, however, that analyses based on raw recall scores, although consistent with the proportional analyses in direction, indicated somewhat fewer statistically significant comparisons. For example, in Experiment 1 the raw PI effect for high spans under load was not significantly greater than that under no load. In Experiment 2, high and low spans under no load showed different PI susceptibility only when comparing Lists 1 and 3, whereas high spans under no load showed less PI than those under load and encoding or retrieval only when comparing Lists 1 and 2.

However, we are not particularly concerned about our general conclusions. The proportional analyses were very clear, and most critical comparisons were significant in the raw recall analyses (and most of those that were not significant were very close). Moreover, the basic findings from Experiment 2 replicated those from Experiment 1, and the basic conclusions from both experiments parallel those from an entirely different task, semantic fluency (Rosen & Engle, 1997). Moreover, in a combined analysis of raw recall data from Experiments 1 and 2 (combining load participants from Experiment 1 with both encoding and retrieval load participants from Experiment 2), low spans under no load showed significantly greater PI than did high spans under no load, $F(2, 332) = 4.84$, $MSE = 1.66$. Under load, high spans actually showed slightly larger PI effects than did low spans, $F(2, 476) = 2.54$, $MSE = 1.54$, $p = .08$. Load significantly increased raw PI effects for high spans, $F(2, 404) = 3.74$, $MSE = 1.76$, whereas it significantly reduced PI for low spans, $F(2, 404) = 3.43$, $MSE = 1.41$.

We do not have an explanation for a significant reduction in PI effects for low spans under load compared with no load

---

[3] In Experiment 2, dividing attention at either encoding or retrieval appears to have had the same detrimental effect on high spans as dividing attention at both encoding and retrieval in Experiment 1. We are not sure whether this under-additivity represents a real phenomenon or not. Recall that the no-load "cascade" condition from Experiment 1 may have provided a small load for high spans, but the no-load condition from Experiment 2 did not. Thus, comparing the Experiment 1 and Experiment 2 load effects may underestimate the differences between them. We are therefore hesitant to venture an explanation for this under-additivity until it has been demonstrated more convincingly.

(a nonsignificant trend that appears in other analyses as well). One possibility is that floor effects may be responsible; that is, given the fairly low recall for low spans overall, there may be insufficient sensitivity to detect an increase in PI under load. In Experiment 2, low spans under no load recalled 5.94, 3.75, and 2.94 words in Lists 1–3, respectively. Low spans recalled 4.78, 3.08, and 2.39 words under load at encoding, and 4.83, 3.25, and 2.61 words under load at retrieval.

We make the following arguments against a floor effect explanation for the lack of load effects on low spans' PI. First, if low spans under load were at floor in recall of List 3, one would expect the standard deviations at List 3 to be smaller than for other lists, or smaller than those for other participant groups. The standard deviations for List 3 recall for low spans under load at encoding and retrieval are 1.31 and 1.68, respectively. These values are actually *larger* than those for low spans' List 3 recall under no load (1.07). Moreover, they are also equivalent to the standard deviations for low spans' retrieval on List 1, which are clearly free from floor effects (SDs = 1.17, 1.40, and 1.11 for no load, encoding load, and retrieval load, respectively). They are also equivalent to high spans' standard deviations on List 3 recall (SDs = 1.66, 1.56, and 1.44 for no load, encoding load, and retrieval load, respectively), and high spans *did* show significant load effects on PI. As a further argument against floor effects, we note that all of our conclusions about span and load may be drawn from comparisons between only Lists 1 and 2. The mean recall on List 2 in no condition drops below three items, and it is clearly a stretch to consider recall rates over 30% to be at floor. Indeed, it might seem a stretch of any boundary effect explanation to suggest that even the List 3 data, with the lowest mean recall at 24%, are at floor.

Interference Resistance at Encoding and Retrieval

How, exactly, did high spans use controlled attention to counteract PI? As we discussed previously, M. C. Anderson and Neely (1996) argued that overcoming PI involves active suppression of competitors at retrieval. High spans may therefore have used controlled attention to inhibit competition from prior list items but were foiled in doing so under retrieval load. Indeed, many theories assume that interference effects reflect a momentary disruption of retrieval (e.g., J. R. Anderson, 1983; M. C. Anderson, Bjork, & Bjork, 1994; Mensink & Raaijmakers, 1988; Nelson, Schreiber & McEvoy, 1992; Postman & Underwood, 1973). By these views, interference may arise under certain encoding conditions, but it does not stem from an encoding failure per se. In fact, several experiments have directly tested whether the interference in PI buildup tasks affects list encoding or retrieval (e.g., Loftus & Patterson, 1975; Watkins & Watkins, 1975). These found typical PI effects on immediate recall tests, with late-list items recalled less accurately than early-list items. However, on a delayed final test, after all lists were studied and tested, late-list items were as accurately recalled as were early-list items. The fact that

early-list and late-list items could be equivalently recalled at a later time indicated that all items had been equivalently encoded, whether in the presence or absence of interference from competitors. Our findings likewise suggest that PI creates temporary *retrieval* failures, and moreover, that those retrieval failures are exacerbated, for high spans, under divided attention.

