

Attention during memory retrieval enhances future remembering

NICOLE M. DUDUKOVIC, SARAH DUBROW, AND ANTHONY D. WAGNER
Stanford University, Stanford, California

Memory retrieval is a powerful learning event that influences whether an experience will be remembered in the future. Although retrieval can succeed in the presence of distraction, dividing attention during retrieval may reduce the power of remembering as an encoding event. In the present experiments, participants studied pictures of objects under full attention and then engaged in item recognition and source memory retrieval under full or divided attention. Two days later, a second recognition and source recollection test assessed the impact of attention during initial retrieval on long-term retention. On this latter test, performance was superior for items that had been tested initially under full versus divided attention. More importantly, even when items were correctly recognized on the first test, divided attention reduced the likelihood of subsequent recognition on the second test. The same held true for source recollection. Additionally, foils presented during the first test were also less likely to be later recognized if they had been encountered initially under divided attention. These findings demonstrate that attentive retrieval is critical for learning through remembering.

Few people would debate the necessity of attention for learning new information. Extensive research has demonstrated that dividing attention during learning results in diminished declarative memory relative to fully dedicating attention during learning (e.g., Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Mulligan, 1998). However, the effects of divided attention during memory retrieval seem to be less severe than they are at encoding, with many studies failing to observe any effects other than reaction time (RT) costs to the secondary, or distractor, task (Baddeley et al., 1984; Craik et al., 1996; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Naveh-Benjamin, Craik, Perretta, & Tonev, 2000). Retrieval costs tend to be maximal when the distractor task and mnemonic task compete for the same representational system (Fernandes & Moscovitch, 2000, 2002, 2003).

The extent to which attention influences memory retrieval may depend on the type of processes involved in the memory decision. Recognition memory decisions can be based either on *recollection*, the retrieval of specific contextual details associated with an item, or on *familiarity*, the sense of having encountered an item without retrieving any specific details (Tulving, 1985; Yonelinas, 2002). Recollection is thought to be a more attention-demanding process, whereas familiarity is considered to be a more automatic process (Jacoby, 1991). Consequently, performance on memory tests that rely on item familiarity may be relatively unaffected by allocating attention to an unrelated distractor task, whereas performance on tests that involve effortful recollection processes may differentially suffer under conditions of divided attention (Lozito

& Mulligan, 2006). Supporting this prediction, recent research suggests that recognition performance suffers under divided attention when study and test conditions promote elaborative encoding and contextual retrieval (Hicks & Marsh, 2000; Lozito & Mulligan, 2006; Troyer, Winocur, Craik, & Moscovitch, 1999).

Even when distraction does not impede successful memory retrieval, it may still produce mnemonic costs. The act of remembering serves to re-encode our experiences, increasing the chances that we will remember them in the future. As such, retrieval tests are particularly effective study events for ensuring subsequent remembering (Carrier & Pashler, 1992; Gates, 1917; Hogan & Kintsch, 1971; Landauer & Bjork, 1978; Roediger & Karpicke, 2006b; Wheeler, Ewers, & Buonanno, 2003). However, it is unknown whether these beneficial testing effects are attention dependent. Because divided attention during encoding diminishes recognition memory performance, we predicted that dividing attention during retrieval would similarly reduce the encoding power of a retrieval event, decreasing the likelihood of subsequent remembering. Distraction during memory retrieval could disrupt memory processes that would give rise to later recollection, familiarity, or both. Recent research has demonstrated that initial recall tests increase subsequent recollection without affecting familiarity processes (Chan & McDermott, 2007), suggesting that divided attention during initial retrieval may be detrimental for future recollection-based memories but not for familiarity-based memories.

In two experiments, we examined whether items that were accurately retrieved on an initial memory test would

N. M. Dudukovic, nicole.dudukovic@trincoll.edu

be less likely to be remembered on a second test if they had been initially retrieved under divided versus full attention. These experiments provide novel insight as to whether the power of retrieval as an encoding event is attention dependent and, if so, whether the encoding consequences of attention at retrieval differ for subsequent source memory, which may be more likely to require recollection than would item recognition. The results have important implications for everyday learning through retrieval practice, which often occurs amid distractions. These distractions may reduce the likelihood that a successful retrieval attempt in the present promotes successful remembering in the future.

EXPERIMENT 1

Method

Participants. Twenty-two Stanford University students (ages 18–32 years, mean = 20.0 years; $SD = 3.0$ years) participated in two experimental sessions in exchange for \$20, in accord with procedures approved by the Stanford University Institutional Review Board.

Materials. The stimuli were 226 color pictures of nameable things (113 living, 113 nonliving) from a previously described set (Simons, Koutstaal, Prince, Wagner, & Schacter, 2003). For each participant, 124 pictures appeared as study items. The rest served as novel items: 2 during practice, 40 during the first test, and the remaining 60 during the second test. Pictures were counterbalanced across participants; each picture appeared as both a studied and novel item and was tested under full and divided attention.

Procedure. The experiment consisted of three phases: study, Test 1, and Test 2 (see Figure 1). Each study trial began with a task cue indicating whether to make a living/nonliving or like/dislike judgment. After 500 msec, a picture appeared beneath the cue for 3,000 msec, during which time participants made their decision using one of two keys. Living/nonliving judgments (62 pictures) and like/dislike judgments (62 pictures) were intermixed randomly. The first 2 and the last 2 pictures were filler items, controlling for primacy and recency effects.

