

# Parallel Memory Retrieval in Dual-Task Situations: I. Semantic Memory

Gordon D. Logan and Matthew D. Schulkind  
University of Illinois

Can participants retrieve information about the 2nd of 2 stimuli while they are processing the 1st? Four experiments suggest they can. Reaction times to the 1st stimulus were faster if it came from the same category as the 2nd than if it came from a different category. This category-match effect was observed for letter–digit discrimination (Experiment 1), magnitude and parity judgments about digits (Experiment 2), and lexical decisions (Experiment 3). Experiment 4 showed that the 2nd stimulus could semantically prime the 1st. The category-match effect was observed only when the same task was performed on the 2 stimuli. When the task changed from the 1st stimulus to the 2nd, there was no advantage of a category match. This dependence on task set may explain previous failures to find parallel retrieval.

This article is concerned with the possibility of parallel memory retrieval in dual-task situations. Can people retrieve information about one stimulus while they are busy processing another? This question has been important in the memory literature for decades, and many studies have addressed it throughout that period (see below). Many of the studies had procedural limitations that undermined their ability to address the question of parallel retrieval precisely. Two recent studies overcame these limitations by using the *psychological refractory period* (PRP) procedure, but they suggested opposite conclusions: Carrier and Pashler (1995) argued that memory retrieval was strictly serial, whereas Hommel's (1998) results suggest that memory retrieval may be parallel. Our research addressed procedural, analytic, and logical differences between the studies and found evidence supporting parallel retrieval that has implications for theories of the PRP as well as theories of memory.

We should make clear that we are addressing parallel retrieval between tasks rather than within tasks. Most formal theories of memory assume parallel retrieval within tasks, in that a single memory probe accesses all traces in memory concurrently (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Humphreys, Bain, & Pike, 1989; Murdock, 1982, 1983, 1993; Raaijmakers & Shiffrin, 1981; Shiffrin & Steyvers, 1997). In some contexts, such as free recall, memory may be probed successively with different retrieval cues, but each probe accesses all of the traces in parallel (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981). The question of parallel retrieval between tasks is logically separate from the

question of parallel retrieval within tasks, so parallel retrieval within tasks does not imply parallel retrieval between tasks. Parallel retrieval within tasks is compatible with parallel retrieval between tasks as we suggest in the General Discussion.

## Dual-Task Studies of Memory Retrieval

### *A Brief History*

In 1968, Atkinson and Shiffrin drew a distinction between memory structures and the control processes that operated on them. Many aspects of encoding and retrieval were thought to be strategic and, consequently, dependent on attention. In the early 1970s, several studies used dual-task techniques to assess the attention demands of encoding and retrieval processes (e.g., Johnston, Wagstaff, & Griffith, 1972; Martin, 1970; Trumbo & Milone, 1971). For example, Johnston, Greenberg, Fisher, and Martin (1970) asked participants to track a visual target while listening to words to be remembered later (i.e., during encoding) and while listening to words they made recognition judgments about (i.e., during retrieval). Typical of similar studies, they found that tracking error was larger during retrieval than during encoding, suggesting that retrieval demanded more attention.

Subsequent researchers have focused more on the memorial consequences of dual-task performance, engaging participants in more continuous concurrent tasks and examining memory performance rather than performance on the concurrent task (e.g., Baddeley, Lewis, Eldridge, & Thompson, 1984; Jacoby, 1991; Moscovitch, 1994; Park, Smith, Dudley, & Lafronza, 1989). For example, Craik, Govoni, Naveh-Benjamin, and Anderson (1996) asked participants to perform a serial reaction-time (RT) task during encoding and retrieval. Typical of similar studies, they found that a concurrent task at encoding interferes with subsequent recognition performance more than performing a concurrent task at retrieval time.

From these results, it is difficult to tell whether dual-task conditions are worse at encoding or at retrieval. It is tempting to interpret the differences between encoding and retrieval in terms of attention allocation strategies: At encoding, participants give higher priority to the concurrent task than to encoding, which results in good concurrent-task performance and poor memory on

---

Gordon D. Logan and Matthew D. Schulkind, Department of Psychology, University of Illinois.

This research was supported by National Science Foundation Grants SBR 9709711 and SBR 9808971. We are grateful to Julie Delheimer for testing the participants, to Jim Townsend and Rich Schweickert for valuable discussions, and to Tram Neill and two anonymous reviewers for challenging comments on the article.

Correspondence concerning this article should be addressed to Gordon D. Logan, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820. Electronic mail may be sent to [glogan@s.psych.uiuc.edu](mailto:glogan@s.psych.uiuc.edu).

a subsequent test. At retrieval, participants give higher priority to the memory task, so retrieval is accurate and concurrent-task performance suffers.

Regardless of the reason for the differences between encoding and retrieval, the results suggest that retrieval can go on in parallel with other activities. None of the concurrent tasks reduced recall or recognition to chance levels, as they would if the concurrent task prevented retrieval. However, three limitations in the procedures undermine that conclusion.

First, many of the studies have not reported performance on the concurrent task, and those that have reported it have often provided summary measures averaged over whole blocks of trials. The interpretation of these studies hinges on the assumption that there were no trade-offs between the concurrent task and memory task. The interpretation assumes that participants performed the concurrent task equally well in each condition of the memory task or that concurrent task performance was worse when memory performance was worse. It assumes that concurrent task performance was not better when memory performance was worse, or vice versa, which would reflect a strategic trade-off between tasks. Without data on concurrent-task performance, it is impossible to tell whether these assumptions are valid. Trade-offs between tasks abound in the attention literature (e.g., Shulman & Fisher, 1972; Strayer & Kramer, 1990), so it is possible, if not likely, that they also occur in the memory literature.

Second, the relative timing of the concurrent tasks was not controlled in some studies and not reported in others. Relative timing is perhaps the most important independent variable in dual-task situations. It is very difficult to perform two tasks if their stimuli appear simultaneously—that is, if the *stimulus onset asynchrony* (SOA) is 0—and much easier to perform them if even small amounts of time elapse between stimulus onsets (e.g., SOAs of 500–1,000 ms; Duncan, 1980; Moray, Fitter, Ostry, Favreau, & Nagy, 1976; Pohlmann & Sorkin, 1976; for a review, see Pashler, 1994). It is possible that participants were unable to retrieve memory items when the tasks were truly concurrent (i.e., SOA = 0) but could retrieve them reasonably well when the tasks followed one another (i.e., SOA > 0). Averaging over long and short SOAs obscures this effect, and publishing only averaged data makes it impossible for subsequent researchers to recover the SOA effects.

Third, in most of the studies, accuracy was used to measure memory performance. This is reasonable, of course, because accuracy was well below ceiling, but it presents a problem in attempting to distinguish between parallel and serial retrieval, because most of the tests that distinguish parallel from serial processes address RT rather than accuracy (see, e.g., Egeth & Dagenbach, 1991; Rohrer, Pashler, & Etchegaray, 1998; Ross & Anderson, 1981; Townsend, 1990). In particular, the methods used in the Carrier and Pashler (1995) and Hommel (1998) studies addressed RT rather than accuracy.

### PRP Procedure

The PRP procedure involves presenting two stimuli, Stimulus 1 (S1) and Stimulus 2 (S2), with some SOA between them, usually ranging from 0 to 1,000 ms. Participants are told to perform Task 1 on S1, making Response 1 (R1) to indicate their decision about S1, with latency RT1. They are told to perform Task 2 on S2, making Response 2 (R2) to indicate their decision about it, with latency

RT2. The stimuli are punctate and the responses are discrete, so timing can be controlled very carefully and measured very accurately. Typically, RT1 is unaffected by SOA, which reflects the priority that participants give to Task 1, and RT2 is strongly affected by SOA, increasing markedly as SOA decreases, which reflects interference from Task 1 on Task 2. These effects are robust. They have been replicated in dozens of experiments over the last 50 years (for reviews, see Bertelson, 1966; Kahneman, 1973; Pashler, 1994; Smith, 1967; Welford, 1952).

*Carrier and Pashler (1995).* Carrier and Pashler (1995) reported two PRP experiments on parallel retrieval, in which Task 2 involved retrieval from episodic memory. In Experiment 1, Task 1 was tone discrimination. Participants had to report which of two tones (high or low) was presented. Task 2 was cued recall. Participants saw a word and had to report the word that was associated with it during a previous study period. The difficulty of retrieval was manipulated by varying the number of presentations (1 vs. 2). The experiment replicated the basic PRP results. RT1 was unaffected by SOA, and RT2 increased sharply as SOA decreased. The key effect was the interaction between retrieval difficulty and SOA. The interaction was clearly null; the effects were additive.

Carrier and Pashler's (1995) Experiment 2 was a replication of Experiment 1 with recognition as the retrieval task rather than cued recall. Task 1 was the same tone discrimination task used in Experiment 1. S2 was a word, and participants had to decide whether or not it was on the list they had studied. The difficulty of retrieval was manipulated by varying the number of presentations (0 for new items, 1 or 5 for old items). Again, the experiment replicated the main PRP results, and again, the interaction between retrieval difficulty and SOA was clearly null. The effects were additive.

Carrier and Pashler (1995) used the *locus of slack* logic (Pashler, 1984; Pashler & Johnston, 1989; Schweickert, 1978; Schweickert & Townsend, 1989; Townsend & Schweickert, 1989) to determine whether retrieval could occur in parallel with Task 1 processing. The logic assumes that each task can be described as a series of stages leading from stimulus to response (e.g., Sternberg, 1969) and that processing in one of the stages requires a central bottleneck that can deal with only one task at a time. Task 2 processing prior to the bottleneck can proceed in parallel with Task 1. Task 2 processing after the bottleneck must wait for Task 1 to be finished with the bottleneck before it can proceed. The logic focuses on interactions between Task 2 difficulty factors and SOA. If the difficulty factor affects a stage in Task 2 before the bottleneck and SOA is short, both the easy and the hard versions may have time to finish their processing before Task 2 gets access to the bottleneck, in which case the difficulty factor will have no effect on RT2. Long SOAs provide the controls against which the effects at short SOAs are evaluated. At a long SOA, neither the easy nor the hard version has to wait to get access to the central bottleneck, so the Task 2 difficulty effect should be as large as it is in single-task conditions. Put together, these effects create an *underadditive* interaction between Task 2 difficulty and SOA.

If the difficulty factor affects a Task 2 stage that is at or after the bottleneck, SOA and difficulty should have additive effects. Bottleneck and postbottleneck processes in Task 2 must wait until Task 1 is finished with the bottleneck before they begin, so the easy version and the hard version will always begin at the same time and end at different times. The differences in finishing time

should be the same whether Task 2 has to wait at short SOAs or not wait at long SOAs. The interaction between difficulty and SOA should be null.

On the basis of the locus of slack logic, Carrier and Pashler (1995) interpreted the additive effects of retrieval difficulty and SOA as evidence that retrieval was a bottleneck or postbottleneck process. This led them to conclude that retrieval cannot go on in parallel with other processes that also require the bottleneck. Their experiments are a serious challenge to the hypothesis that retrieval can go on in parallel with other activities, because they rely on rigorous logic and the data are very clear and convincing.

*Hommel (1998).* Hommel (1998) conducted five experiments in which participants reported the color or the identity (or both) of a red or green *H* or *S*. In Experiment 1 participants made manual responses to color, pressing keys with their left and right hands to indicate which color the letter was, and they made vocal responses to letter identity, saying "left" if it was an *H* and "right" if it was an *S*, for example. The main result was a 75-ms R2-R1 compatibility effect on RT1: Participants were faster to press the left key if the response to the letter was also "left" than if it was "right." Similarly, right keypresses were faster if the vocal response was "right" than if it was "left."

In Experiment 2, Hommel (1998) again presented red or green *H*'s or *S*'s and had participants respond manually to the color and vocally to the letter identity. The manual responses were left and right keypresses, as before, but the vocal responses were "red" and "green" (e.g., "red" for *H* and "green" for *S*). This experiment showed a 104-ms R2-S1 compatibility effect on RT1, in that manual responses to the color were faster if the vocal response matched the color (e.g., "red" and red) than if it did not match (e.g., "red" and green). Experiment 3 was a replication of Experiment 2, except that the color was first presented on a rectangle, which changed into a letter of the same color at some SOA (50, 150, or 650 ms). Task 1 was to respond manually to the color, and Task 2 was to respond vocally to letter identity (with "red" and "green" as responses). A 28-ms R2-S1 compatibility effect appeared at the 50-ms SOA, albeit weaker than in the previous experiment (in which SOA was 0).