However, we also found that high spans' PI increased under a secondary load at *encoding*. We see at least two possible explanations for this result that are not mutually exclusive. First, and consistent with an inhibitory view of interference resistance, high spans under no load may have blocked, or inhibited, prior lists as each subsequent list was presented for study. A large literature in "directed forgetting" demonstrates that explicit instructions to forget a prior list before studying a new list increase recall of that new list, often to the level of single-list recall (for reviews see Golding & MacLeod, 1998). Thus, the "forget" cue effectively eliminates PI, and it may do so through an inhibitory process in which prior list items are suppressed below a resting baseline level of accessibility (Geiselman, Bjork, & Fishman, 1983). Indeed, consistent with an inhibition interpretation, "forgotten" items are only temporarily inaccessible and rebound to normal levels of accessibility under certain testing conditions (e.g., Bjork et al., 1973; Bjork et al., 1984). Work with the negative priming task has already suggested that higher working-memory capacity affords more efficient attentional inhibition (Conway, Tuholski, Shisler, & Engle, 1999; Engle, Conway, et al., 1995), and so an inhibitory explanation of encoding-load effects for high spans seems tenable.

An inhibitory interpretation may be even more appropriate in light of recent paired-associate data collected by Rosen and Engle (1998). High and low spans in an interference condition learned an A–B list, then an A–C list, and then relearned the A–B list. Control participants learned an E–F list, then a C–D list, and then an A–B list. All participants were equated at high levels of learning for Lists 1 and 2. The measure of interest here was response time on the first learning–relearning trial for the A–B List 3. If high spans in the interference condition suppressed List 1 during learning of List 2, then they should have responded slowly to these List 1 items in trying to relearn them on List 3. In fact, high spans in the interference condition were significantly slower on List 3 than were their controls (M difference = 114 ms). A within-subjects analysis also showed that interference participants were much slower on List 3 than they had been with these same pairs on List 1. In contrast, low spans in the interference condition responded significantly *faster* on List 3 than did their controls (M difference = 106 ms), and they were not significantly slower than they themselves had been on List 1. Thus, only high spans showed evidence of inhibiting interfering information when attempting to learn new, related information.

A second possible explanation for high spans' attentional engagement during encoding under interference is a noninhibitory one. Some theorists have proposed that PI results from a list discrimination failure (e.g., Underwood &

Ekstrand, 1966, 1967; Winograd & Smith, 1966). Particularly in PI buildup tasks where lists are presented in rapid succession, participants might, in theory, confuse the lists in which specific stimuli had appeared; the more lists are presented, the greater the confusability (Loftus & Patterson, 1975; Wickens & Gittis, 1974). This confusion would lead participants to incorrectly recall prior-list stimuli when trying to report only the current-list stimuli. By this view, high spans may have used controlled attention at encoding to establish a "time tag" (Yntema & Trask, 1963) to each successive list, which then served as a discriminative retrieval cue.

Although a reasonable hypothesis, it is contradicted by a small set of studies indicating that list discrimination is actually quite good in PI buildup tasks (Dillon & Thomas, 1975) and that manipulations designed to facilitate discrimination do not reduce PI (e.g., Dillon, 1973). However, even if list discrimination does not worsen across lists, there is evidence that PI impairs some aspect of encoding quality. In AB–AC list-learning tasks, for example, participants encode the pairs in each subsequent list with less stable elaborators than on prior lists (e.g., Hasher & Johnson, 1975; Keppel, 1968). Thus, high spans may have used controlled attention to generate a unique encoding context for each list.