Following the study phase, participants practiced the auditory discrimination task and the memory task. The auditory discrimination task consisted of two 1,500-msec auditory patterns with distinct rhythms, presented continuously (Kensinger, Clarke, & Corkin, 2003). Participants were instructed to detect the patterns and respond with their left index finger whenever one pattern switched to the other. The memory practice task consisted of six pictures: the four filler pictures from the study phase and two novel pictures. Each picture appeared for 3,500 msec, during which time participants made a keypress with their right hand to indicate one of four item recognition and source memory judgments (old–living/nonliving task, old–like/dislike task, old–unsure of task, or new). Participants practiced the auditory task twice, the memory task once, and then the auditory and memory tasks simultaneously. Participants were instructed that both tasks were equally important and that they should respond as accurately as possible on both tasks.

During Test 1, participants were presented with 80 studied pictures and 40 novel pictures. Half were presented while the participants simultaneously performed the auditory task; half were presented under full attention. The divided attention and full attention conditions (DA and FA, respectively) alternated in blocks of 10 pictures, for a total of 12 blocks. At the beginning of each block, an instruction screen appeared for 2,000 msec informing participants of the type of block (FA or DA) to follow.

Two days later, participants returned to complete Test 2, during which they again made recognition and source memory decisions, all under full attention. Participants were presented with the 80 studied items that appeared in the first memory test (40 DA, 40 FA),

the 40 studied items that were never tested, and 60 new items. Participants made a two-step memory decision, first indicating whether each picture was old or new. For pictures judged old, participants indicated which task they had performed at study by responding “living/nonliving task,” “like/dislike task,” or “unsure of task.” All items were intermixed randomly, and all memory decisions on Test 2 were self-paced.

Results

Test 1. The proportion of studied pictures correctly recognized as old on Test 1 was higher for items tested under FA versus DA [$F(1,21) = 9.90, p < .001$] (see Table 1). When the hit rate was corrected for false alarms (hits – false alarms), corrected recognition still tended to be higher for FA versus DA [$F(1,21) = 3.74, p = .07$].

Source memory accuracy was calculated as the proportion of hits that were accompanied by accurate source decisions [i.e., $P(\text{source correct} | \text{hit})$], referred to as the *identification-of-origin score* (Johnson, Hashtroudi, & Lindsay, 1993). Source memory was significantly better for pictures tested under FA versus DA [$F(1,21) = 15.52, p = .001$] (see Table 1). Furthermore, the proportion of incorrect source responses was higher under DA [$F(1,21) = 13.55, p = .001$], whereas the proportion of “unsure” source responses was equivalent ($p > .10$).

The identification-of-origin measure of source memory may still be biased by recognition memory levels, particularly when recognition levels differ between groups (Batchelder & Riefer, 1990). For example, the identification-of-origin measure assumes that a “new” response is made whenever item-detection processes fail even when source discrimination processes succeed, an assumption that has been shown to be violated in some circumstances (Murnane & Bayen, 1996). We conducted an alternative assessment of source memory based on Bayen, Murnane, and Erdfelder’s (1996) multinomial model by using the HMMTree computer program (Stahl & Klauer, 2007) to perform an analysis similar to that described in Mulligan (2004, Appendix). The multinomial model represents purported processes underlying source memory decisions and can be used to test whether source memory is equivalent across conditions by equating the source discrimination parameters and testing whether this reduces the fit of the model to the data. For Test 1, this assessment produced results consistent with the identification-of-origin scores. Equating the DA and FA source discrimination parameters in the model significantly reduced the fit of the model relative to the unconstrained model ($p < .05$), suggesting that DA and FA source memory were not equivalent.

In terms of response speed, the average median RT for hits was significantly faster under DA (1,578 msec) than under FA (1,673 msec) [$F(1,21) = 7.04, p < .05$]. The average median RT for correct source responses was also significantly faster for items tested under DA (1,515 msec) versus FA (1,632 msec) [$F(1,21) = 8.83, p < .01$].

Because memory decisions were faster under DA versus FA, we looked for evidence of a speed–accuracy trade-off. First, we computed correlations between RTs and accuracy for both item and source decisions under DA. None of these correlations approached significance ($ps > .10$).

Study



Test 1

- 1 = Old–Living/Nonliving Task
- 2 = Old–Like/Dislike Task
- 3 = Old–Unsure of Task
- 4 = New

Full Attention (FA)



Studied

Novel

Divided Attention (DA)

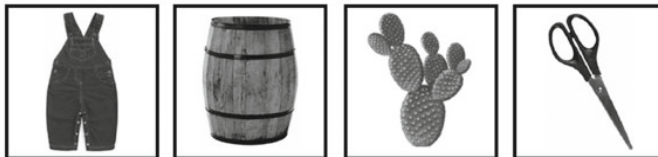


Studied

Novel

Test 2 (2 days later)

- 1 = Old–Living/Nonliving Task
- 2 = Old–Like/Dislike Task
- 3 = Old–Unsure of Task
- 4 = New



Studied
Untested

Studied
FA Tested

Studied
DA Tested

Novel

Exp 2 only

- 1 = Study Only
- 2 = Study & Test
- 3 = Test Only
- 4 = Unsure When



Test 1
FA Novel

Test 1
DA Novel

Figure 1. The experimental design for the three phases: study, Test 1, and Test 2.