Experiment 4 returned to the colored letter stimuli (i.e., SOA = 0), but required participants to respond rapidly only to color. The vocal response to identity (again "red" or "green") was to be withheld for 1,500 ms after stimulus onset. The R2-S1 compatibility effect replicated once again (29 ms). Finally, Experiment 5 required participants to prepare both responses but execute only one of them on each trial. The response to be executed was indicated by a cue that appeared before the colored letter. The compatibility effect appeared once again. Participants were 22 ms faster if the manual response they executed to the color matched the color name they would have given the letter than if it did not match, and they were 39 ms faster if the vocal response matched the (unreported) color.

Hommel's (1998) compatibility effects are important, because they suggest that participants were able to retrieve R2 while they were busy processing S1. The nature of R2 interacted with the manual responses (Experiment 1) and with the perceptual discriminations or semantic judgments (Experiments 2–5) required for the first task, speeding RT1 if they matched and slowing it if they mismatched. It seems clear that the compatibility effects were due to R2 rather than S2. The second stimulus attribute was always

letter identity, and it was always unrelated to the first stimulus attribute (color) and unassociated with the manual responses required for the first attribute. By contrast, the response to the second attribute was either the same as the response given to the first ("left" and left hand) or the same as the categorization given to the first ("red" and red). In order for R2 to affect RT1, participants must have retrieved it at least partially before they finished processing S1.

More formally, Hommel's (1998) results address a fundamental distinction between serial and parallel processes described by Townsend and Ashby (1983). In a serial system, the current process cannot be affected by processes that have not yet finished, because those processes have not yet begun (they cannot begin until the current process is finished). In a parallel system, the current process can be affected by all processes that are currently active, and those are the ones that have not yet finished. Hommel's R2-S1 and R2-R1 compatibility effects show that Task 2 processes are active while Task 1 processes are active, and that is consistent with parallel processing and inconsistent with serial processing.

*Critical differences.* Carrier and Pashler's (1995) experiments differ from Hommel's (1998) experiments in important aspects of the procedure and in the logic of the analysis. First, Carrier and Pashler used a broad range of SOAs, whereas Hommel used only one (SOA = 0) in four of his five experiments. It is important to replicate effects like Hommel's over a broader range of SOAs.

Second, in Carrier and Pashler's (1995) experiments, the task set changed from Task 1 to Task 2. Task 1 involved tone discrimination, while Task 2 involved cued recall or recognition. The two task sets had nothing in common. By contrast, the task sets in Hommel's (1998) experiments overlapped in a very specific manner: The responses for Task 2 were related to the stimuli or the responses to Task 1. It is possible that the task sets must overlap before parallel retrieval can occur. The set for Task 1 may have to include some elements in common with the set for Task 2. Put differently, a change in task set from Task 1 to Task 2 may block parallel retrieval. Retrieval may be only *conditionally automatic* (Bargh, 1992; DeJong, Liang, & Lauber, 1994; Logan & Etherton, 1994), requiring an appropriate attentional set before it is triggered. The overlapping elements in Hommel's Task 1 and Task 2 may have provided that triggering condition, whereas the nonoverlapping task sets in Carrier and Pashler's experiments may have taken it away.

Third, Carrier and Pashler's (1995) conclusions were derived from the locus of slack logic, which dominates much of the research on the PRP (Pashler, 1994). The locus of slack logic manipulates Task 2 difficulty and looks for interactions with SOA in RT2 to distinguish serial from parallel retrieval (Schweickert, 1978; Schweickert & Townsend, 1989; Townsend & Schweickert, 1989). By contrast, Hommel's (1998) conclusions were based on an analysis of *crosstalk* from Task 2 to Task 1 to distinguish serial from parallel retrieval (Townsend & Ashby, 1983). He manipulated the compatibility of the two tasks and looked for evidence that responses to Task 2 affected RT1.

At first glance, it looks like the two methods of analysis may be compatible. We could use both in one experiment, applying the crosstalk logic to RT1 and the locus of slack logic to RT2. Unfortunately, on deeper analysis, the two methods seem fundamentally incompatible. Situations in which there is crosstalk between Task 1 processes and bottleneck or postbottleneck Task 2

processes violate the basic assumptions that are required to apply the locus of slack logic. Consider, for example, trying to apply the locus of slack logic to Hommel's (1998) Experiment 2. Consider the type of vocal response (saying "red" or "green") as a Task 2 difficulty manipulation. At what level of Task 1 should vocal response types be compared? The results will be quite different if S1 is green than if it is red. Averaging over red S1 and green S1 does not seem appropriate because it obscures an important difference (i.e., the R2-S1 compatibility effect). There does not seem to be an appropriate baseline condition in Hommel's Task 1. By contrast, most successful applications of the locus of slack logic use nonoverlapping task sets for Task 1 and Task 2. There is no crosstalk between Task 1 and Task 2, so any Task 1 condition is an appropriate baseline, and averaging over Task 1 stimuli does not obscure important effects that may influence RT2 (e.g., Carrier & Pashler, 1995).

More formally, evidence of crosstalk between Task 1 and Task 2 questions the assumption of *selective influence* (Sternberg, 1969) on which the locus of slack logic is founded (Schweickert & Townsend, 1989). The locus of slack analysis of Task 2 difficulty effects relies on the assumption that the difficulty manipulation affects one and only one stage. Crosstalk between stages of different tasks suggests that stage durations are correlated such that factors that affect one stage also affect the other in the same manner. That is, factors can no longer selectively influence one stage and not the other. This violation of the assumption of selective influence means that the locus of slack logic cannot be applied properly to situations in which there is crosstalk from Task 2 to Task 1 (e.g., Hommel, 1998), even if appropriate baseline conditions could be found.

### The Present Experiments

The present experiments were intended to bridge the procedural gap between Carrier and Pashler's (1995) experiments and Hommel's (1998). Like Carrier and Pashler's, our experiments used separate stimuli for Task 1 and Task 2 and we manipulated SOA between S1 and S2 (0, 100, 300, or 900 ms). Like Hommel, our experiments used overlapping task sets. Participants performed the same task on S1 and S2. In Experiment 1, for example, S1 was a letter or a digit and so was S2. Participants had to decide whether S1 was a letter or a digit and whether S2 was a letter or a digit. Like Hommel, the main focus of our analysis was on RT1 rather than RT2. We tested for parallel retrieval by examining *category-match effects*: If participants can retrieve the categorization of S2 while they are retrieving the categorization of S1, then both RT1 and RT2 should be faster when the categories match (e.g., both are digits or both are letters) than when they mismatch (e.g., one is a letter and the other is a digit). The effects on RT1 are critical. A serial process cannot be influenced by processes that have not yet finished, because those processes have not yet begun. By contrast, a parallel process can be influenced by processes that have not yet finished, because it can be influenced by all currently active processes. Thus, evidence that the S2 categorization influences RT1 is evidence that retrieval cannot be serial and must be parallel (Townsend & Ashby, 1983).

Like Hommel (1998) and unlike Carrier and Pashler (1995), we focused on retrieval from *semantic memory* rather than *episodic memory* (Tulving, 1983). Participants were asked to discriminate

between letters and digits (Experiment 1), between digit magnitude and parity (Experiment 2), and between words and nonwords (Experiments 3 and 4). We chose these tasks because they clearly required retrieval from memory—there is nothing inherent in letters, digits, words, and nonwords that makes their meaning clear without prior learning. We felt justified in analyzing semantic memory retrieval rather than episodic memory retrieval because the difference between semantic and episodic memory is not very clear. First, most formal models of memory do not distinguish between them (see, e.g., Humphreys et al., 1989; Shiffrin & Steyvers, 1997). The distinction between them becomes blurred from the perspective of instance or exemplar theories (e.g., Hintzman, 1986; Landauer, 1975; Logan, 1988; Nosofsky, 1986). Second, several studies suggest that the similarities between episodic and semantic memory are more important than the differences (e.g., Anderson & Ross, 1980; McKoon & Ratcliff, 1979, 1986; McKoon, Ratcliff, & Dell, 1986). Third, the differences between them may not be very important from the perspective of other processes, such as attention, that are taxed in dual-task situations. Finally, we thought it was important to explore parallel retrieval effects in tasks in which accuracy is high enough for us to focus our analyses on RTs. The analytic techniques we use to make inferences about parallel or serial retrieval focus on RT rather than accuracy. Our semantic memory tasks were relatively simple, so accuracy was not a problem.

### Experiment 1

In the first experiment, participants saw two characters on each trial, one above the other. The task was to indicate whether each character was a letter or a digit. Participants responded to the top character with their right hand and the bottom character with their left one. If participants are able to retrieve the category (letter vs. digit) of S2 while they are processing S1, then RT1 should be faster when the category of S2 matches the category of S1. That is, participants should be faster when S1 and S2 are both letters or both digits than when one is a letter and the other is a digit. If participants are not able to retrieve the category of S2 while processing S1, then RT1 should be no faster when S2 matches the category of S1.

### Method

*Participants.* The participants were 16 volunteers from the general university community who were paid \$5 for participating in one 1-hr session.

*Apparatus and stimuli.* The stimuli were the capital letters Z, B, K, S, G, L, H, and F and the digits 2, 3, 4, 5, 6, 7, 8, and 9. They were displayed on Gateway 2000 Crystalscan 1024 NI monitors controlled by Gateway 2000 486 computers. Responses to the top stimulus were collected from the period and ?/ keys; responses to the bottom stimulus were collected from the Z and X keys. Timing was accurate to 1 ms. SOA was 0, 100, 300, or 900 ms.

Each trial involved three displays if SOA was 0 and four displays if SOA was greater than zero. The first display was a fixation field, which indicated the positions the character stimuli would occupy. It consisted of two rows, each made of two dashes separated by three blanks, centered on the screen. The top row began on row 13 and column 37 of the standard IBM text screen. The bottom row began on row 14 and column 37. The second display contained two characters if SOA was zero and one if SOA was greater than zero. In the one-character display, the character appeared in

the top row indicated by the fixation field in a position that was centered between the two dashes. It appeared in row 13, column 38 of the IBM text screen. In the two-character display, the top character appeared in the same position as in the one-character display and was always identical to that character (i.e., the top character looked like it stayed on the screen throughout the second and third display). The bottom character appeared immediately below the top one in row 14, column 38 of the IBM text screen. The third display on zero SOA trials and the fourth display on greater-than-zero SOA trials was blank.

The fixation display was exposed for 500 ms. On zero-SOA trials, the two-character display was exposed for 1,000 ms. It was then extinguished and replaced with a blank display for a 3,500-ms intertrial interval. On trials with SOA greater than zero, the one-character display was exposed for the duration of the SOA and then it was replaced immediately by the two-character display, which was exposed for 1,000 ms and then replaced with a blank screen for a 3,500-ms intertrial interval.

**Procedure.** The basic design involved 16 trials: letter or digit for S1 × letter or digit for S2 × four SOAs. There were 32 replications of this basic design, which produced a total of 512 trials. The eight letters and eight digits were assigned randomly to trials, with two constraints: Each letter and digit should occur equally often, and no letter or digit was allowed to repeat on given trial (i.e., S1 and S2 were always different).

Participants were given verbal instructions that told them which keys to press in response to S1 and S2. Task 1 was emphasized. They were told to respond to it as quickly as possible and not to wait for S2. They were told to rest their fingers on the appropriate response keys throughout the experiment. They were allowed to rest after every 64 trials.

Participants always responded to S1 with their right hand and to S2 with their left hand. The assignment of keys to stimuli was counterbalanced between subjects. There were four response mappings, LDDL, DLLD, DLDL, and LDLD, where L refers to letter, D refers to digit, and the left-to-right order refers to the middle and index fingers of the left hand and then the index and middle fingers of the right hand, respectively.

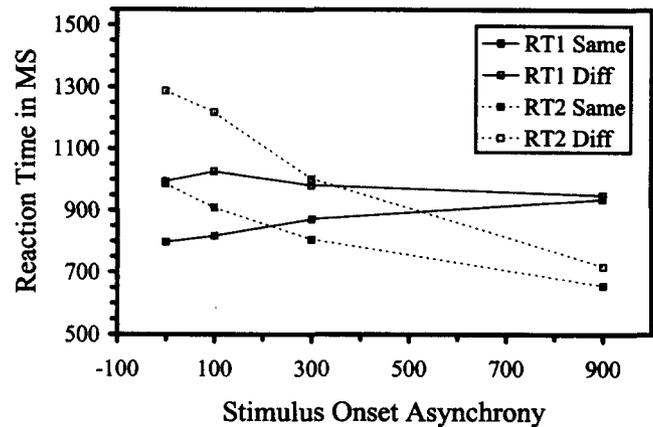
## Results

The main focus of the experiment was the contrast between trials on which S1 and S2 were in the same category (both letters or both digits) and trials on which S1 and S2 were in different categories (one letter and one digit). We computed mean RT in each combination same versus different category and SOA for each task for each subject, including data only if both responses were correct. The means across subjects are plotted in Figure 1. We analyzed the RT data separately for RT1 and RT2, in 2 (category match: same vs. different) × 4 (SOA: 0, 100, 300, and 900 ms) analyses of variance (ANOVAs). The accuracy data were analyzed in a similar ANOVA. Because we selected trials for analysis in which both responses were correct, the accuracy data are the same for Task 1 and Task 2, so we performed a single ANOVA. The accuracy data are presented in Table 1.