The Attentional Demands of Retrieval

The PI findings from Experiment 2 certainly suggest some boundary conditions for producing, or detecting, relatively controlled *retrieval* from long-term memory. Simple list-learning experiments have found retrieval to be minimally disrupted by a secondary task; in contrast, encoding may be severely impaired under divided attention (e.g., Baddeley et al., 1984; Park et al., 1989). However, *when averaged across many successive lists that may induce PI,* free recall may produce larger divided-attention costs at retrieval than do cued-recall and recognition tests (Craik et al., 1996). This is especially true in free-recall tests like verbal fluency, when the recall periods in these tasks are long enough to allow for significant output interference (Baddeley et al., 1984; Martin et al., 1994; Moscovitch, 1994; Phillips, 1997; Rosen & Engle, 1997). Indeed, along with these fluency studies, our Experiment 2 suggests that the potential for memory interference may be a sufficient, if not necessary, condition for detecting substantial divided-attention costs at retrieval (assuming that one's experimental participants are high spans).

But why might low spans use controlled attention to encode and retrieve when the potential for interference is low (i.e., on List 1) and then not use controlled attention when interference is high? We speculate that the control capabilities of low spans are so impoverished that List 1 encoding and retrieval processes alone push control to its limit. Low spans may therefore have no capacity remaining to modify or supplement processing in the face of the interference Lists 2 and 3. If this is so, what remains to be explained is why low spans should require so much effort to encode and retrieve even a single categorized list.

We suspect that the answer lies in the specifics of the stimuli used in our experiments. Although the words all came from well-learned categories, we avoided including words that were normatively among the top 12 associates for each category (e.g., in "Animals," we excluded *dog, cat, horse, cow, lion, tiger, elephant, pig, bear, mouse, rat,* and *deer*). Recall that the fluency studies by Rosen and Engle (1997) suggested that low spans dedicate significant attentional processing to monitoring their output for repetitions and intrusions. Perhaps, here, low spans used control capabilities to prevent intrusions of highly dominant category exemplars, both while encoding and retrieving List 1. That is, given the implicit cue "animals" during study and recall of List 1, low spans may have had difficulty ignoring the high-probability exemplars that were not presented on the list. Moreover, if low spans invested control to prevent intrusions on List 1, then they were less likely to be able to modify processing under interference conditions on List 2 and List 3, where the potential for covert intrusions increased.

Of course, a related question is how high spans were able to effectively encode and recall a single word list without significant use of controlled attention. Again, the answer may lie in our categorized word lists. With such obvious and strong associative links among the words in the first list, high spans may have been able to do minimal mental "work" in encoding and retrieving these items, relying largely on automatic spreading activation among the studied associates. Consistent with Rosen and Engle (1997), perhaps high spans, but not low spans, were able to completely block the high-dominance exemplars from coming to mind, and so did not need to expend attention on monitoring output for extralist intrusions. Indeed, such blocking or inhibition may have been reflected by the significant reduction in tapping performance for high spans on List 1.

Working Memory Capacity and the Prefrontal Cortex

Our findings of working-memory-span differences in PI susceptibility are consistent with the idea that central-executive processes are particularly reliant on prefrontal cortex (PFC) functioning (e.g., Baddeley, 1996; Duncan, 1995; Shallice & Burgess, 1993). Working-memory/central-executive tasks are widely found to be impaired after PFC damage in both macaques and humans, and imaging and single-unit recording studies show that PFC activation increases with the working-memory demands of a task (for reviews see Fuster, 1989; Engle, Kane, & Tuholski, 1999; Goldman-Rakic, 1987). Similar to Rosen and Engle (1997), moreover, we found that high and low spans differed significantly in a task, PI buildup, that is sensitive and somewhat selective to PFC injury and activation (e.g., Shimamura, Jurica, Mangels, Gershberg, & Knight, 1995; Uhl, Podreka, & Deecke, 1994). Our findings thus support the idea that encoding and retrieval under certain conditions, particularly those of interference, rely on PFC circuits (see Moscovitch, 1992a, 1992b). These results also contribute to a small but growing body of work suggesting that normal

individual differences in working-memory capacity may serve as a "human model" for levels of PFC functioning (e.g., Engle, Kane, & Tuholski, 1999; Rosen & Engle, 1997).

A look toward neuroscience may additionally help us interpret our most unusual finding. Namely, in contrast to much prior work (e.g., Baddeley et al., 1984; Craik et al., 1996), we found that divided attention at retrieval reduced recall *equivalently* to divided attention at encoding, whether for low spans on List 1, or for high spans on Lists 2 and 3. We note that recent neuroimaging research has suggested different broad patterns of PFC activation during episodic encoding and retrieval, and such findings have lead to a descriptive model called HERA ("Hemispheric Encoding/ Retrieval Asymmetry"; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). HERA holds that left PFC areas are particularly active during encoding processes (e.g., Kapur et al., 1994; Kapur et al., 1996; Shallice et al., 1994). In contrast, right PFC areas are particularly active during retrieval (e.g., Moscovitch, Kapur, Kohler, & Houle, 1995; Nyberg et al., 1996; Shallice et al., 1994). We speculate that HERA may be relevant here because our concurrent tapping task, a novel motor sequence, may have targeted the right PFC areas involved in retrieval more so than the left PFC areas involved in encoding.