Second, we performed a median split and repeated all of our analyses using the fastest half of the FA trials and the slowest half of the DA trials. Using this sample, memory decisions were significantly faster under FA than under DA ($ps < .0005$), but the pattern of data for Test 1 (and Test 2) remained the same. Thus, we found no evidence of a speed–accuracy trade-off.

Secondary task. Performance on the auditory discrimination task in the DA condition was impaired relative to performance during practice, which can be considered as a pretest baseline. It is likely that auditory task performance improved over the course of the experiment due

to practice effects, which suggests that comparing DA performance with the pretest baseline is a conservative estimate of secondary task costs. Under DA, participants were less likely to detect auditory pattern changes [DA mean = 78%; pretest mean = 90%; $F(1,21) = 13.15, p < .005$] and were slower at responding to these changes [DA average median RT = 688 msec; pretest average median RT = 423 msec; $F(1,21) = 121.61, p < .0005$]. False alarm rates to nonchanges did not differ between the DA condition and the pretest baseline (DA false alarms = 17%; pretest false alarms = 15%; $F < 1$). The secondary performance decrements in the DA condition provide

Table 1
Experiment 1: Recognition and Source Memory
Performance on Test 1 and Test 2

Test	Attention Condition	Hits	Hits – False Alarms	Correct Source	Incorrect Source	Unsure of Source
1	Full	.89	.86	.76	.10	.14
	Divided	.83	.83	.67	.18	.15
2	Full	.89	.84	.64	.16	.20
	Divided	.80	.75	.58	.19	.23
	Untested	.55	.50	.57	.14	.29

Note—Scores are mean proportions. Source memory scores are proportions of hits.

evidence that participants were dual-tasking; however, the practice task was identical across participants, leaving open the possibility that the practice task was on average easier than were the auditory discrimination tasks performed in the DA condition.

Test 2. The Test 2 hit rate was significantly different for pictures initially tested under FA, initially tested under DA, and initially untested [$F(2,42) = 113.93, p < .0005$] (see Table 1). Hit rates were higher for pictures previously tested compared with those previously untested ($F_s > 102.21, p_s < .0005$). Importantly, the hit rate was higher for pictures previously tested under FA versus DA [$F(1,21) = 29.43, p < .0005$].

We measured Test 2 source memory accuracy by calculating the proportion of hits for each condition that were also accompanied by accurate source decisions (i.e., identification-of-origin scores). Source memory was impacted significantly by whether and how a picture was tested previously [$F(2,42) = 4.01, p < .05$] (see Table 1). Source memory was superior for pictures previously tested under FA compared with pictures previously tested under DA and pictures previously untested ($F_s > 4.63, p_s < .05$). Strikingly, the proportion of correct source memory decisions for pictures initially tested under DA did not differ from that for previously untested pictures ($F < 1$). However, when source failure occurred, it was more likely to result in an “unsure” response for previously untested pictures compared with an incorrect response for pictures previously tested under DA ($F_s > 5.97, p_s < .05$).

An alternative multinomial model assessment of source memory confirmed that source memory discrimination levels differed depending on whether and how a picture was previously tested (i.e., equating the source discrimination parameters for DA, FA, and previously untested items significantly reduced the fit relative to the unconstrained model; $p < .05$). Equating DA and FA source discrimination parameters did reduce the fit of the unconstrained model, although this reduction was not significant ($p > .10$). Equating the source memory parameters for pictures initially tested under DA and previously untested pictures produced a marginally significant reduction in the fit of the model ($p < .10$), suggesting that source discrimination may potentially differ between these two groups of items.

Mnemonic fate of Test 1 memories. Finally, and critically, we examined whether the mnemonic fate of items that were correctly retrieved on Test 1 differed depending

on whether attention was divided during this initial act of remembering (see Figure 2). Pictures that were correctly recognized on Test 1 were more likely to be correctly recognized on Test 2 if they had been initially retrieved under FA (.91) versus DA (.86) [$F(1,21) = 17.91, p < .0005$]. Similarly, pictures whose source was recollected correctly on Test 1 were more likely to be accompanied by accurate source memory on Test 2 if they had been initially recollected under FA (.69) versus DA (.60) [$F(1,21) = 9.68, p = .005$].

Discussion

In contrast to previous research demonstrating minimal effects of divided attention on memory retrieval (Baddeley et al., 1984; Craik et al., 1996; Naveh-Benjamin et al., 1998; Naveh-Benjamin et al., 2000), both recognition and source memory performance were affected negatively by attention being divided at retrieval. However, when recognition hit rates were corrected for false alarm rates, the difference for recognition memory was only marginally significant. The nature of our encoding procedures, which involved relatively deep semantic tasks, may account for any decrements in recognition memory under conditions of limited attention (Hicks & Marsh, 2000). Further, our results confirm the hypothesis that performance on tests that often involve effortful recollection processes, such as source memory decisions, suffer under conditions of divided attention (Lozito & Mulligan, 2006).