The data replicate patterns typical of PRP studies: RT1 was fast and unaffected by SOA, and RT2 was slower, initially at least, and was strongly affected by SOA. More important to our present purposes, there were strong effects of the match between Task 1 category and Task 2 category in both RT1 ( $M = 133$  ms) and RT2 ( $M = 217$  ms). For both RT1 and RT2, the category-match effect was stronger when the SOA was short than when it was long.

These conclusions were supported by ANOVA. In the RT1 data, there was a significant main effect of category match,  $F(1, 15) = 42.70$ ,  $p < .01$ ,  $MSE = 13,116.49$ , and a significant interaction between category match and SOA,  $F(3, 45) = 22.12$ ,  $p < .01$ ,  $MSE = 2,983.65$ , but the main effect of SOA was not

## Letter/Digit Discrimination



**Figure 1.** Reaction time on Task 1 (RT1; solid lines) and reaction time on Task 2 (RT2; broken lines) for category-match (filled squares: same) and category-mismatch (open squares: different) trials as a function of stimulus onset asynchrony in Experiment 1. Diff = different.

significant ( $F < 1$ ). Fisher's LSD, derived from the error term for the interaction between category match and SOA, was 39 ms ( $p < .05$ ). By this criterion, the category-match effect was significant at the 0-, 100-, and 300-ms SOA but not at the 900-ms SOA. The differences were 197, 209, 110, and 14 ms, respectively.

In the RT2 data, the main effect of category match,  $F(1, 15) = 73.68$ ,  $p < .01$ ,  $MSE = 20,483.01$ , the main effect of SOA,  $F(3, 45) = 234.39$ ,  $p < .01$ ,  $MSE = 5,410.56$ , and the interaction between category match and SOA,  $F(3, 45) = 39.68$ ,  $p < .01$ ,  $MSE = 2,697.94$ , were all significant. Fisher's LSD ( $p < .05$ ), based on the Category Match × SOA error term, was 37 ms. By this criterion, the category-match effect was significant at each SOA. The differences were 301, 310, 196, and 62 ms, respectively.

In the accuracy data, there were significant main effects of category match,  $F(1, 15) = 12.12$ ,  $p < .01$ ,  $MSE = 5.70$ , and SOA,  $F(3, 45) = 2.86$ ,  $p < .05$ ,  $MSE = 9.17$ , but there was no significant interaction between them ( $F < 1$ ).

## Discussion

The RT1 data were strongly affected by the match between the category of S1 and S2. This suggests that participants retrieved the categorization of S2 while they were processing S1. Thus, the RT1 data support the parallel retrieval hypothesis (Hommel, 1998) and contradict the hypothesis that retrieval cannot go on in parallel with S1 processing (Carrier & Pashler, 1995). The category-match effects on RT2 are less important to the issue of parallel retrieval. They could be predicted on several grounds, because the response to S2 (R2) followed the response to S1 (R1) on each trial.

The RT1 data were largely unaffected by SOA, which suggests that participants protected Task 1 against interference from Task 2. However, RT1 was quite long, and RT2 was faster than RT1 at the longest SOA. As we shall see, these effects occurred in each experiment, whenever Task 1 and Task 2 were the same. There are several possible explanations for these effects. We address them in

Table 1  
*Mean Percentage of Correct Responses in Task 1 and Task 2 in Experiment 1 as a Function of Category Match and Stimulus Onset Asynchrony (SOA)*

Category match	SOA (ms)			
	0	100	300	900
Same category	94	96	96	96
Different category	93	94	94	95

the General Discussion and conclude that they do not compromise our interpretation of the category-match effects.

### Experiment 2

The second experiment had two purposes. First, it was a conceptual replication of Experiment 1, testing the hypothesis that participants could retrieve categorical information about S2 while they were processing S1, provided that the task set was the same for Task 1 and Task 2. Second, it was intended to test our interpretation of Carrier and Pashler's (1995) failure to find evidence of parallel retrieval. In their experiments, the task set changed from Task 1 to Task 2. Changes in task set produce large costs (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995; Sudevan & Taylor, 1987), which may have obscured retrieval effects or blocked them. Participants may not be able to retrieve information unless an appropriate task set is established.

Participants saw two digits on each trial and processed them with the same or different task sets. One task set was *magnitude judgment*; participants made one response if the digit was greater than five and another if the digit was less than five. The other task set was *parity judgment*; participants made one response if the digit was odd and another if the digit was even. Sudevan and Taylor (1987) showed that the task sets for magnitude and parity judgments could not be maintained simultaneously. Switching between magnitude and parity judgments produced substantial costs that persisted over several sessions of practice.

Our participants served in four sessions, and each session involved a different combination of task sets, formed by factorially combining the two task sets and the two tasks. Thus, in one session, Task 1 and Task 2 both required magnitude judgments; in another, Task 1 and Task 2 both required parity judgments; in another, Task 1 required a parity judgment and Task 2 required a magnitude judgment; and in another, Task 1 required a magnitude judgment and Task 2 required a parity judgment.

As in Experiment 1, we used RT1 category-match effects as our criterion for parallel retrieval. If RT1 is affected by the categorization of S2, the Task 2 retrieval process must have been active at the same time as the Task 1 retrieval process, violating the seriality assumption. If a common task set is necessary to ensure parallel retrieval, we should see an effect of category match on RT1 in the same-set conditions (i.e., Task 1 and Task 2 parity and Task 1 and Task 2 magnitude) but not in the different-set conditions (i.e., Task 1 parity–Task 2 magnitude and Task 1 magnitude–Task 2 parity).

### Method

**Participants.** The participants were 8 volunteers from the general university community who were paid \$20 for serving in four 1-hr sessions. One subject also served in Experiment 1.

**Apparatus and stimuli.** The apparatus was the same as in Experiment 1. The stimuli were the digits 1, 2, 3, 4, 6, 7, 8, and 9 (note that 5 was excluded). The temporal and spatial parameters were the same as in Experiment 1.

**Procedure.** Each session involved 512 trials, formed from two replications of a basic 256-trial design (8 digits for S1  $\times$  8 digits for S2  $\times$  4 SOAs). The stimuli were the same for each task; only the instructions varied. Participants performed a different combination of tasks on each session (i.e., parity–parity, magnitude–magnitude, parity–magnitude, and magnitude–parity for Task 1 and Task 2, respectively). The order of tasks was determined by a balanced Latin square. All participants pressed the keys under their index fingers for high and even, respectively, and the keys under their middle fingers for low and odd, respectively.

Participants were told about the task combination for each day at the beginning of the session. They were told to respond as quickly as possible to S1 and not to wait for S2 before responding.

### Results

Mean RT was computed for RT1 and RT2 for each session on trials on which both responses were correct. The mean RTs are plotted in Figure 2; the accuracy data are presented in Table 2. The RT1 and RT2 data were analyzed in separate 2 (judgment: magnitude vs. parity)  $\times$  2 (task set: same vs. different)  $\times$  2 (category match: same vs. different)  $\times$  4 (SOA: 0, 100, 300, or 900 ms) ANOVAs. The accuracy data were analyzed in a single ANOVA with the same design. The ANOVA summaries are presented in Table 3.

The results in each task combination were typical of PRP studies: RT1 was unaffected by SOA, and RT2 was strongly affected. There was a large category-match effect on RT1 in the same-task-set conditions, averaging 107 ms in the magnitude–magnitude condition and 129 ms in the parity–parity condition, suggesting that category information about S2 was retrieved during S1 processing. The category-match effect on RT1 was negligible in the different-task-set conditions, averaging  $-18$  ms in the magnitude–parity condition and 14 ms in the parity–magnitude condition, suggesting that category information about S2 was not retrieved during S1 processing.

These conclusions were supported by inferential statistics. In the RT1 ANOVA, the main effects of task set and category match were significant, as was the interaction between them. Fisher's LSD, computed with the error term for the interaction between judgment, task set, category match, and SOA, indicated that differences larger than 44 ms were significant ( $p < .05$ ). By this criterion, the category-match effect was significant at the 0-, 100-, and 300-ms SOAs but not at the 900-ms SOA for same-task-set magnitude judgments, and it was not significant at any SOA for different-task-set magnitude judgments. By the LSD ( $p < .05$ ) criterion, the category-match effect was significant at the 0-, 100-, and 300-ms SOAs but not at the 900-ms SOA for same-task-set parity judgments, and it was not significant at any SOA for different-task-set parity judgments.

Similar effects were found for RT2, although they were less interesting theoretically: The category-match effect was large in the same-task-set conditions ( $M_s = 146$  and 180 ms in the

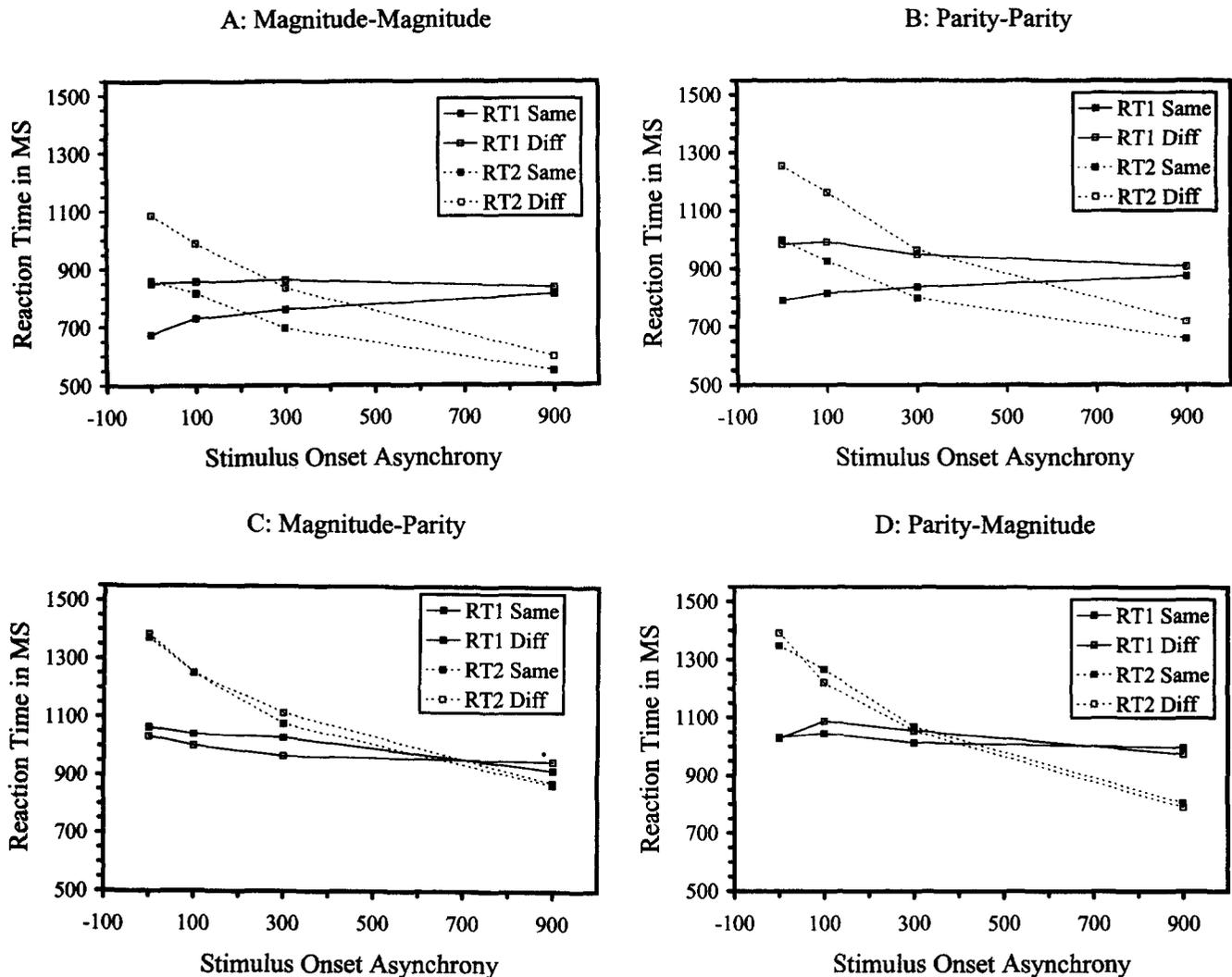


Figure 2. Reaction time on Task 1 (RT1; solid lines) and reaction time on Task 2 (RT2; broken lines) for category-match (filled squares: same) and category-mismatch (open squares: different) trials as a function of stimulus onset asynchrony in Experiment 2. A: Same task set for Stimulus 1 (S1) and Stimulus 2 (S2; magnitude judgment). B: Same task set for S1 and S2 (parity judgment). C: Different task sets for S1 and S2 (Task 1 = magnitude judgment, Task 2 = parity judgment). D: Different task sets for S1 and S2 (Task 1 = parity judgment, Task 2 = magnitude judgment). Diff = different.

magnitude–magnitude and parity–parity conditions, respectively) and small in the different-task-set conditions ( $M_s = 14$  and  $33$  ms in the magnitude–parity and parity–magnitude conditions, respectively).