All of our participants tapped the "complex" finger sequence with their nondominant hands. We used this procedure in an effort to make the tapping task as attention demanding and nonautomatic as possible. However, a potential side effect of this decision was that left-handed tapping—the nondominant hand for most of our participants—should have relied more on right PFC areas than left (Frith et al., 1991; Jenkins et al., 1994). Tapping may therefore have interfered with retrieval, which is a predominantly right PFC function, more significantly than with encoding, which according to HERA is a predominantly left PFC function. Indeed, Towell, Burton, and Burton (1994) found that encoding lists of verbal materials was more disruptive to repeatedly tapping a single finger with the right hand than with the left hand: Maintaining performance on a novel left-hemisphere task was more difficult when performed with encoding, which is also a left-hemisphere task. Although a significant divided-attention cost would likely have arisen at retrieval even with a right-hand tapping task, the divided-attention cost at encoding may have been larger, and therefore may have been larger than the retrieval effect. Future research should examine the differential effects of putative right-hemisphere versus left-hemisphere secondary tasks on encoding and retrieval processes.

Conclusions

In two experiments, individuals of high and low working-memory capacity were tested in a buildup of proactive interference task, with some participants tested under single-task, and others tested under dual-task, conditions. Under single-task conditions, high working-memory span participants showed less susceptibility to PI than did low-span participants. However, under dual-task conditions, high and low spans were equivalently vulnerable to PI. Dividing attention, whether at encoding or retrieval, increased high spans' PI effects, but it had no measurable effect on low spans' PI effects. High spans thus appeared to reduce the effect of PI with controlled attention at both encoding and retrieval, whereas low spans did not seem to increase their use of controlled processing under interference conditions compared with noninterference conditions. Moreover, only low spans showed divided attention costs on List 1 encoding and retrieval, suggesting that whatever control capabilities low spans had were being exhausted on List 1, before the potential for PI was high. Low spans may have required controlled processing on List 1 because of extralist interference from high-dominance category exemplars.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, M. C., & Bjork, R. A. (1994). Mechanisms of inhibition in long-term memory: A new taxonomy. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 265–325). San Diego, CA: Academic Press.

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.

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

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs, 60* (Whole No. 277).

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

Baddeley, A. D. (1993). Working memory or working attention? In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 152–170). Oxford, England: Clarendon Press.

Baddeley, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5–28.

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

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

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80,* 1–46.

Bjork, E. L., Bjork, R. A., & Glenberg, A. (1973, November). *Reinstatement of interference owing to to-be-forgotten items.* Paper presented at the annual meeting of the Psychonomic Society, St. Louis, MO.

Bjork, E. L., Bjork, R. A., & White, S. A. (1984, November). *On the induced recovery of proactive interference.* Paper presented at the annual meeting of the Psychonomic Society, San Antonio, TX.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10,* 12–21.

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

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

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

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

Dagenbach, D., & Carr, T. H. (1994). Inhibitory processes in perceptual recognition: Evidence for a center-surround attentional mechanism. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 327–357). New York: Academic Press.

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

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89,* 63–100.

Dempster, F. N. (1991). Inhibitory processes: A neglected dimension in intelligence. *Intelligence, 15,* 157–173.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12,* 45–75.

Dempster, F. N., & Cooney, J. B. (1982). Individual differences in digit span, susceptibility to proactive interference, and aptitude/achievement test scores. *Intelligence, 6,* 399–416.

Dillon, R. F. (1973). Locus of proactive interference effects in short-term memory. *Journal of Experimental Psychology, 99,* 75–81.

Dillon, R. F., & Thomas, H. (1975). The role of response confusion in proactive interference. *Journal of Verbal Learning and Verbal Behavior, 14,* 603–615.

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

Engle, R. W. (1996). Working memory and retrieval: An inhibition-resource approach. 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* (pp. 89–119). New York: Oxford University Press.

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

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

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

Foucalt, M. (1928). Les inhibitions internes de fixation [Internal inhibitions of fixation]. *L'Annee Psychologie, 29,* 92–113.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. J. (1991). Willed action and the prefrontal cortex in man: A study with PET. *Proceedings of the Royal Society of London, 244B,* 241–246.

Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). New York: Raven Press.

Geiselman, R. E., Bjork, R. A., & Fishman, D. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112,* 58–72.