More interestingly, our results indicate that dividing attention during memory retrieval is detrimental for future remembering. Although pictures that were initially tested under divided attention were more likely to be recognized than pictures that had not been initially tested, they were less likely to be recognized than pictures that had been initially tested under full attention. Further, the proportion of recognized pictures that were accompanied by accurate source memory was the same for pictures previously tested under divided attention and for pictures previously untested, although there is evidence that the underlying source discrimination processes may differ depending on whether a picture has been previously tested. Recognized pictures were more likely to be attributed to incorrect sources when they were previously retrieved under divided attention versus not retrieved at all. Participants were more hesitant about making source attributions for previously untested pictures, making more “unsure” responses for those pictures.

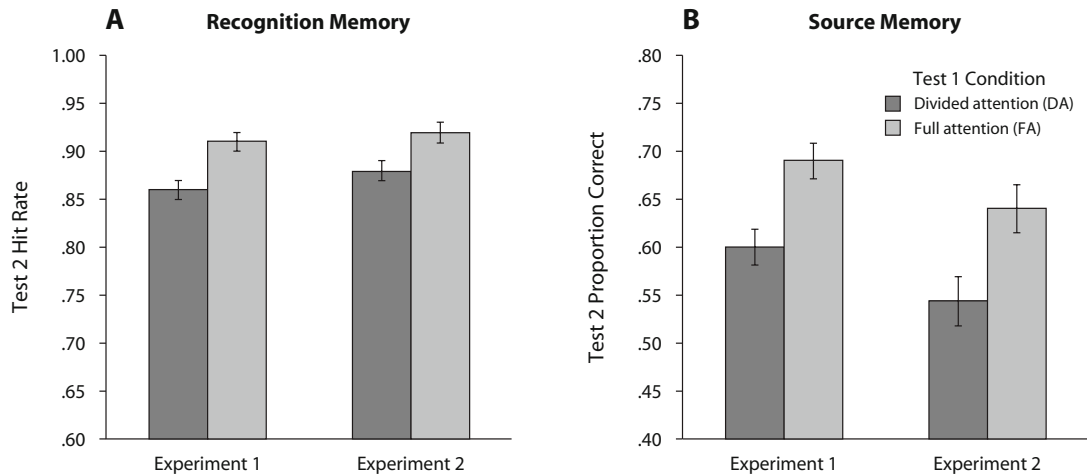


Figure 2. The mnemonic fate of Test 1 recognition and source memory hits in Experiments 1 and 2. (A) Test 2 recognition hit rates for pictures recognized on Test 1. (B) Test 2 source memory performance for pictures whose source was correctly retrieved on Test 1.

Most importantly, the results of Experiment 1 support our hypothesis that the power of retrieval as an encoding event is attention dependent. Pictures that were accurately retrieved on the first test were less likely to be accurately retrieved on the second test when attention was divided on the first test. This held true for source memory retrieval as well as for recognition, suggesting that divided attention during retrieval may be detrimental for future recollection-based memories.

EXPERIMENT 2

In Experiment 1, we assessed the mnemonic fate of studied pictures, but we did not examine the fate of pictures that appeared as foils on the first memory test. Given the wealth of research demonstrating the negative effects of divided attention on memory encoding (e.g., Baddeley et al., 1984; Craik et al., 1996; Mulligan, 1998; Naveh-Benjamin et al., 1998; Naveh-Benjamin et al., 2000), it is possible that these foils are more likely to be encoded when attention is fully dedicated during memory retrieval. Recent research suggests that recognition memory for foils may be superior under conditions in which initial retrieval is constrained by a focus on deep (as opposed to more shallow) source information (Jacoby, Shimizu, Daniels, & Rhodes, 2005). Older adults are less likely to constrain retrieval in a way that benefits later foil memory (Jacoby, Shimizu, Velanova, & Rhodes, 2005), suggesting that the process of recapitulating deep study processes during retrieval may be attention demanding. Thus, we predicted that foil memory would be superior when attention is fully engaged during initial retrieval attempts.

However, foils that appear during a recognition memory test are often later falsely remembered as being correct (Butler, Marsh, Goode, & Roediger, 2006; Roediger & Marsh, 2005), suggesting that when foils are encoded

during memory tests, relevant source information may not always be encoded as well. Therefore, people may later falsely attribute the source of a novel item that appeared on an initial test as having been presented during study, and they may be more likely to do so if the item was initially presented under divided attention. In Experiment 2, we included foils from the first memory test on the second memory test to determine whether subjects would remember encountering these pictures and the circumstances under which they encountered them.

Method

Participants. Twenty Stanford University students and affiliates (ages 18–23 years; mean = 20.5 years; $SD = 1.2$ years) participated in two experimental sessions in exchange for \$20. Informed consent was obtained in accordance with the Stanford University Institutional Review Board.

Materials. The stimuli were the same 226 color pictures of nameable objects used in Experiment 1 along with 20 additional pictures drawn from the same set and used as additional novel items during Test 2.