These conclusions were supported by inferential statistics. There was a strong main effect of SOA and significant main effects for task set and category match. The interaction between task set and category match was significant as well. Fisher's LSD ( $p < .05$ ) based on the error term for the highest order interaction was 48 ms. By this criterion, the category-match effect was significant at every SOA for same-task-set magnitude judgments, and it was not significant at any SOA for different-task-set magnitude judgments. By the LSD ( $p < .05$ ) criterion, the category-match effect was significant at every SOA for same-task-set parity judgments, and it

was significantly negative at the 0-ms SOA and nonsignificant at the other SOAs for different-task-set parity judgments.

The accuracy data showed effects that were similar to the RT data, suggesting no problems with speed–accuracy trade-offs. The LSD ( $p < .05$ ) for the highest order interaction was 2%.

### Discussion

This experiment revealed an important boundary condition on the category-match effect observed in Experiment 1: The category-match effect depends on the adoption of the same task set for Task 1 and Task 2. It can be eliminated by changing task set from Task 1 to Task 2. The positive category-match effect we observed in the same-set conditions extends and confirms the conclusions of

Table 2  
Mean Percentage of Correct Responses in Task 1 and Task 2 in Experiment 2 as a Function of Task Set, Category Match, and Stimulus Onset Asynchrony (SOA)

Category match	SOA (ms)			
	0	100	300	900
<b>Magnitude–magnitude</b>				
Same category	97	99	98	98
Different category	98	98	98	99
<b>Magnitude–parity</b>				
Same category	95	93	94	94
Different category	94	93	94	95
<b>Parity–parity</b>				
Same category	98	96	97	98
Different category	96	95	97	96
<b>Parity–magnitude</b>				
Same category	94	95	96	96
Different category	95	96	93	95

Experiment 1: It is possible for participants to retrieve information about S2 while they are engaged in processing S1. The null category-match effect we observed in the different-set conditions provides an explanation of Carrier and Pashler's (1995) failure to find evidence of parallel retrieval: Their participants switched task sets from Task 1 (tone discrimination) to Task 2 (recognition or recall).

RT for both tasks was strongly affected by task set. For Task 1, same-task-set RTs were 170 ms faster than different-task-set RTs; for Task 2, same-task-set RTs were 258 ms faster than different-task-set RTs. These effects are important and deserve comment, which we defer to the General Discussion.

### Experiment 3

Experiment 1 showed that letter–digit categorizations could be retrieved in parallel, and Experiment 2 showed that magnitude and parity categorizations could be retrieved in parallel, provided that the task set was the same for Task 1 and Task 2. Experiment 3 examined word–nonword categorizations in a pair of lexical-decision tasks. The purpose was to test the generality of the conclusions in Experiments 1 and 2 by attempting to extend them to more complex stimuli (five characters vs. one) and a more complex category (i.e., words). Experiments 1 and 2 involved a small set of stimuli that were repeated many times throughout the experiment. It is possible that multiple repetitions are necessary to produce parallel retrieval. Experiment 3 addressed that possibility by presenting stimuli only once during the experiment.

Experiment 3 manipulated the difficulty of the lexical decision between subjects. One group had to discriminate words from pronounceable nonwords, which is relatively difficult, and another group had to discriminate words from unpronounceable nonwords, which is relatively easy.

### Method

**Participants.** Two groups of 16 volunteers from the general university community participated in the experiment. Each received \$5 for a 1-hr session. No participant had taken part in either of the previous experiments.

**Apparatus and stimuli.** The stimuli were 339 five-letter words and five-letter nonwords. The words were five-letter nouns selected from the Kucera and Francis (1967) word frequency norms, with a mean frequency of 75.27 per million and a range of 8 to 787 per million. Two nonwords were paired with each word, one pronounceable and the other unpronounceable (e.g., *world*, *wirld*, and *wrsld*). Pronounceable nonwords were constructed by substituting vowels for vowels and consonants for consonants in a manner that produced a letter string that the first author could

Table 3  
Summaries for Analyses of Variance for RT1, RT2, and Accuracy in Experiment 2

Source	MSE	df	F
<b>RT1</b>			
Judgment (J)	146,473.21	1, 7	1.58
Task set (T)	153,204.31	1, 7	12.15*
J × T	18,575.16	1, 7	4.16
Category match (C)	25,193.69	1, 7	8.55*
J × C	8,533.07	1, 7	1.34
T × C	12,658.67	1, 7	18.42**
J × T × C	3,806.02	1, 7	0.13
Stimulus onset asynchrony (S)	44,172.53	3, 21	0.39
J × S	2,207.66	3, 21	1.09
T × S	5,309.15	3, 21	7.33**
C × S	1,513.04	3, 21	9.93**
J × T × S	3,638.94	3, 21	3.19*
J × C × S	1,808.54	3, 21	3.08*
T × C × S	3,783.93	3, 21	6.06**
J × T × C × S	1,767.32	3, 21	1.63
<b>RT2</b>			
J	93,791.11	1, 7	1.30
T	128,496.56	1, 7	33.08**
J × T	33,763.68	1, 7	14.50**
C	19,348.08	1, 7	19.53**
J × C	4,504.62	1, 7	0.17
T × C	19,495.63	1, 7	24.57**
J × T × C	9,308.17	1, 7	2.73
S	14,280.47	3, 21	189.20**
J × S	4,313.05	3, 21	1.30
T × S	5,163.25	3, 21	5.09**
C × S	3,772.99	3, 21	5.63**
J × T × S	3,153.54	3, 21	1.26
J × C × S	2,169.05	3, 21	0.34
T × C × S	3,396.38	3, 21	10.06**
J × T × C × S	2,158.92	3, 21	0.57
<b>Accuracy</b>			
J	61.52	1, 7	0.21
T	18.17	1, 7	33.04**
J × T	31.97	1, 7	2.90
C	8.40	1, 7	0.74
J × C	1.72	1, 7	13.81**
T × C	14.84	1, 7	0.21
J × T × C	6.88	1, 7	0.66
S	8.27	3, 21	0.72
J × S	6.00	3, 21	0.05
T × S	6.51	3, 21	0.16
C × S	6.24	3, 21	0.03
J × T × S	6.78	3, 21	2.90
J × C × S	7.22	3, 21	0.64
T × C × S	9.08	3, 21	1.03
J × T × C × S	4.20	3, 21	2.32

\*  $p < .05$ . \*\*  $p < .01$ .

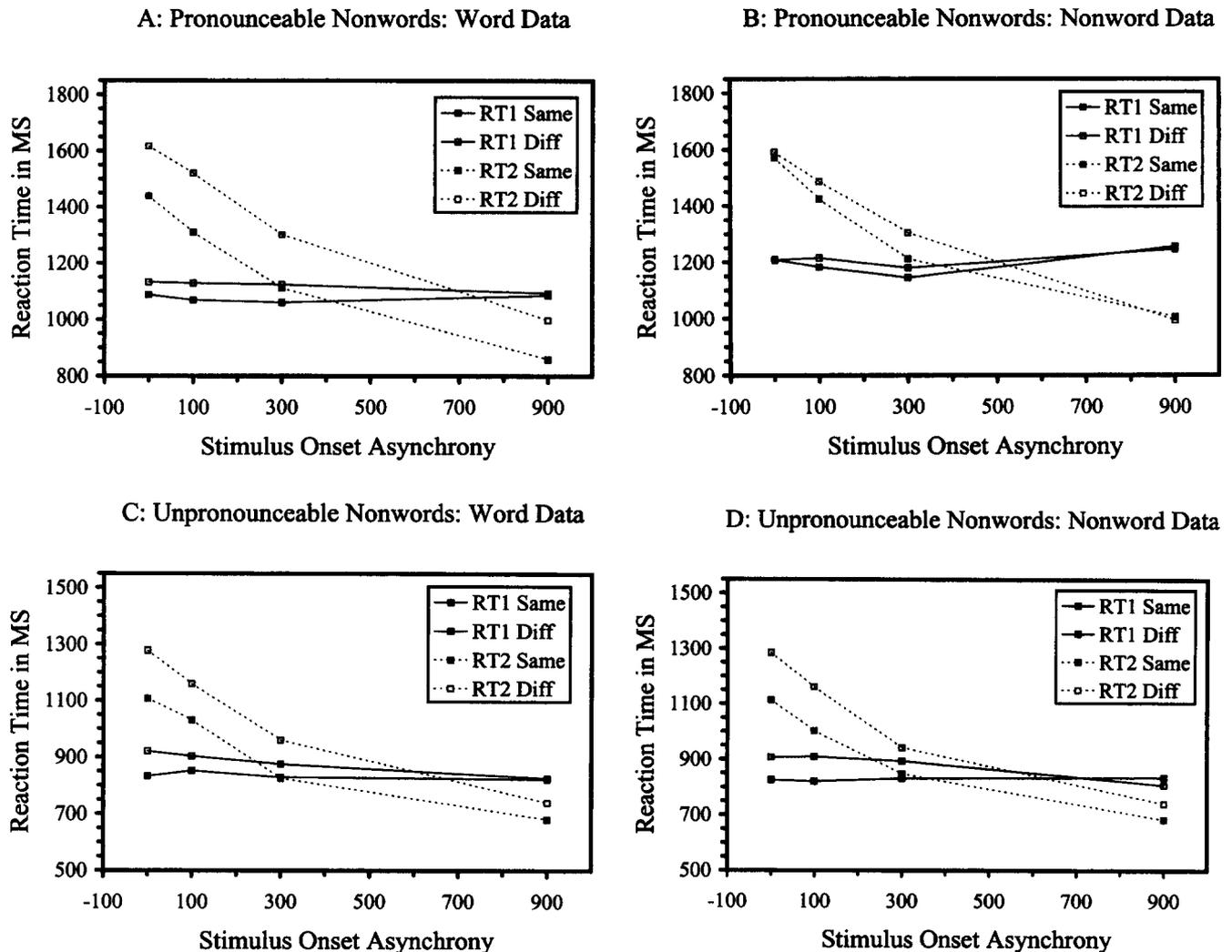


Figure 3. Reaction time on Task 1 (RT1; solid lines) and reaction time on Task 2 (RT2; broken lines) for category-match (filled squares: same) and category-mismatch (open squares: different) trials as a function of stimulus onset asynchrony in Experiment 3. A: Word data for words presented with pronounceable nonwords. B: Data from pronounceable nonwords. C: Word data for words presented with unpronounceable nonwords. D: Data from unpronounceable nonwords. Diff = different.

pronounce. Unpronounceable nonwords were constructed by substituting consonants for vowels. Almost all of the unpronounceable nonwords were strings of consonants. Seven of the strings contained a *y*, and two contained a terminal *e* (i.e., *grfne* for *grade* and *tcfle* for *table*).

As in the previous experiments, there were three (SOA = 0) or four (SOA > 0) displays. The fixation display contained two rows of four dashes separated by spaces (e.g., ----). One row appeared one text line above the location of the top letter string, beginning at row 12, column 35 of the IBM text screen. The other appeared one text line below the location of the bottom word, beginning at row 15, column 35 of the IBM text screen. The fixation display measured 22.5 mm horizontally by 23.5 mm vertically, which corresponds to  $2.14^\circ \times 2.24^\circ$  of visual angle from a viewing distance of 60 cm.

The letter strings were presented in lowercase. Each letter string was 15 mm wide and 4 to 7 mm high, depending on whether it had ascenders (e.g., *t*) or descenders (e.g., *g*) or both (e.g., *j*). This corresponded to visual angles of  $1.43^\circ \times 0.38^\circ$  to  $0.67^\circ$  from a viewing distance of 60 cm. The two letter

strings were written on adjacent lines of the IBM text screen, beginning at row 13, column 35 and row 14, column 35.