Gibson, J. J., & Raffel, G. (1936). A technique for investigating retroactive and other inhibitory effects in immediate memory. *Journal of General Psychology, 15,* 107–116.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

Golding, J. M., & MacLeod, C. M. (1998). *Intentional forgetting: Interdisciplinary approaches.* Mahwah, NJ: Erlbaum.

Goldman-Rakic, P. S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology—The nervous system* (Vol. 5, pp. 373–417). Bethesda, MD: American Physiological Society.

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

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Hawkins, H. L., Pardo, V. J., & Cox, R. D. (1972). Proactive interference in short-term recognition: Trace interaction or competition? *Journal of Experimental Psychology, 92,* 43–48.

Houghton, G., & Tipper, S. P. (1994). A model of inhibitory mechanisms in selective attention. In D. Dagenbach & T. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 53–112). San Diego, CA: Academic Press.

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience, 14,* 3775–3790.

Kapur, S., Craik, F. I. M., Tulving, E., Wilson, A. A., Houle, S., & Brown, G. M. (1994). Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proceedings of the National Academy of Science, 91,* 2008–2011.

Kapur, S., Tulving, E., Cabeza, R., McIntosh, A. R., Houle, S., & Craik, F. I. M. (1996). The neural correlates of intentional learning of verbal materials: A PET study in humans. *Cognitive Brain Research, 4,* 243–249.

Keppel, G. (1968). Retroactive and proactive inhibition. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 172–213). Englewood Cliffs, NJ: Prentice-Hall.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1118–1133.

Loess, H., & Waugh, N. C. (1967). Short-term memory and

inter-trial interval. *Journal of Verbal Learning and Verbal Behavior, 6,* 455–460.

Loftus, G. R., & Patterson, K. K. (1975). Components of short-term proactive interference. *Journal of Verbal Learning and Verbal Behavior, 14,* 105–121.

Martin, A., Wiggs, C. L., Lalonde, F., & Mack, C. (1994). Word retrieval to letter and semantic cues: A double dissociation in normal subjects using interference tasks. *Neuropsychologia, 32,* 1487–1494.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53,* 173–203.

Mensink, G. J., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434–455.

Moscovitch, M. (1992a). A neuropsychological model of memory and consciousness. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 5–22). New York: Guilford Press.

Moscovitch, M. (1992b). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience, 4,* 257–267.

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

Moscovitch, M., Kapur, S., Kohler, S., & Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity: A positron emission tomography (PET) study in humans. *Proceedings of the National Academy of Science, 92,* 3721–3725.

Murdock, B. B., Jr. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology, 56,* 413–419.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1091–1104.

Neill, W. T. (1977). Inhibition and facilitation processes in selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 444–450.

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review, 99,* 322–348.

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

Nyberg, L., McIntosh, A. R., Cabeza, R., Nilsson, L. G., Houle, S., Habib, R., & Tulving, E. (1996). Network analysis of positron emission tomography regional cerebral blood flow data: Ensemble inhibition during episodic memory retrieval. *Journal of Neuroscience, 16,* 3753–3759.

Park, D. C., Smith, A. D., Dudley, W. N., & Lafronza, V. N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1185–1191.

Partington, J. E., & Leiter, R. G. (1949). *Partington Pathways Test.* Washington, DC: Psychological Services Center.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

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

Postman, L., & Hasher, L. (1972). Conditions of proactive inhibition in free recall. *Journal of Experimental Psychology, 92,* 276–284.

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

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

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology, 5,* 393–394.

Ricker, J. H., & Axelrod, B. N. (1994). Analysis of an oral paradigm for the Trail Making Test. *Assessment, 1,* 47–51.

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

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

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

Shallice, T., & Burgess, P. W. (1993). Supervisory control of action and thought selection. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 171–187). Oxford, England: Clarendon Press.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994, April). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature, 368,* 633–635.

Shimamura, A. P., Jurica, P. J., Mangels, J. A., Gershberg, F. B., & Knight, R. T. (1995). Susceptibility to memory interference effects following frontal lobe damage: Findings from tests of paired-associate learning. *Journal of Cognitive Neuroscience, 7,* 144–152.

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

Towell, N., Burton, A., & Burton, E. (1994). The effects of two matched memory tasks on concurrent tapping. *Neuropsychologia, 32,* 125–129.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tulving, E., Kapur, S., Craik, F. I. M., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Science, 91,* 2016–2020.