Procedure. The procedures for the study phase, practice tasks, and Test 1 were identical to those used in Experiment 1. Two days after the first experimental session, participants returned to the laboratory and completed Test 2. They were presented with the 80 studied items from Test 1, the 40 studied items that were never tested, the 40 novel items from Test 1 (20 DA, 20 FA), and 80 new items. Participants were told that some of the pictures had appeared only during study, that some had appeared during study and Test 1, that some had appeared only during Test 1, and that some were new. Participants made a three-step memory decision, first indicating whether each picture was old or new. Classifying an item as *old* signified that it had been seen at some point during the first experimental session. For pictures classified as *old*, participants indicated when the picture had been seen (study only, study and test, test only, or unsure when). Finally, for pictures classified as *viewed during study*, participants indicated which task they had performed (living/nonliving task, like/dislike task, or unsure of task). As in Experiment 1, items were randomly intermixed, and memory decisions on Test 2 were self-paced.

Results

Test 1. In contrast to Experiment 1, both the hit rate and corrected recognition were similar on Test 1 for pictures tested under FA versus DA ($ps > .10$). These results are summarized in Table 2. As in Experiment 1, source memory was significantly better for pictures tested under FA versus DA [$F(1,19) = 16.70, p = .001$]. The proportion of incorrect source responses was higher under DA [$F(1,19) = 5.59, p < .05$], whereas the proportion of “unsure” source responses did not differ ($p > .10$). The results of an alternative multinomial model assessment of source memory were consistent with the finding that Test 1 source memory discrimination differed for pictures tested under FA versus DA ($p < .05$).

The average median RT for hits was significantly faster under DA (1,619 msec) compared with FA (1,789 msec) [$F(1,19) = 9.57, p < .01$]. Similarly, the average median RT for correct source responses was significantly faster for items tested under DA (1,569 msec) versus FA (1,715 msec) [$F(1,19) = 7.01, p < .05$].

Because memory decisions were again faster under DA versus FA, we tested for a speed–accuracy trade-off. None of the correlations between RTs and item and source accuracy under DA approached significance ($ps > .10$). When we performed a median split and repeated all of our analyses using the fastest half of the FA trials and the slowest half of the DA trials, memory decisions were significantly faster under FA than under DA ($p < .0005$), but the pattern of data for Test 1 (and Test 2) remained the same. The one exception was that the hit rate for the slowest DA trials was higher than that for the fastest FA trials at Test 1; however, this effect reversed at Test 2 [$F(1,19) = 8.71, p < .01$]. Thus, once again, we found no evidence of a speed–accuracy trade-off.

Secondary task. As in Experiment 1, performance on the auditory discrimination task in the DA condition was impaired relative to performance during practice, the pretest baseline. When performing the auditory task and the memory task simultaneously, participants were less likely to detect auditory pattern changes [DA mean = 77%; pretest mean = 88%; $F(1,19) = 12.31, p < .005$] and were slower at responding to these changes [DA average median RT = 721 msec; pretest average median RT = 418 msec; $F(1,19) = 169.68, p < .0005$]. False alarm rates to non-

changes also differed between the DA condition and the pretest baseline [DA false alarms = 19%; pretest false alarms = 9%; $F(1,19) = 15.48, p = .001$]. As in Experiment 1, the secondary performance decrements in the DA condition provide evidence that participants were engaging in both the memory task and the auditory task.

Test 2. As in Experiment 1, the Test 2 hit rate was significantly different for pictures initially tested under FA, initially tested under DA, and initially untested [$F(2,38) = 109.00, p < .0005$]. These results are summarized in Table 2. Hit rates were higher for pictures previously tested than for those previously untested ($F_s > 96.88, ps < .0005$). Again, the hit rate was higher for pictures previously tested under FA versus DA [$F(1,19) = 14.55, p = .001$]. Additionally, pictures that had appeared as novel foils in Test 1 were better recognized when they had appeared under FA versus DA [$F(1,19) = 18.32, p < .0005$].

We assessed list discrimination success (i.e., correct decisions about phases in which pictures were previously seen) by calculating the proportion of hits that were accompanied by accurate list discrimination decisions. List discrimination was significantly affected by whether and how a studied picture was previously tested [$F(2,38) = 24.36, p < .0005$] (see Table 2). List discrimination for pictures previously tested under FA was superior to that for pictures previously tested under DA and for pictures previously untested ($F_s > 25.40, ps < .0005$). Likewise, list discrimination for pictures previously tested under DA was superior to that for pictures previously untested [$F(1,19) = 8.20, p = .01$]. For pictures that appeared as novel foils during Test 1, list discrimination, or correct attribution to *test only*, did not differ for foils that appeared under DA versus FA ($F < 1$).

An additional multinomial model assessment of list discrimination, or context memory, supported our findings that list discrimination significantly differed between pictures previously tested under DA versus FA ($p < .05$) and did not differ between pictures that appeared as novel foils under DA versus FA ($p > .10$).

Mnemonic fate of Test 1 memories. We again examined whether the mnemonic fate of items that were correctly retrieved on Test 1 differed depending on whether attention was divided during this initial remembering (see Figure 2). Pictures that were correctly recognized on Test 1 were more likely to be correctly recognized on Test 2 if they had been initially retrieved under FA (.92) versus DA (.88) [$F(1,19) = 5.46, p < .05$]. Pictures whose source was correctly recollected on Test 1 were also more likely to be accompanied by accurate source memory on Test 2 if they had been initially recollected under FA (.64) versus DA (.54) [$F(1,19) = 6.95, p < .05$].