The temporal parameters were the same as in the previous experiments: The fixation display was exposed for 500 ms; on zero-SOA trials, S1 and S2 were exposed for 1,000 ms; on greater than zero-SOA trials, S1 was exposed for SOA + 1,000 ms and S2 was exposed for 1,000 ms; on all trials, the blank screen intertrial interval lasted 3,500 ms.

**Procedure.** The experiment involved 20 replications of the basic 16-trial design—that is, S1 (word or nonword)  $\times$  S2 (word or nonword)  $\times$  SOA (0, 100, 300, or 900 ms)—for a total of 320 trials. Words and nonwords were assigned randomly to each trial, and the order in which the trials occurred was randomized separately for each participant. The mapping of stimulus categories onto responses was counterbalanced between participants with 4 participants in each group receiving one of four mappings: WNNW, NWWN, WNWN, or NWNW, where W refers to word, N refers to nonword, and the left-to-right ordering refers to the left middle, left index, right index, and right middle fingers, respectively. Participants

were instructed as in the previous experiments, except that they were told to make lexical decisions about S1 and S2.

**Results**

Mean RT was computed for trials on which both responses were correct in each combination of lexical status, category match, and SOA. The means across participants appear in Figure 3; the accuracy data are presented in Table 4. The RT1 and RT2 data were subjected to separate 2 (nonword pronounceability: pronounceable vs. unpronounceable) × 2 (lexical status: word vs. nonword) × 2 (category match: same vs. different) × 4 (SOA: 0, 100, 300, or 900 ms) ANOVAs. Nonword pronounceability was a between-subjects factor; the others were within-subjects factors. The ANOVA summaries are presented in Table 5.

The RT1 and RT2 data show effects typical of the PRP procedure: RT1 was unaffected by SOA, and RT2 was strongly affected. More relevant to our purposes, there were category-match effects for words and nonwords on both RT1 and RT2. The RT1 effects were smaller than in the previous experiments, averaging 44 ms for words presented with pronounceable nonwords and 47 ms for words presented with unpronounceable nonwords. They averaged 14 ms for pronounceable and 51 ms for unpronounceable nonwords. These results suggest that the lexical status of S2 can be retrieved while participants are assessing the lexical status of S1.

These conclusions are supported by a significant main effect of category match and a significant interaction between category match and SOA in the RT1 ANOVA (see Table 5). Fisher's LSD, computed from the highest order interaction error term, indicated that differences in RT1 larger than 41 ms were significant ( $p < .05$ ). By this criterion, the category-match effect was significant for the 0-, 100-, and 300-ms SOAs but not the 900-ms SOA for words paired with pronounceable nonwords (the differences were 45, 60, 64, and 8 ms, respectively) and for words paired with

**Table 4**  
*Mean Percentage of Correct Responses in Task 1 and Task 2 in Experiment 3 as a Function of Nonword Pronounceability, Lexical Status, Category Match, and Stimulus Onset Asynchrony (SOA)*

Category match	SOA (ms)			
	0	100	300	900
Words with pronounceable nonwords				
Same category	86	90	92	92
Different category	89	88	88	87
Pronounceable nonwords				
Same category	90	88	88	91
Different category	88	89	87	85
Words with unpronounceable nonwords				
Same category	98	99	98	99
Different category	96	96	96	98
Unpronounceable nonwords				
Same category	98	97	98	98
Different category	96	96	97	97

**Table 5**  
*Summaries for Analyses of Variance for RT1, RT2, and Accuracy in Experiment 3*

Source	MSE	df	F
RT1			
Nonword pronounceability (N)	1,156,943.38	1, 30	9.63**
Lexical status (L)	19,372.75	1, 30	17.75**
N × L	19,372.75	1, 30	20.53**
Category match (C)	24,070.28	1, 30	8.14**
N × C	24,070.28	1, 30	0.53
L × C	5,914.62	1, 30	0.93
N × L × C	5,914.62	1, 30	1.54
Stimulus onset asynchrony (S)	26,008.41	3, 90	0.61
N × S	26,008.41	3, 90	1.53
L × S	5,182.39	3, 90	2.20
N × L × S	5,182.39	3, 90	2.76*
C × S	6,200.19	3, 90	4.83**
N × C × S	6,200.19	3, 90	1.34
L × C × S	6,680.69	3, 90	0.27
N × L × C × S	6,680.39	3, 90	0.32
RT2			
N	1,294,267.71	1, 30	10.38**
L	14,086.56	1, 30	6.51*
N × L	14,086.56	1, 30	6.85*
C	39,839.20	1, 30	43.10**
N × C	39,839.20	1, 30	0.11
L × C	14,233.50	1, 30	11.37**
N × L × C	14,233.50	1, 30	10.39**
S	14,267.68	3, 90	502.19**
N × S	14,267.68	3, 90	5.28**
L × S	7,219.13	3, 90	0.42
N × L × S	7,219.13	3, 90	0.11
C × S	6,204.46	3, 90	7.21**
N × C × S	6,204.46	3, 90	2.36
L × C × S	7,412.64	3, 90	0.08
N × L × C × S	7,412.64	3, 90	0.80
Accuracy			
N	158.08	1, 30	60.62**
L	56.40	1, 30	1.19
N × L	56.40	1, 30	0.15
C	53.98	1, 30	7.82**
N × C	53.98	1, 30	0.01
L × C	27.82	1, 30	0.51
N × L × C	27.82	1, 30	0.51
S	38.99	3, 90	0.49
N × S	38.99	3, 90	0.30
L × S	34.64	3, 90	1.16
N × L × S	34.64	3, 90	0.61
C × S	41.06	3, 90	1.00
N × C × S	41.06	3, 90	2.10
L × C × S	40.63	3, 90	1.19
N × L × C × S	40.63	3, 90	0.52

\*  $p < .05$ . \*\*  $p < .01$ .

unpronounceable nonwords (the differences were 88, 51, 47, and 3 ms, respectively). By the LSD criterion, the category-match effect was not significant at any SOA for pronounceable nonwords (the differences were -1, 32, 35, and -10 ms, respectively). It was significant for the 0-, 100-, and 300-ms SOAs for unpronounceable nonwords (the differences were 82, 89, 62, and -28 ms, respectively).

The RT2 data show similar effects, though they were much stronger. The category-match effects averaged 179 and 123 ms for words presented with pronounceable and unpronounceable nonwords, respectively, and 41 and 121 ms for pronounceable and unpronounceable nonwords. These conclusions were supported by a significant main effect of category match and a significant interaction between category match and SOA in the RT2 ANOVA (see Table 5). Using the highest order interaction error term, LSD = 43 ms ( $p < .05$ ). By this criterion, the category-match effect was significant at every SOA for words paired with pronounceable nonwords (the differences were 178, 211, 190, and 138 ms, respectively) and for words paired with unpronounceable nonwords (the differences were 171, 128, 135, and 59 ms, respectively). The category-match effect for pronounceable nonwords was significant only for the 100- and 300-ms SOAs (the differences were 20, 62, 93, and -13 ms, respectively). It was significant at every SOA for unpronounceable nonwords (the differences were 172, 158, 95, and 57 ms, respectively).

The accuracy data were generally consistent with the RTs. The LSD ( $p < .05$ ) from the highest order interaction term was 3.1%.

### Discussion

The category-match effect in the RT1 data supports the parallel retrieval hypothesis (Hommel, 1998) and disconfirms the serial retrieval hypothesis (Carrier & Pashler, 1995). Apparently, participants were able to retrieve information about the lexical status of S2 while they were processing S1. Thus, the conclusions from Experiments 1 and 2 extend to more complex stimuli (five-character strings vs. single characters) and more complex categories (words). Experiment 3 also shows that stimulus repetition is not necessary to produce the category-match effect. Words and nonwords were presented only once in the experiment, yet robust category-match effects were observed.

### Experiment 4

Experiment 3 demonstrated that participants could retrieve semantic information about S2 while processing S1. However, the semantic information was relatively crude: Was S2 a word? The next experiment examined more detailed semantic information, asking whether associations between words could be retrieved while S1 was processed. Participants performed a lexical-decision task on both S1 and S2. On half of the trials on which S1 and S2 were both words, S1 and S2 were semantically associated. On the other half of those trials, S1 and S2 were not associated. The question was whether participants could retrieve information about the relation between S1 and S2 while they were processing S1. If they could, then RT1 should be faster when S1 and S2 were associated than when they were not associated.

In essence, we are asking whether S2 can prime S1 even when S2 follows S1 by an appreciable SOA. Several researchers have found semantic priming from primes presented 100 ms after the target (Briand, den Heyer, & Dannenbring, 1988; Dark, 1988; Kiger & Glass, 1983), so we should be able to find a similar effect, unless processing S1 prevents retrieval of associations to S2. The RT2 data provide a useful manipulation check: RT2 should be faster when S1 and S2 are related because participants will be likely to have processed S1 by the time S2 is presented. Many

researchers have found that the first stimulus primes the second with SOAs in the range we used.

The present experiment also provides an opportunity to replicate the category-match effect observed with lexical-decision tasks in Experiment 3, using a different set of words and nonwords. If participants can retrieve information about the lexical status of S2 while they are processing S1, RT1 should be faster when S1 and S2 are both words or both nonwords than when one is a word and the other is a nonword.

### Method

**Participants.** The participants were 24 volunteers from the general university community who were paid \$5 for serving in one 1-hr session. Seven participants served in one or more of the previous experiments. One served in Experiment 1, 1 served in Experiments 1 and 2, and 5 served in Experiment 2. None served in Experiment 3, which also used a lexical-decision task.

**Apparatus and stimuli.** This experiment used the same apparatus and temporal parameters as the previous experiments, except that the intertrial interval was reduced to 2,500 ms. The fixation display was the same as in Experiment 3, and the spatial parameters were the same, except that letter strings were not restricted in length to five letters.

The stimuli were 512 words and 512 nonwords. The words were selected in pairs from the University of South Florida word association, rhyme, and word fragment norms (Nelson, McEvoy, & Schreiber, 1994), which measure the probability of generating one word as the first associate of another. In our stimuli, the average association strength (probability of generation) was .35, ranging from .10 to .94. The words were selected without regard to part of speech. The average frequency of occurrence in the Kucera and Francis (1967) norms was 282.3 per million, ranging from 20 per million to 9,999 per million.<sup>1</sup>

Nonwords were constructed from the words by substituting vowels. Consonants were substituted in a small number of cases in which all vowel substitutions created either English words or homophones of English words. All nonwords were judged to be pronounceable by the second author. The length of the words and nonwords averaged 5.19 characters, ranging from 2 to 12.

**Procedure.** There were 512 trials. To construct the trials, words and nonwords were arranged in sets of eight: two associated words, two nonwords constructed from them, two more associated words, and two nonwords constructed from the second pair of associated words (e.g., 1 = *awake*, 2 = *asleep*, 3 = *oweke*, 4 = *eslope*, 5 = *false*, 6 = *true*, 7 = *folse*, 8 = *troe*). Each set of eight letter strings was used to construct one of four different types of four trials. Two of the types involved related stimuli, and two involved unrelated stimuli. Two of the types paired words with words and nonwords with nonwords, and two paired words with nonwords and nonwords with words. In the related, word-word, nonword-nonword condition, the first string was paired with the second, the third with the fourth, the fifth with the sixth, and the seventh with the eighth (e.g., *awake* + *asleep*, *oweke* + *eslope*, *false* + *true*, *folse* + *troe*). In the unrelated, word-word, nonword-nonword condition, String 1 was paired with 5, 2 with 6, 3 with 7, and 4 with 8 (e.g., *awake* + *false*, *asleep* + *true*, *oweke* + *folse*, *eslope* + *troe*). In the related word-nonword, nonword-word condition, String 1 was paired with 4, 3 with 2, 5 with 8, and 7 with 6 (e.g., *awake* + *eslope*, *oweke* + *asleep*, *false* + *troe*, *folse* + *true*). Finally, in the unrelated word-nonword, nonword-word condition, String 1 was paired with 7, 2 with 8, 3 with 5, and 4 with 6 (e.g., *awake* + *folse*, *asleep* + *troe*, *oweke* + *false*, *eslope* + *true*). The four trials constructed from each set of eight letter strings were assigned to one of four SOAs (0,

<sup>1</sup> We are extremely grateful to Doug Nelson for selecting words for us that matched our constraints.