Uhl, F., Podreka, I., & Deecke, L. (1994). Anterior frontal cortex and the effect of proactive interference in word pair learning—Results of Brain-SPECT. *Neuropsychologia, 32,* 241–247.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49–60.

Underwood, B. J., Boruch, R. F., & Malmi, R. A. (1978). Composition of episodic memory. *Journal of Experimental Psychology: General, 107,* 393–419.

Underwood, B. J., & Ekstrand, B. R. (1966). An analysis of some shortcomings in the interference theory of forgetting. *Journal of Experimental Psychology, 73,* 540–549.

Underwood, B. J., & Ekstrand, B. R. (1967). Studies of distributed

practice XXIV: Differentiation and proactive inhibition. *Journal of Experimental Psychology, 74*, 574–580.

Watkins, O. C., & Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 1*, 442–452.

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

Wickens, D. D., & Gittis, M. M. (1974). The temporal course of recovery from interference and degree of learning in the Brown-Peterson paradigm. *Journal of Experimental Psychology, 102*, 1021–1026.

Winograd, E., & Smith, W. S. (1966). List differentiation with varied trials on both lists. *Science, 152*, 1101–1102.

Yntema, D. B., & Trask, F. P. (1963). Recall as a search process. *Journal of Verbal Learning and Verbal Behavior, 2*, 65–74.

Young, C. W., & Supa, M. (1941). Mnemic inhibition as a factor in the limitation of the memory span. *American Journal of Psychology, 54*, 546–552.

Zacks, R. T., & Hasher, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 241–264). New York: Academic Press.

Appendix A

Four Sample Stimulus Sets (the "Animal" PI Buildup Sets) Used in Experiment 1

Set	List 1	List 2	List 3	List 4
1	SHEEP	GIRAFFE	GOAT	SCIENTIST
	SQUIRREL	WOLF	DONKEY	POLICEMAN
	LEOPARD	MULE	FOX	SECRETARY
	BUFFALO	SKUNK	RHINOCEROS	ARTIST
	ANTELOPE	MOOSE	RACCOON	SAILOR
	LLAMA	GAZELLE	CHEETAH	FIREMAN
	BEAVER	COYOTE	HAMSTER	PRESIDENT
	CHIPMUNK	PUMA	TURTLE	BARBER
	LYNX	MONKEY	GROUNDHOG	DRIVER
	OPOSSUM	ALLIGATOR	HYENA	MUSICIAN
2	ZEBRA	SHEEP	GIRAFFE	FARMER
	RABBIT	SQUIRREL	WOLF	BANKER
	BULL	LEOPARD	MULE	JUDGE
	CAMEL	BUFFALO	SKUNK	ARCHITECT
	PANTHER	ANTELOPE	MOOSE	WRITER
	ELK	LLAMA	GAZELLE	SENATOR
	ANTEATER	BEAVER	COYOTE	MACHINIST
	JAGUAR	CHIPMUNK	PUMA	PILOT
	OX	LYNX	MONKEY	STUDENT
	LIZARD	OPOSSUM	ALLIGATOR	ASTRONAUT
3	GOAT	ZEBRA	SHEEP	BRAZIL
	DONKEY	RABBIT	SQUIRREL	INDIA
	FOX	BULL	LEOPARD	GREENLAND
	RHINOCEROS	CAMEL	BUFFALO	GREECE
	RACCOON	PANTHER	ANTELOPE	IRAQ
	CHEETAH	ELK	LLAMA	EGYPT
	HAMSTER	ANTEATER	BEAVER	SCOTLAND
	TURTLE	JAGUAR	CHIPMUNK	PANAMA
	GROUNDHOG	OX	LYNX	BULGARIA
	HYENA	LIZARD	OPOSSUM	ICELAND
4	GIRAFFE	GOAT	ZEBRA	AUSTRALIA
	WOLF	DONKEY	RABBIT	VIETNAM
	MULE	FOX	BULL	BOLIVIA
	SKUNK	RHINOCEROS	CAMEL	ISRAEL
	MOOSE	RACCOON	PANTHER	VENEZUELA
	GAZELLE	CHEETAH	ELK	ALBANIA
	COYOTE	HAMSTER	ANTEATER	CUBA
	PUMA	TURTLE	JAGUAR	BOLIVIA
	MONKEY	GROUNDHOG	OX	KOREA
	ALLIGATOR	HYENA	LIZARD	COLOMBIA