Discussion

In Experiment 2, recognition memory performance on Test 1 was equivalent, regardless of whether attention was simultaneously allocated to the auditory task, in accordance with previous research demonstrating minimal effects of divided attention on recognition memory tasks (Baddeley et al., 1984; Craik et al., 1996; Naveh-

Table 2
Experiment 2: Recognition and Source Memory
Performance on Test 1 and Test 2

Test	Attention Condition	Hits	Hits – False Alarms	Correct List	Correct Source
1	Full	.84	.79	–	.74
	Divided	.81	.76	–	.64
2	Full	.88	.82	.46	.62
	Divided	.82	.76	.32	.61
	Untested	.58	.52	.20	.51
	Full, foils	.57	.51	.40	–
	Divided, foils	.39	.33	.38	–

Note—Scores are mean proportions. List discrimination scores are proportions of hits. Source memory scores are proportions of hits correctly recognized as appearing during the study phase.

Benjamin et al., 1998; Naveh-Benjamin et al., 2000). Source memory, on the other hand, was significantly less accurate when retrieval took place under DA. This result replicates our finding in Experiment 1 and supports the hypothesis that memory tests that involve effortful recollection processes, such as source memory decisions, are attention dependent (Lozito & Mulligan, 2006).

Dividing attention during memory retrieval once again reduced the power of retrieval as an encoding event. As in Experiment 1, pictures that were initially tested under divided attention were less likely to be later recognized than were pictures that were initially tested under full attention, even if they were correctly recognized on the first test. Similarly, accurate source memories from the first test were more likely to be maintained when full attention was allocated to retrieval during the first test.

List discrimination performance was also more accurate for pictures tested under full attention. In other words, participants were better able to remember encountering items during *both* study and Test 1 when attention was undivided at Test 1, suggesting that limiting attention during Test 1 exposure impaired memory for the study episode, the Test 1 episode, or both. It is unclear whether participants were making their list discrimination decisions on the basis of memory for prior encounters with the stimuli or on the basis of memory for their previous responses to stimuli. For example, a participant could correctly respond that an item was previously encountered at both “study and test” because s/he remembers encountering the item on two separate occasions, or alternatively because s/he remembers responding “living” to the “living/nonliving” task at study and responding “old–living/nonliving” at Test 1. This latter possibility, which entails memory for past responses or previous remembrances, is referred to as *output monitoring* (e.g., Gardiner & Klee, 1976; Marsh, Hicks, Hancock, & Munsayac, 2002). Recent evidence suggests that output monitoring suffers when attentional resources are limited during initial retrieval (Sugimori & Kusumi, 2009). Our results support these findings by demonstrating that when a picture is retrieved under divided attention, the retrieval event itself is less likely to be subsequently remembered, as measured on a list discrimination test. However, participants were equally likely to respond “test only” as “study only” to pictures that were presented at both study and Test 1 under divided attention, suggesting that memory for the initial study episode may have also suffered when attentional resources were limited during testing.

In Experiment 2, we also assessed the mnemonic fate of pictures that appeared as foils on the first memory test. These foils were more likely to be remembered subsequently if they were encountered under full versus divided attention during memory retrieval, demonstrating that the negative effects of divided attention on memory encoding can also occur in the context of retrieval. Contrary to our predictions, list discrimination (i.e., memory for when foils were encountered) was no better for foils presented under full versus divided attention. This null effect may be the product of weak memory traces for all foils, evidenced by the low recognition rates and a large proportion of “un-

sure” list discrimination responses (.44 for DA foils; .49 for FA foils).

GENERAL DISCUSSION

Memory retrieval often occurs amid distractions. For example, while writing to an acquaintance, you might be reminded that you need to call a colleague. Are you ultimately less likely to remember to make that phone call because you were distracted by other tasks? The results of the present experiments suggest that the answer is yes. Attention is not only necessary during memory encoding (e.g., Baddeley et al., 1984; Craik et al., 1996; Mulligan, 1998), it is also important during memory retrieval to ensure future remembering.

In the present experiments, dividing attention during memory retrieval impaired source memory while impacting item recognition to a lesser extent, supporting the hypothesis that performance on tests that involve effortful contextual retrieval processes suffers under conditions of divided attention (Lozito & Mulligan, 2006). Furthermore, dividing attention during retrieval increased incorrect source memory responses, yet left the proportion of unsure responses the same, providing further evidence that distraction during retrieval may increase false recollection (Knott & Dewhurst, 2007; Skinner & Fernandes, 2008).

Foils presented on an initial memory test were more likely to be recognized subsequently if they had been encountered under FA versus DA, demonstrating that the negative effects of divided attention during memory encoding can also occur in the context of retrieval. Foils may have been less likely to be encoded under DA, because there were limited encoding resources available. Alternatively, the retrieval processes that were engaged when foils were encountered under DA may have been less likely to be constrained to the recapitulation of study processes, resulting in shallower processing of these foils (Jacoby, Shimizu, Daniels, & Rhodes, 2005; Jacoby, Shimizu, Velanova, & Rhodes, 2005). However, because list discrimination did not differ for foils encountered under FA compared with DA, even foils presented under FA may have received relatively little processing during the initial test, or at least little of the type of processing necessary for performing the challenging list discrimination task.