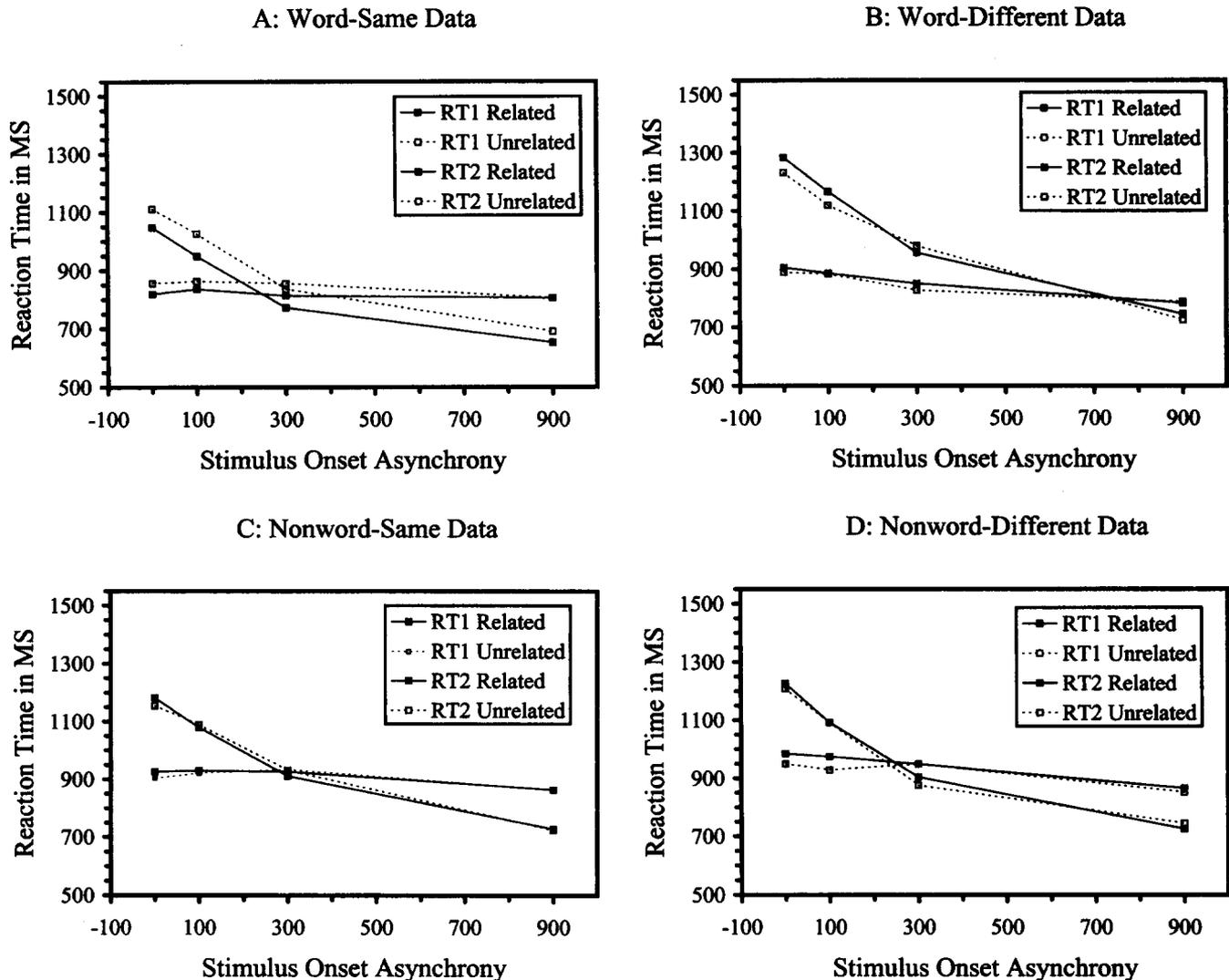


Figure 4. Reaction time on Task 1 (RT1) and reaction time on Task 2 (RT2) for related (solid lines) and unrelated (broken lines) stimuli as a function of stimulus onset asynchrony in Experiment 4. A: Category match; Stimulus 1 (S1) and Stimulus 2 (S2) both words. B: Category mismatch; RT1 data from S1 words and RT2 data from S2 words. C: Category match; S1 and S2 both nonwords. D: Category mismatch; RT1 data from S1 nonwords and RT2 data from S2 nonwords.

100, 300, and 900 ms), and the order of trials and SOAs was randomized separately for each participant.

Participants were instructed as in Experiment 3. They were allowed brief rests after every 64 trials.

### Results

The mean RTs were computed for trials on which both R1 and R2 were correct in each combination of lexical status, relatedness, category match, and SOA. The means across participants are presented in Figure 4. The accuracy data are presented in Table 6. The RT1, RT2, and accuracy data were analyzed in 2 (relatedness: associated vs. not associated)  $\times$  2 (lexical status: word vs. nonword)  $\times$  2 (category match)  $\times$  4 (SOA: 0, 100, 300, or 900 ms)

ANOVAs. The summaries for these ANOVAs are presented in Table 7.

**Priming effects.** The RT data showed effects typical of PRP studies: RT1 was unaffected by SOA, and RT2 was strongly affected. Most important for our purposes, RT1 was affected by associative relations between S1 and S2. In the word-word conditions, RT1 was 27 ms faster when S2 was related to S1 than when it was not related. Thus, participants appear able to retrieve semantic associates of S2 while processing S1. The relation between S1 and S2 also affected RT2. In the word-word conditions, RT2 was 58 ms faster when S1 was associatively related to S2 than when it was not, replicating typical semantic priming effects.

Table 6  
*Mean Percentage of Correct Responses in Task 1 and Task 2 in Experiment 4 as a Function of Lexical Status, Relatedness, Category Match, and Stimulus Onset Asynchrony (SOA)*

Category match	SOA (ms)							
	Words				Nonwords			
	0	100	300	900	0	100	300	900
<b>Related</b>								
Same category	95	98	96	96	92	92	92	94
Different category	90	92	90	92	85	85	91	91
<b>Unrelated</b>								
Same category	92	92	92	94	89	91	92	93
Different category	89	90	90	92	86	89	90	93

These conclusions were supported by inferential statistics. In the RT1 data, the priming effect of S2 on S1 was evaluated in a contrast based on the interaction between relatedness, lexical status, and category match from the RT1 ANOVA (see Table 7) that compared the related, word, same-category (word-word) RTs with the unrelated, word, same-category RTs (i.e., the RT1 data in Figure 4A). The contrast was highly significant,  $F(1, 23) = 53.95$ ,  $p < .01$ ,  $MSE = 2,546.57$ . The priming effect at each SOA was assessed by calculating Fisher's LSD from the highest order interaction. The value of LSD ( $p < .05$ ) was 33 ms. By this criterion, the priming effect was significant at the 0- and 300-ms SOA but not at the 100- or 900-ms SOA (the differences were 37, 27, 42, and 0 ms, for SOAs of 0, 100, 300, and 900 ms, respectively).

In the RT2 data, the contrast comparing related, word, same-category RTs with unrelated, word, same-category RTs was significant,  $F(1, 23) = 221.61$ ,  $p < .01$ ,  $MSE = 2,939.67$ . Fisher's LSD ( $p < .05$ ) from the highest order interaction was 29 ms. By this criterion, the priming effect was significant at the 0-, 100-, and 300-ms SOAs but not at the 900-ms SOA (the differences were 64, 77, 64, and 28 ms, respectively).

**Category-match effects.** The RT data also replicated the category-match effect observed in the previous experiments. Participants responded to S1 20 ms faster if S2 was in the same category than if it was in the opposite category. They responded to S2 73 ms faster if it was in the same category as S1 than if it was in the opposite category.

These conclusions were supported by inferential statistics. The main effect of category match was significant for both RT1 and RT2 (see Table 7). Fisher's LSD ( $p < .05$ ) based on the error term from the highest order interaction was 33 ms for RT1 and 29 ms for RT2. By these criteria, the RT1 effects were significant at the 0-, 100-, and 300-ms SOAs for related words, at the 0-ms SOA for unrelated words, at the 0- and 100-ms SOAs for related nonwords, and at the 0-ms SOA for unrelated nonwords. The RT2 effects were significant at all SOAs for related and unrelated words and at the 0-ms SOA for related and unrelated nonwords.

**Accuracy.** The accuracy data, presented in Table 6, are consistent with the RTs. The LSD ( $p < .05$ ) for the highest order interaction was 4%.

### Discussion

The category-match effect, observed in the previous experiments, was replicated in this experiment as well, using a new set

of materials. The category-match effect suggests that participants were able to retrieve information about the lexical status of S2 while they were processing S1. As in Experiment 3, specific words and nonwords were presented only once in the experiment, confirming the conclusion that stimulus repetition is not necessary to produce parallel retrieval.

Experiment 4 goes beyond the previous ones in showing that S2 can produce semantic priming while participants are busy processing S1. Participants must have been able to retrieve associations between S2 and S1 in parallel with processing S1. This is an important extension of the previous experiments, because it shows that very specific information about S2 can be retrieved in parallel. The category-match effect in this experiment and in Experiment 3 showed that S2 could activate lexical memory in a general way while S1 was being processed, and Experiments 1 and 2 showed that S2 could activate other general aspects of semantic memory. It is important to show that S2 produces specific rather than general activation, because the memory theories at issue assume that retrieval cues activate specific traces.

### General Discussion

The experiments were intended to distinguish between parallel and serial retrieval in dual-task situations. If parallel retrieval was possible, then things retrieved from S2 should affect S1 processing. If parallel retrieval was not possible, then nothing should be retrieved from S2 until S1 processing was complete. The results of each experiment supported the parallel retrieval hypothesis. Experiment 1 showed that participants could retrieve information that indicated whether S2 was a letter or a digit while processing S1. Experiment 2 showed that participants could retrieve information about the magnitude and parity of S2 while processing S1, with the important caveat that the task set had to be the same for S1 and S2. Changing task sets from S1 to S2 apparently blocked retrieval. Experiments 3 and 4 showed that participants could retrieve information about the lexical status of S2 while processing S1, and Experiment 4 showed that S2 could semantically prime S1 while S1 was being processed.

### Caveats

**Slow RT1.** One possible limitation on the generality of these conclusions stems from the fact that RT1 was so long in each

**Table 7**  
Summaries for Analyses of Variance for RT1, RT2, and Accuracy in Experiment 4

Source	MSE	df	F
RT1			
Relatedness (R)	3,223.37	1, 23	0.61
Lexical status (L)	30,955.02	1, 23	37.98**
R × L	3,340.89	1, 23	7.89*
Category match (C)	10,162.18	1, 23	8.15**
R × C	4,335.96	1, 23	8.03**
L × C	2,191.60	1, 23	0.15
R × L × C	2,546.57	1, 23	1.55
Stimulus onset asynchrony (S)	25,964.01	3, 69	9.53**
R × S	3,395.45	3, 69	0.76
L × S	4,523.60	3, 69	2.92*
R × L × S	3,360.06	3, 69	1.24
C × S	4,497.87	3, 69	8.96**
R × C × S	3,006.85	3, 69	0.89
L × C × S	2,976.87	3, 69	0.70
R × L × C × S	3,303.69	3, 69	1.20
RT2			
R	3,236.38	1, 23	1.65
L	7,539.58	1, 23	16.26**
R × L	2,897.96	1, 23	4.93*
C	31,849.85	1, 23	32.32**
R × C	3,162.38	1, 23	33.92**
L × C	12,179.09	1, 23	67.18**
R × L × C	2,939.67	1, 23	27.02**
S	19,377.84	3, 69	417.32**
R × S	2,841.98	3, 69	1.57
L × S	4,981.28	3, 69	0.09
R × L × S	2,873.30	3, 69	1.33
C × S	4,885.90	3, 69	10.00**
R × C × S	2,960.41	3, 69	2.42
L × C × S	5,063.11	3, 69	6.34**
R × L × C × S	2,501.21	3, 69	2.78*
Accuracy			
R	47.37	1, 23	5.01*
L	96.40	1, 23	10.25**
R × L	35.27	1, 23	7.18*
C	71.42	1, 23	30.29**
R × C	29.85	1, 23	10.52**
L × C	81.57	1, 23	0.11
R × L × C	29.57	1, 23	0.21
S	71.01	3, 69	4.46**
R × S	44.88	3, 69	0.33
L × S	54.80	3, 69	2.35
R × L × S	48.49	3, 69	0.65
C × S	47.22	3, 69	1.54
R × C × S	52.75	3, 69	0.48
L × C × S	66.04	3, 69	0.66
R × L × C × S	54.66	3, 69	0.59

\*  $p < .05$ . \*\*  $p < .01$ .

experiment. It was longer than what would be expected from single-task controls, and it was longer than RT2 for the same task at the 900-ms SOA. It is possible that participants strategically delayed R1 in order to gain the benefits of category repetition. We think such strategic slowing is unlikely for two reasons. First, there is no evidence of strategic slowing in the data. If participants waited for S2 before responding with R1, then RT1 should in-

crease with increasing SOA. There was no effect of SOA on RT1 in any experiment, contrary to this prediction. Second, S1 and S2 categories mismatched as often as they matched, so any benefits gained from repetitions would be offset by costs from nonrepetitions. The interactions between S1 and S2 produced no net benefit that would reward strategic slowing. Note as well that RT1 slowing was greatest in the different-task-set conditions of Experiment 2, where no category-match effects were observed. Participants could not have waited for S2 to benefit from nonexistent effects. Thus, we conclude that participants did not wait for S2, so strategic slowing cannot account for the category-match effects.