Appendix B

Analyses of SAT Scores for Participants in Experiment 1 and Experiment 2

For Experiment 1, university records of Verbal (VSAT) and Math (MSAT) scores were available for 151 participants (76 high spans and 75 low spans). On the VSAT, high spans ($M = 615$, $SD = 88$) scored significantly higher than did low spans ($M = 490$, $SD = 75$), $F(1, 149) = 88.34$, $MSE = 6,719.38$. On the MSAT, high spans ($M = 604$, $SD = 88$) also scored significantly higher than did low spans ($M = 503$, $SD = 82$), $F(1, 149) = 53.01$, $MSE = 7,246.63$. High spans tested under no load had equivalent VSATs and MSATs to those high spans tested under load, for VSAT, $F(1, 74) < 1$; for MSAT, $F(1, 74) = 1.13$, $MSE = 7,719.52$, $p > .25$; low spans tested under load also had equivalent VSATs and MSATs to those low spans tested under load, for VSAT, $F(1, 73) < 1$; for MSAT, $F(1, 73) = 1.06$, $MSE = 6,748.14$, $p > .30$.

For Experiment 2, scores were available for 153 participants (78 high spans and 75 low spans). On the VSAT, high spans ($M = 560$, $SD = 94$) scored significantly higher than did low spans ($M = 486$, $SD = 82$), $F(1, 151) = 26.50$, $MSE = 7,817.57$; on the MSAT, high spans ($M = 550$, $SD = 81$) scored significantly higher than did low spans ($M = 480$, $SD = 75$), $F(1, 151) = 30.51$, $MSE = 6,121.08$. High spans tested under no load scored equivalently to those tested under load at encoding, for VSAT, $F(1, 50) = 1.83$, $MSE = 6,713.69$, $p = .18$; for MSAT, $F(1, 50) < 1$, and under load at retrieval, for both VSAT and MSAT, $F(1, 51) < 1$. Low spans tested under no load had equivalent VSATs to those tested under load at encoding, $F(1, 49) < 1$, but no-load participants had significantly lower MSAT scores than encoding-load participants, $F(1, 49) = 4.66$, $MSE = 5,015.59$. Low spans tested under no load had equivalent VSAT and MSAT scores to low spans tested under load at retrieval, for VSAT, $F(1, 49) < 1$; for MSAT, $F(1, 49) = 1.95$, $MSE = 6,011.44$, $p = .17$.

Appendix C

Analyses of Variance on the Raw Recall Data From Experiment 1

Omnibus Test: 2 (spans) $\times$ 2 (tapping conditions) $\times$ 3 (Lists 1–3)

Span effect	$F(1, 188) = 32.18$	$MSE = 3.21$	$p < .05$
Load effect	$F(1, 188) = 40.58$	$MSE = 3.21$	$p < .05$
Span $\times$ Load	$F(1, 188) = 1.14$	$MSE = 3.21$	$p = .29$
Trial effect	$F(2, 376) = 221.84$	$MSE = 1.14$	$p < .05$
Span $\times$ List	$F(2, 376) = 1.53$	$MSE = 1.14$	$p = .22$
Load $\times$ List	$F(2, 376) < 1$		
Span $\times$ Load $\times$ List	$F(2, 376) = 1.72$	$MSE = 1.14$	$p = .18$

High spans vs. low spans, under cascade tapping (no load): 2 (spans) $\times$ 3 (lists)

Span effect	$F(1, 94) = 23.60$	$MSE = 3.09$	$p < .05$
List effect	$F(2, 188) = 106.54$	$MSE = 1.68$	$p < .05$
Span $\times$ List	**$F(2, 188) = 3.11$**	**$MSE = 1.68$**	**$p < .05$**

High spans vs. low spans, under complex tapping (load): 2 (spans) $\times$ 3 (lists)

Span effect	$F(1, 94) = 10.22$	$MSE = 3.33$	$p < .05$
List effect	$F(2, 188) = 115.74$	$MSE = 1.56$	$p < .05$
Span $\times$ List	**$F(2, 188) < 1$**		

High spans under cascade (no load) vs. complex tapping (load): 2 (tapping) $\times$ 3 (lists)

Load effect	$F(1, 94) = 26.42$	$MSE = 3.37$	$p < .05$
List effect	$F(2, 188) = 89.29$	$MSE = 1.75$	$p < .05$
Load $\times$ List	$F(2, 188) < 1$		

Low spans under cascade (no load) vs. complex tapping (load): 2 (tapping) $\times$ 3 (lists)

Load effect	$F(1, 94) = 14.75$	$MSE = 3.06$	$p < .05$
List effect	$F(2, 188) = 138.04$	$MSE = 1.48$	$p < .05$
Load $\times$ List	**$F(2, 188) = 1.28$**	**$MSE = 1.48$**	**$p = .28$ (favoring *load*)**

Note. PI-relevant findings that are consistent with the proportion-loss analyses reported in the text appear in bold type.