The present data also suggest that dividing attention during memory retrieval is detrimental for subsequent remembering of studied items. Although pictures that were initially tested under DA were more likely to be recognized than pictures that were initially untested, they were not as likely to be recognized as pictures that were initially tested under full attention. Moreover, source memory and list discrimination decisions were superior for pictures that were previously tested under full versus divided attention.

Although it is well established that memory tests are particularly effective study events (for a review, see Roediger & Karpicke, 2006a), the results of the present experiments support the hypothesis that the power of retrieval as an encoding event is attention dependent. This seems

to be true both for item recognition and for source recollection. Importantly, pictures that were recognized on the first test were less likely to be recognized on the second test when attention was divided during initial retrieval. Similarly, when the source of a picture was correctly identified on the first test, this source was less likely to be accurately remembered on the second test if attention had been divided during initial remembering. Because participants were given (and used) an "unsure" option for source memory decisions on both tests, it is unlikely that many of the correct source responses from the initial test reflect guessing or that the observed retrieval benefits can be attributed to the same mnemonic processes that gave rise to better subsequent memory for items that were not accurately recognized on the initial test.

In a follow-up study, we found similar effects using verbal stimuli, suggesting that these effects are not limited to pictorial stimuli. Object names that were initially retrieved under DA were less likely to be retrieved on a later test compared with object names initially retrieved under FA. Similarly, memory for the encoding task, or source memory, was impaired on a second test if it had previously been retrieved under DA versus FA. This replication suggests that the detrimental effects of distraction on future remembering may extend to multiple forms of study materials, highlighting a potential implication for educational situations in which students test themselves on course material under conditions of distraction.

Our results suggest that distraction during memory retrieval disrupts memory processes that give rise to later item recognition and source memory. Recent research has shown that testing enhances subsequent recollection but not familiarity (Chan & McDermott, 2007), suggesting that divided attention during testing should impair tasks that require recollection to a greater extent than it should impair tasks that rely more on familiarity processes. The relatively deep semantic encoding tasks used in the present experiments likely promoted the role of both recollection and familiarity processes in item recognition (e.g., Hicks & Marsh, 2000). Future research could elucidate the effects of distraction during memory retrieval on subsequent recollection- versus familiarity-based memories.

Additionally, in the present experiments, re-exposure to studied pictures during the initial memory test may have provided additional opportunities to strengthen item representations that enhanced subsequent familiarity, and this could be true even for items that were accurately retrieved on the initial test. Thus, it is unclear whether the observed decrements in subsequent recognition memory for pictures that were retrieved under DA can be fully accounted for by the effects of DA on memory retrieval processes as opposed to the effects of encoding the second presentation of the picture. However, because source information was not presented during the initial test (it was only retrieved), the decreased likelihood of subsequent source retrieval for pictures previously retrieved under DA suggests that encoding of the second presentation of the pictures cannot fully account for our effects. Future research could further specify the precise conditions under which an initial test

affects subsequent item and source memory and could explore the role of attention in these processes.

The results of the present experiments provide novel insight about the power of retrieval as an encoding event, demonstrating that recognition memory and source memory testing benefits are attention dependent. Distractions during memory retrieval reduce the likelihood that a successful act of remembering in the present will support successful remembering in the future, highlighting the importance of attentive remembering for learning.

AUTHOR NOTE

This research was supported by the National Institute of Mental Health (R01-MH076932), the National Alliance for Research on Schizophrenia and Depression, and the Alfred P. Sloan Foundation. We thank Arpeet Shah and Meredith Ayres for assisting with data collection and Elizabeth Kensinger and Ian Dobbins for providing stimuli. We also thank Fergus Craik and two anonymous reviewers for their insightful comments on a version of the manuscript. Address correspondence to N. M. Dudukovic, Department of Psychology, Trinity College, Hartford, CT 06106 (e-mail: nicole.dudukovic@trincoll.edu).

REFERENCES

- BADDELEY, A., LEWIS, V., ELDRIDGE, M., & THOMSON, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, **113**, 518-540. doi:10.1037/0096-3445.113.4.518
- BATCHELDER, W. H., & RIEFER, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, **97**, 548-564. doi:10.1037/0033-295X.97.4.548
- BAYEN, U. J., MURNANE, K., & ERDFELDER, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 197-215. doi:10.1037/0278-7393.22.1.197
- BUTLER, A. C., MARSH, E. J., GOODE, M. K., & ROEDIGER, H. L., III (2006). When additional multiple-choice lures aid versus hinder later memory. *Applied Cognitive Psychology*, **20**, 941-956. doi:10.1002/acp.1239
- CARRIER, M., & PASHLER, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, **20**, 633-642.
- CHAN, J. C. K., & McDERMOTT, K. B. (2007). The testing effect in recognition memory: A dual process account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 431-437. doi:10.1037/0278-7393.33.2.431
- 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. doi:10.1037/0096-3445.125.2.159
- FERNANDES, M. A., & MOSCOVITCH, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General*, **129**, 155-176. doi:10.1037/0096-3445.129.2.155
- FERNANDES, M. A., & MOSCOVITCH, M. (2002). Factors modulating the effect of divided attention during retrieval of words. *Memory & Cognition*, **30**, 731-744.
- FERNANDES, M. A., & MOSCOVITCH, M. (2003). Interference effects from divided attention during retrieval in younger and older adults. *Psychology & Aging*, **18**, 219-230. doi:10.1037/0882-7974.18.2.219
- GARDINER, J. M., & KLEE, H. (1976). Memory for remembered events: An assessment of output monitoring in free recall. *Journal of Verbal Learning & Verbal Behavior*, **15**, 227-233. doi:10.1016/S0022-5371(76)90042-6
- GATES, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, **6**, 1-104.
- HICKS, J. L., & MARSH, R. L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1483-1498. doi:10.1037/0278-7393.26.6.1483