In our view, RT1 slowing is a common occurrence in PRP experiments. There is an implicit assumption in most PRP theories that RT1 is as fast as single-task controls, based on the idea that Task 1 gets first access to the bottleneck in single-task and dual-task conditions. The invariance of RT1 over SOA suggests that Task 1 got first access to the bottleneck, so single-task controls for Task 1 are relatively rare in the literature. Whenever single-task controls have been run, however, RTs are faster than dual-task RT1s (Bertelson, 1967; Gottsdanker, Broadbent, & Van Sant, 1963; Herman & McCauley, 1969; Hommel, 1998). The difference between dual-task RT1 and single-task controls warrants explanation, of course (see, e.g., Herman & Kantowitz, 1970), but it does not appear to compromise our conclusions.

*Locus of category-match effects.* Throughout this article, we have assumed that the category-match effect reflected memory retrieval; participants retrieved the category of S2 while they were retrieving the category of S1. However, other interpretations are possible, and it is important to rule them out. One possibility is that the category-match effects reflect stimulus encoding. For example, the stimuli may have been more similar when the categories matched than when they mismatched, and similarity may have facilitated processing of both S1 and S2, resulting in a category-match effect. The encoding-stage interpretation is an important alternative to rule out, because several PRP studies suggest that S2 can be encoded while S1 is being processed (e.g., DeJong, 1993; Pashler, 1984; Pashler & Johnston, 1989). Thus, the category-match effect would be expected on the basis of those prior results. However, stimulus encoding seems unlikely to be the locus of the category-match effect, because the stimuli in the different categories in each experiment were quite similar. The letters in Experiment 1 were selected to be similar to the digits. The words and nonwords in Experiments 3 and 4 were all formed from the same letters, and it is likely that the words and nonwords sampled the letters with the same frequency, at least when the nonwords were pronounceable. Experiment 2 provides the strongest evidence against a stimulus encoding interpretation, because the two tasks (magnitude and parity judgments) used exactly the same stimuli, so similarity of S1 to S2 should have been the same whether or not Task 1 and Task 2 used the same task set. However, the category-match effect was strongly modulated by task set. It was strong when the task sets were the same, and it was almost eliminated when the task sets were different.

Note that repetition of the exact stimulus is not necessary to produce the category-match effect. There were no stimulus repetitions in Experiments 1, 3, and 4, yet there were category-match effects in each case. Exact stimulus repetitions occurred on occasion in Experiment 2, but the effects were quite small, compared with the advantage of category matches over category mismatches.

In the magnitude–magnitude condition, RT1 was 730 ms for stimulus repetitions, 750 ms for response repetitions (category matches), and 852 ms for nonrepetitions (category mismatches; LSD = 68 ms,  $p < .05$ ). In the parity–parity condition, RT1 was 796 ms for stimulus repetitions, 840 ms for response repetitions, and 958 ms for nonrepetitions (LSD = 74 ms,  $p < .05$ ).

Another possible locus for the category-match effect is in post-bottleneck response processes. Response processes could benefit from repetition of motor commands or from residual activation from prior responses that were similar. This seems unlikely as well, because Task 1 and Task 2 used different responses. In each experiment, participants responded to Task 1 with their right hands and to Task 2 with their left hands. Moreover, in all experiments, the mapping of categories onto responses was counterbalanced, so that the same category was not always mapped onto homologous fingers (e.g., in Experiment 1, some participants used their right index and left middle fingers to indicate “digit”). Thus, it is likely that the category-match effect occurred at a more central locus (also see Hommel, 1998).

*Variation in effect size.* The category-match effects were larger in Experiments 1 and 2 ( $M_s = 126$  and  $190$  ms for RT1 and RT2, respectively) than in Experiments 3 and 4 ( $M_s = 30$  and  $95$  ms for RT1 and RT2, respectively). Several factors could account for this difference. First, the stimuli in Experiments 1 and 2 were simpler than the stimuli in Experiments 3 and 4, single letters and digits versus multiletter words and nonwords. A simple stimulus may activate memory more readily than a complex one. Second, the categories in Experiments 1 and 2 were simpler and smaller than the ones in Experiments 3 and 4. It may be easier to detect activation of one of 10 digits or one of 26 letters than of one of several hundred thousand words. Third, the stimuli were repeated many times in Experiments 1 and 2 but were presented only once in Experiments 3 and 4. There was more opportunity for learning and repetition effects to benefit retrieval in Experiments 1 and 2. Further research will be necessary to distinguish among these alternatives.

*Automaticity.* One could argue that we obtained category-match effects because the stimuli could be categorized automatically. Jacoby (1991), among others, argued that automatic memory retrieval can occur without interference from a concurrent task. It is difficult to evaluate Jacoby’s evidence for that claim because his concurrent tasks were not synchronized and because the claim depends on the validity of Jacoby’s process dissociation procedure, which has been challenged many times since he first proposed it (e.g., Curran & Hintzman, 1995; Dodson & Johnson, 1996; Hillstrom & Logan, 1997). Nevertheless, the claim that automatic retrieval can occur in parallel has some currency in the literature and should be considered despite the controversy over the data that appear to support it.

There is reason to believe that each of our tasks could be performed automatically. The automaticity of letter–digit discrimination (Experiment 1) is supported by visual search studies. Searching for letters among digits or digits among letters often produces “pop out” that is characteristic of automatic processing (i.e., search time is unaffected by the number of distractors; see Egeth, Jonides, & Wall, 1972; Jonides & Gleitman, 1972; Schneider & Shiffrin, 1977; but see Duncan, 1983; Krueger, 1984). The automaticity of magnitude judgments (Experiment 2) is supported by Stroop-type interference from magnitude when judging parity

(Otten, Sudevan, Logan, & Coles, 1996; Tzelgov, Meyer, & Henik, 1992). The automaticity of lexical access (Experiments 3 and 4) is supported by many Stroop studies (e.g., Stroop, 1935) and priming studies (e.g., Neely, 1977).

There is nothing in our data to rule out the possibility that the category-match effect is restricted to stimuli that can be processed automatically or to tasks that can be performed automatically. The data from Experiment 2 suggest that category-match effects are conditionally automatic if they are automatic at all (Bargh, 1992; DeJong et al., 1994). They occur if the task set is the same for S1 and S2, but not if it is different. Stroop (Bauer & Besner, 1997; Besner, Stolz, & Butilier, 1997) and priming effects (Chiappe, Smith, & Besner, 1996; Henik, Friedrich, & Kellogg, 1983; McKoon & Ratcliff, 1995; Smith, 1979; but see Stolz & Besner, 1996) show a similar conditionality. Smith, Theodor, and Franklin (1983), for example, showed that one word semantically primed another only if it was treated as a word. Performing letter search on the prime blocked priming.

*Generalization to episodic memory.* The present experiments examined retrieval from semantic memory. Carrier and Pashler (1995) examined retrieval from episodic memory. To what extent do our conclusions bear on theirs? The answer depends on the reader’s faith in the similarity between episodic and semantic memory. Readers who think they are similar will be comfortable with the generalization. Readers who think they are dissimilar will want to see our results replicated in an episodic memory task. There is no reason in principle why the crosstalk analysis cannot be applied to episodic memory if accuracy can be controlled in a manner that makes the RTs interpretable. The present results take one step toward demonstrating parallel retrieval in episodic memory. They show that the crosstalk analysis can discriminate parallel retrieval from serial retrieval in semantic memory. Thus, there is no practical reason why the method cannot be extended to episodic memory.

### *Implications for Theories of the PRP*

*Problems.* The present experiments have important implications for the two dominant theories of the PRP: the *response selection bottleneck* view (Pashler, 1989; Pashler & Johnston, 1989) and the *strategic response deferral* model (Meyer & Keiras, 1997). The category-match effect is problematic for the response selection bottleneck view because it suggests that response selection (i.e., category retrieval) can go on in parallel in Task 1 and Task 2. The response selection bottleneck view claims that response selection is strictly serial (Pashler, 1989). In fact, the category-match effect directly contradicts data obtained by Pashler and Johnston (1989). They examined S2 repetition effects in a PRP task with tone discrimination as Task 1 and letter discrimination as Task 2. Repetition effects were additive with SOA, which suggests by the locus of slack logic that repetition affected bottleneck or postbottleneck processes. By contrast, our data suggest that response repetition affects stages that go on in parallel between Task 1 and Task 2, which must be before the bottleneck if there is one. Below we consider ways to reconcile these results.

The category-match effect is not as problematic for the strategic response deferral model, because it assumes that response selection can go on in parallel in Task 1 and Task 2—there are no central capacity limitations (Meyer & Keiras, 1997). The category-

match effect suggests that information may cascade continuously through the stages underlying performance (Eriksen & Schultz, 1979; McClelland, 1979), which contradicts the strategic response deferral model's assumption of discrete, serial stages (Sternberg, 1969). That assumption appears to have been motivated primarily by convenience. It is less central to the strategic response deferral model than the assumption of no central capacity limitations.

In addition to the category-match effect, two other effects challenge the dominant views of the PRP: the finding that RT1 was generally slow (slower than RT2 at the longest SOA) and the finding of large task-set switching costs in RT1 and RT2 in Experiment 2. Neither of these effects is predicted by the response selection bottleneck view (Pashler, 1989; Pashler & Johnston, 1989) or by the strategic response deferral model (Meyer & Keiras, 1997). Both perspectives assume that Task 1 gets first access to the bottleneck, as it would in single-task conditions, so there should be no RT1 slowing in dual-task conditions. The response selection bottleneck view is a rather stark stage model and has no processes to account for RT1 slowing. The strategic response deferral model has many processes at its disposal, but none of them account for RT1 slowing. The most fundamental principle of the model is that there are no central capacity limitations. Improved versions of these models may be able to explain RT1 slowing in terms of limitations on maintaining two task sets in working memory at the same time (Logan, 1979), in terms of limitations on preparation (DeJong, 1995; DeJong & Sweet, 1994), or in terms of competition among the responses that are active on a trial (R1 only on single-task trials; R1 and R2 on dual-task trials; Herman & Kantowitz, 1970).

The set switching costs in Experiment 2 are reminiscent of set switching costs in procedures in which participants alternate between tasks (Allport et al., 1994; Rogers & Monsell, 1995). It may be profitable to think of them in the same way. The response selection bottleneck view provides no mechanisms to account for set switching, so the present set switching costs challenge the sufficiency of its explanation of PRP phenomena. The strategic response deferral model contains mechanisms for switching sets, but under the assumption that central processes have no capacity limitations (Meyer & Keiras, 1997), it may not account for set switching costs. It seems likely that interpretations of set switching that are not based on capacity limitations could be implemented in the strategic response deferral model, however (e.g., the *task set inertia* hypothesis of Allport et al., 1994).

*One possible reconciliation: Conditional and unconditional automaticity.* The automaticity literature provides one way to reconcile the present findings and those of Hommel (1998) with those of Carrier and Pashler (1995). More generally, it provides a way to reconcile evidence of parallel retrieval seen in experiments that employ the crosstalk logic with the evidence of serial retrieval seen in experiments that employ the locus of slack logic. Bargh (1992) distinguished between *conditional* and *unconditional* automaticity (see also DeJong et al., 1994). A process is unconditionally automatic if it occurs whenever a relevant stimulus is presented, regardless of the person's intentions and regardless of the current direction of attention. Unconditioned reflexes and processes like dark adaptation are good examples of unconditionally automatic processes. By contrast, a process is conditionally automatic if it occurs whenever a relevant stimulus occurs and the person is in a

particular intentional or attentional state. The category-match effects in the present experiments and the Stroop (Bauer & Besner, 1997; Besner et al., 1997) and priming effects (Chiappe et al., 1996; Henik et al., 1983; McKoon & Ratcliff, 1995; Smith, 1979; Smith et al., 1983; Stolz & Besner, 1996) in previous experiments seem to be examples of this kind of automaticity.

The locus of slack logic could be sensitive to unconditionally automatic processes, whereas the crosstalk logic could be sensitive to conditionally automatic processes. The locus of slack logic seems to require a change of task set from Task 1 to Task 2 in order to provide a neutral baseline against which to evaluate Task 2 difficulty effects. The only Task 2 processes that can go on in parallel with a different task set in Task 1 may be those processes that are not contingent on task set, namely, the unconditionally automatic processes. S2 intensity is a good example of a difficulty factor that is likely to be unconditionally automatic. Any task would benefit from a "cleaned up" stimulus. Indeed, several studies have shown that S2 intensity effects interact underadditively with SOA, suggesting that the process it affects goes on concurrently with Task 1 automatic (DeJong, 1993; Pashler, 1984; Pashler & Johnston, 1989).