(Appendixes continue)

Appendix D

Analyses of Variance on the Raw Recall Data From Experiment 2

Omnibus Test: 2 (span) $\times$ 3 (load) $\times$ 3 (list)
Span effect	$F(1, 210) = 25.68$	$MSE = 3.04$	$p < .05$
Load effect	$F(2, 210) = 13.82$	$MSE = 3.04$	$p < .05$
Span $\times$ Load	$F(2, 210) < 1$		
List	$F(2, 420) = 244.51$	$MSE = 1.56$	$p < .05$
Span $\times$ List	$F(2, 420) = 1.52$	$MSE = 1.56$	$p < .23$
Load $\times$ List	$F(4, 420) < 1$		
Span $\times$ Load $\times$ List	**$F(4, 420) = 2.78$**	**$MSE = 1.56$**	**$p < .05$**

High spans vs. low spans, under no load: 2 (span) $\times$ 3 (list)
Span effect	$F(1, 70) = 11.27$	$MSE = 2.70$	$p < .05$
List effect	$F(2, 140) = 76.07$	$MSE = 1.66$	$p < .05$
Span $\times$ List	$F(2, 140) = 2.24$	$MSE = 1.66$	$p = .11$

Planned contrast for span difference in PI between
Lists 1 and 2	$F(1, 70) = 2.01$	$MSE = 3.35$	$p = .16$
Lists 1 and 3	**$F(1, 70) = 4.75$**	**$MSE = 2.99$**	**$p < .05$**

High spans vs. low spans, under load at encoding: 2 (span) $\times$ 3 (list)
Span effect	$F(1, 70) = 10.59$	$MSE = 3.54$	$p < .05$
List effect	$F(2, 140) = 85.10$	$MSE = 1.43$	$p < .05$
Span $\times$ List	**$F(2, 140) = 2.11$**	**$MSE = 1.43$**	**$p = .13$ (favoring *low spans*)**

High spans vs. low spans, under load at retrieval: 2 (span) $\times$ 3 (list)
Span effect	$F(1, 70) = 4.67$	$MSE = 2.89$	$p < .05$
List effect	$F(2, 140) = 84.65$	$MSE = 1.59$	$p < .05$
Span $\times$ List	**$F(2, 140) = 2.76$**	**$MSE = 1.59$**	**$p = .07$ (favoring *low* spans)**

High spans under no load vs. load at encoding: 2 (load) $\times$ 3 (list)
Load effect	$F(1, 70) = 8.81$	$MSE = 3.12$	$p < .05$
List effect	$F(2, 140) = 58.31$	$MSE = 1.89$	$p < .05$
Load $\times$ List	$F(2, 140) = 1.69$	$MSE = 1.89$	$p = .19$

Planned contrast for Load difference in PI between
Lists 1 and 2	**$F(1, 70) = 3.54$**	**$MSE = 3.54$**	**$p = .06$**
Lists 1 and 3	$F(1, 70) < 1$		

High spans under no load vs. load at retrieval: 2 (load) $\times$ 3 (list)
Load effect	$F(1, 70) = 15.12$	$MSE = 2.88$	$p < .05$
List effect	$F(2, 140) = 69.21$	$MSE = 1.89$	$p < .05$
Load $\times$ List	**$F(2, 140) = 2.52$**	**$MSE = 1.89$**	**$p = .08$**

Planned contrast for load difference in PI between
Lists 1 and 2	**$F(1, 70) = 4.34$**	**$MSE = 3.48$**	**$p < .05$**
Lists 1 and 3	**$F(1, 70) = 3.73$**	**$MSE = 3.73$**	**$p = .06$**

Low spans under no load vs. load at encoding
Load effect	$F(1, 70) = 10.98$	$MSE = 3.12$	$p < .05$
List effect	$F(2, 140) = 116.24$	$MSE = 1.20$	$p < .05$
Load $\times$ List	**$F(2, 140) = 1.59$**	**$MSE = 1.20$**	**$p = .21$ (favoring *load*)**

Low spans under no load vs. load at retrieval
Load effect	$F(1, 70) = 8.39$	$MSE = 2.70$	$p < .05$
List effect	$F(2, 140) = 95.68$	$MSE = 1.37$	$p < .05$
Load $\times$ List	**$F(2, 140) = 2.21$**	**$MSE = 1.37$**	**$p = .11$ (favoring *load*)**

Note. PI-relevant findings that are consistent with the proportion-loss analyses reported in the text appear in bold type.