- HOGAN, R. M., & KINTSCH, W. (1971). Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning & Verbal Behavior*, **10**, 562-567. doi:10.1016/S0022-5371(71)80029-4
- JACOBY, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541. doi:10.1016/0749-596X(91)90025-F
- JACOBY, L. L., SHIMIZU, Y., DANIELS, K. A., & RHODES, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, **12**, 852-857.
- JACOBY, L. L., SHIMIZU, Y., VELANOVA, K., & RHODES, M. G. (2005). Age differences in depth of retrieval: Memory for foils. *Journal of Memory & Language*, **52**, 493-504. doi:10.1016/j.jml.2005.01.007
- JOHNSON, M. K., HASHTRUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28. doi:10.1037/0033-2909.114.1.3
- KENSINGER, E. A., CLARKE, R. J., & CORKIN, S. (2003). What neural correlates underlie successful encoding and retrieval? A functional magnetic resonance imaging study using a divided attention paradigm. *Journal of Neuroscience*, **23**, 2407-2415.
- KNOTT, L. M., & DEWHURST, S. A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition*, **35**, 1954-1965.
- LANDAUER, T. K., & BJORK, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625-632). London: Academic Press.
- LOZITO, J. P., & MULLIGAN, N. W. (2006). Exploring the role of attention during memory retrieval: Effects of semantic encoding and divided attention. *Memory & Cognition*, **34**, 986-998.
- MARSH, R. L., HICKS, J. L., HANCOCK, T. W., & MUNSAYAC, K. (2002). Investigating the output monitoring component of event-based prospective memory performance. *Memory & Cognition*, **30**, 302-311.
- MULLIGAN, N. W. (1998). The role of attention during encoding in implicit and explicit memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 27-47. doi:10.1037/0278-7393.24.1.27
- MULLIGAN, N. W. (2004). Generation and memory for contextual detail. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 838-855. doi:10.1037/0278-7393.30.4.838
- MURNANE, K., & BAYEN, U. J. (1996). An evaluation of empirical measures of source identification. *Memory & Cognition*, **24**, 417-428.
- 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 asymmetry. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1091-1104. doi:10.1037/0278-7393.24.5.1091
- NAVEH-BENJAMIN, M., CRAIK, F. I. M., PERRETTA, J. G., & TONEY, S. T. (2000). The effects of divided attention on encoding and retrieval processes: The resiliency of retrieval processes. *Quarterly Journal of Experimental Psychology*, **53A**, 609-625. doi:10.1080/027249800410454
- ROEDIGER, H. L., III, & KARPICKE, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, **1**, 181-210. doi:10.1111/j.1745-6916.2006.00012.x
- ROEDIGER, H. L., III, & KARPICKE, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, **17**, 249-255. doi:10.1111/j.1467-9280.2006.01693.x
- ROEDIGER, H. L., III, & MARSH, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 1155-1159. doi:10.1037/0278-7393.31.5.1155
- SIMONS, J. S., KOUTSTAAL, W., PRINCE, S., WAGNER, A. D., & SCHACTER, D. L. (2003). Neural mechanisms of visual object priming: Evidence for perceptual and semantic distinctions in fusiform cortex. *Neuro-Image*, **19**, 613-626. doi:10.1016/S1053-8119(03)00096-X
- SKINNER, E. I., & FERNANDES, M. A. (2008). Interfering with remembering and knowing: Effects of divided attention at retrieval. *Acta Psychologica*, **127**, 211-221. doi:10.1016/j.actpsy.2007.05.001
- STAHL, C., & KLAUER, K. C. (2007). HMMTree: A computer program for latent-class hierarchical multinomial processing tree models. *Behavior Research Methods*, **39**, 267-273.
- SUGIMORI, E., & KUSUMI, T. (2009). Limiting attentional resources influences performance and output monitoring of an event-based prospective memory task. *European Journal of Cognitive Psychology*, **21**, 112-128. doi:10.1080/09541440802003108
- TROYER, A. K., WINOCUR, G., CRAIK, F. I. M., & MOSCOVITCH, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology*, **13**, 467-474. doi:10.1037/0894-4105.13.4.467
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychology*, **26**, 1-12. doi:10.1037/h0080017
- WHEELER, M. A., EWERS, M., & BUONANNO, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory*, **11**, 571-580. doi:10.1080/09658210244000414
- YONELINAS, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517. doi:10.1006/jmla.2002.2864

(Manuscript received June 26, 2008;
revision accepted for publication April 26, 2009.)