The crosstalk logic shows evidence of parallel processing when the task set is the same from Task 1 to Task 2. Under these conditions, the task set necessary for performing Task 2 is already in place when participants perform Task 1. That fulfills one of the conditions necessary for a conditionally automatic process to be triggered. The presentation of a task-relevant stimulus—either S1 or S2—fulfills the other condition, and the process will be triggered. If the task set is different for Task 1 and Task 2, then the task set for Task 1 does not fulfill the conditions necessary for conditionally automatic Task 2 processes to fire, and there will be no crosstalk.

The instance theory of automaticity (Logan, 1988; Logan & Etherton, 1994) assumes that acquired automaticity is conditional in this manner. The current state of attention determines what is encoded into memory, and it determines what is retrieved from memory. The contents of the focus of attention act as a retrieval cue that brings relevant associations to mind (Bororat & Logan, 1997; see also Moscovitch, 1994).

*Another possibility: Parallel sampling and serial recovery.* The memory literature provides another way to reconcile the present findings and those of Hommel (1998) with those of Carrier and Pashler (1995) and to reconcile parallel retrieval in crosstalk experiments with serial retrieval in locus of slack experiments. Formal models of memory generally distinguish two steps in retrieval, one that samples information from memory, given the retrieval cue as input, and one that attempts to recover a possible response from the sampled information (Gillund & Shiffrin, 1984; Humphreys et al., 1989; Murdock, 1982, 1983, 1993; Shiffrin & Steyvers, 1997). Following Raaijmakers and Shiffrin (1981), we call these steps *sampling* and *recovery*, respectively. Following Hintzman (1988), we call the output of the sampling process (that serves as input to the recovery process) the *echo*. In all of the models, the sampling process is assumed to be parallel, in the sense that all of the traces in memory contribute, in principle, to the echo. In all of the models, the recovery process is serial, in the sense that it produces exactly one output. In recall tasks, it may be a word. In recognition tasks, it is a decision about whether the memory probe was on the study list.

The same two steps most likely occur in the kind of semantic retrieval we studied in our tasks. In each case, the stimulus probes memory for relevant information, and some process has to come up with a discrete response. It may be that crosstalk experiments address the sampling process, whereas locus of slack experiments address the recovery process. S1 and S2 may initiate separate sampling processes at the same time (or SOA ms apart), and the parallel nature of sampling may result in crosstalk, if Task 1 and Task 2 are related. The echo produced by S2 may mix with the echo produced by S1 and consequently affect RT1. The recovery process is necessary to choose R1, and it chooses one and only one response. The act of recovering R1 from the echo may somehow reset the sampling process, so that partial results accumulated while waiting for R1 will be lost. Alternatively, execution of R1 may produce a large amount of self-inhibition to prevent the motor system from perseverating on R1 (Dell, Burger, & Svec, 1997). That self-inhibition may destroy partial echos of S2 that built up before R1.

### Implications for the Memory Literature

The evidence for parallel retrieval in the present experiments and in Hommel's (1998) experiments has implications for the long-standing literature on the attention demands of retrieval (Baddeley et al., 1984; Craik et al., 1996; Jacoby, 1991; Johnston et al., 1970, 1972; Martin, 1970; Moscovitch, 1994; Park et al., 1989; Trumbo & Milone, 1971). It suggests that parallel sampling is possible. Carrier and Pashler's (1995) results suggest that recovery may be strictly serial (but see Rohrer et al., 1998; Ross & Anderson, 1981). It should be possible to design PRP experiments using methodology from the memory literature to test these hypotheses precisely.

### Conclusions

The present experiments suggest that parallel memory retrieval is possible in dual-task situations. They go beyond previous experiments in exerting tight temporal control over the concurrent tasks, and they go beyond previous PRP experiments by examining situations in which the task set stays the same for S1 and S2. Parallel retrieval appears to require that the same task set be applied to both tasks.

### References

- Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance* (Vol. 15, pp. 420–452). Cambridge, MA: MIT Press.
- Anderson, J. R., & Ross, B. H. (1980). Evidence against a semantic-episodic distinction. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 441–466.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.
- Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General*, 113, 518–540.
- Bargh, J. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing. *American Journal of Psychology*, 105, 181–199.
- Bauer, B., & Besner, D. (1997). Processing in the Stroop task: Mental set as a determinant of performance. *Canadian Journal of Experimental Psychology*, 51, 61–68.
- Bertelson, P. (1966). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology*, 18, 153–163.
- Bertelson, P. (1967). The refractory period of choice reactions with regular and irregular interstimulus intervals. *Acta Psychologica*, 27, 45–56.
- Besner, D., Stolz, J. A., & Boutilier, C. (1997). The Stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, 4, 221–225.
- Boronat, C. B., & Logan, G. D. (1997). The role of attention in automatization: Does attention operate at encoding, or retrieval, or both? *Memory & Cognition*, 25, 36–46.
- Briand, K., den Heyer, K., & Dannenbring, G. L. (1988). Retroactive semantic priming in a lexical decision task. *Quarterly Journal of Experimental Psychology*, 40A, 341–359.
- Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1339–1348.
- Chiappe, P. R., Smith, M. C., & Besner, D. (1996). Semantic priming in visual word recognition: Activation blocking and domains of processing. *Psychonomic Bulletin & Review*, 3, 249–253.
- 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.
- Curran, T., & Hintzman, D. L. (1995). Violations of the independence assumption in process dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 531–547.
- Dark, V. J. (1988). Semantic priming, prime reportability, and retroactive priming are interdependent. *Memory & Cognition*, 16, 299–308.
- DeJong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 965–980.
- DeJong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology*, 48A, 2–25.
- DeJong, R., Liang, C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 731–750.
- DeJong, R., & Sweet, J. B. (1994). Preparatory strategies in overlapping-task performance. *Perception & Psychophysics*, 55, 142–151.
- Dell, G. S., Burger, L. S., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, 104, 123–147.
- Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General*, 125, 181–194.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, 87, 272–300.
- Duncan, J. (1983). Category effects in visual search: A failure to replicate the "oh-zero" effect. *Perception & Psychophysics*, 34, 221–232.
- Egeth, H., & Dagenbach, D. (1991). Parallel vs. serial processing in visual search: Further evidence from subadditive effects of visual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 551–560.
- Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multiple displays. *Cognitive Psychology*, 3, 674–698.
- Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics*, 25, 249–263.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1–67.

- Gottsdanker, R., Broadbent, L., & Van Sant, C. (1963). Reaction time to single and to first signals. *Journal of Experimental Psychology*, *66*, 163–167.
- Henik, A., Friedrich, F. J., & Kellogg, W. A. (1983). The dependence of semantic relatedness effects on prime processing. *Memory & Cognition*, *11*, 366–373.
- Herman, L. M., & Kantowitz, B. H. (1970). The psychological refractory period effect: Only half the double-stimulation story? *Psychological Bulletin*, *73*, 74–88.
- Herman, L. M., & McCauley, M. E. (1969). Delay in responding to the first stimulus in the “psychological refractory period” experiment: Comparisons with delay produced by a second stimulus not requiring a response. *Journal of Experimental Psychology*, *81*, 344–350.
- Hillstrom, A. P., & Logan, G. D. (1997). Process dissociation, cognitive architecture, and response time: Comments on Lindsay and Jacoby (1994). *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1561–1578.
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace model. *Psychological Review*, *93*, 411–428.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528–551.
- Hommel, B. (1998). Automatic stimulus–response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1368–1384.
- Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, *96*, 208–233.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Johnston, W. A., Greenberg, S. N., Fisher, R. P., & Martin, D. W. (1970). Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology*, *83*, 164–171.
- Johnston, W. A., Wagstaff, R. R., & Griffith, D. (1972). Information-processing analysis of verbal learning. *Journal of Experimental Psychology*, *96*, 307–314.
- Jonides, J., & Gleitman, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, *12*, 457–460.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kiger, J. I., & Glass, A. L. (1983). The facilitation of lexical decision by a prime occurring after the target. *Memory & Cognition*, *11*, 356–365.
- Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics*, *35*, 558–564.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Landauer, T. K. (1975). Memory without organization: Properties of a model with random storage and undirected recall. *Cognitive Psychology*, *7*, 495–531.
- Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 189–207.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Logan, G. D., & Etherton, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1022–1050.
- Martin, D. W. (1970). Residual processing capacity during verbal organization in memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 391–397.
- McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.
- McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *18*, 463–480.
- McKoon, G., & Ratcliff, R. (1986). Automatic activation of episodic information in a semantic memory task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 108–115.
- McKoon, G., & Ratcliff, R. (1995). Conceptual combinations and relational contexts in free association and in priming in lexical decision and naming. *Psychonomic Bulletin & Review*, *2*, 527–533.
- McKoon, G., Ratcliff, R., & Dell, G. S. (1986). A critical evaluation of the semantic–episodic distinction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 295–306.
- Meyer, D. E., & Keiras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review*, *104*, 3–65.
- Moray, N., Fitter, M., Ostry, D., Favreau, D., & Nagy, V. (1976). Attention to pure tones. *Quarterly Journal of Experimental Psychology*, *28*, 271–283.
- 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.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609–626.
- Murdock, B. B. (1983). A distributed memory model for serial-order information. *Psychological Review*, *90*, 316–338.
- Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, *100*, 183–203.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1994). *The University of South Florida word association, rhyme, and word fragment norms*. Unpublished manuscript.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.
- Otten, L. J., Sudevan, P., Logan, G. D., & Coles, M. G. H. (1996). Magnitude versus parity in numerical judgments: Event-related potentials implicate response conflict as the source of interference. *Acta Psychologica*, *94*, 21–40.
- 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.
- Pashler, H. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 358–377.
- Pashler, H. (1989). Dissociations and contingencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology*, *21*, 469–514.
- Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *16*, 220–244.
- Pashler, H., & Johnston, J. C. (1989). Interference between temporally overlapping tasks: Chronometric evidence for central postponement with or without response grouping. *Quarterly Journal of Experimental Psychology*, *41A*, 19–45.
- Pohlmann, L. D., & Sorkin, R. D. (1976). Simultaneous three-channel signal detection: Performance and criterion as a function of order of report. *Perception & Psychophysics*, *20*, 179–186.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 552–572.

- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 207-231.
- Rohrer, D., Pashler, H., & Etchegaray, J. (1998). When two memories can and cannot be retrieved concurrently. *Memory & Cognition*, *26*, 731-739.
- Ross, B. H., & Anderson, J. R. (1981). A test of parallel versus serial processing applied to memory retrieval. *Journal of Mathematical Psychology*, *24*, 183-223.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1-66.
- Schweickert, R. (1978). A critical path generalization of the additive factor method. *Journal of Mathematical Psychology*, *18*, 105-139.
- Schweickert, R., & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology*, *33*, 328-347.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145-166.
- Shulman, H. G., & Fisher, R. P. (1972). Expected value as a determinant of the distribution of attention. *Journal of Experimental Psychology*, *93*, 343-348.
- Smith, M. C. (1967). Theories of the psychological refractory period. *Psychological Bulletin*, *67*, 202-213.
- Smith, M. C. (1979). Contextual facilitation in a letter search task depends on how the prime is processed. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 239-251.
- Smith, M. C., Theodor, L., & Franklin, P. (1983). The relationship between contextual facilitation and depth of processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 697-712.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance* (Vol. 2, pp. 276-315). Amsterdam: Elsevier/North-Holland.
- Stolz, J. A., & Besner, D. (1996). Role of set in visual word recognition: Activation and activation blocking as a nonautomatic process. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1166-1177.
- Strayer, D. L., & Kramer, A. F. (1990). Attentional requirements of automatic and controlled processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 67-82.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.
- Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 89-103.
- Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*, 46-54.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, England: Cambridge University Press.
- Townsend, J. T., & Schweickert, R. (1989). Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. *Journal of Mathematical Psychology*, *33*, 309-327.
- Trumbo, D., & Milone, F. (1971). Primary task performance as a function of encoding, retention, and recall in a secondary task. *Journal of Experimental Psychology*, *91*, 273-279.
- Tulving, E. (1983). *Elements of episodic memory*. New York: Oxford University Press.
- Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 166-179.
- Welford, A. T. (1952). The 'psychological refractory period' and the timing of high-speed performance. *British Journal of Psychology*, *43*, 2-19.

Received September 8, 1998

Revision received March 8, 1999

Accepted June 23, 1999 ■