Primining in Recognition Memory for Categorized Lists

Stephan Lewandowsky
University of Toronto, Toronto, Ontario, Canada

Previous research on primed recognition of categorized lists has shown two discrepant patterns of results, for which there has not been a conclusive explanation. Sometimes rejection of negative recognition test items (lures) was inhibited as a consequence of semantic priming, whereas at other times processing of negative test items was facilitated. This article investigates the reasons for that discrepancy by focusing on the differences between the tasks used to effect priming in the various previous studies. The first two experiments in this paper showed that lure processing is facilitated when priming is achieved through another recognition test item, whereas inhibition is obtained if a semantic category judgment task is performed on the priming items. Thus, both patterns can be reproduced under nearly identical circumstances, with the type of prime processing being the only difference. Two additional experiments are reported that served to generalize the inhibition found in the second experiment to other semantic priming tasks. The data show that the type of processing done on the prime determines whether inhibition or facilitation of lure rejection is obtained. Inhibition is obtained when a semantic task is used to prime a recognition judgment, whereas facilitation results from priming with an episodic task. The results are interpreted in the framework of the semantic/episodic distinction.

According to some models of information processing (e.g., Collins & Loftus, 1975; Anderson 1976, 1983a, 1983b), words are represented in memory through individual nodes that are linked together in a network. Nodes are grouped according to their semantic similarity. One phenomenon consistent with these models is semantic priming. Priming refers to the fact that an item (the prime) facilitates subsequent responding to a related word (the critical item). Since the work of Meyer and Schvaneveldt (1971), the lexical-decision task has been a popular way of investigating priming. The typical finding is that when subjects must decide whether a letter string constitutes a word, they respond to, say, nurse faster if it was preceded by doctor than if it was preceded by bread. Within the framework of network models, the prime increases the level of activation of the corresponding node. This activation then spreads to neighboring (related) nodes, thereby readying them for subsequent access.¹

More recently, priming has also been studied with paradigms involving episodic memory, that is, situations in which subjects' memory for specific occurrences of words is tested. Typically, the type of memory test used is item recognition, because it lends itself to an easy measurement of reaction time (RT). The response to a critical recognition test item is observed as a function of some prior stimulus (the prime) that has a semantic relation to the critical item. The semantic relation is usually defined by a joint category membership. Several recent priming studies in episodic memory have yielded conflicting results. On the one hand, Macht and O'Brien (1980) found that priming inhibits rejection of negative recognition test items (lures). On the other hand, some studies have found that rejection of lures is facilitated through semantic priming (e.g., Taylor & Juola, 1974).

Macht and O'Brien (1980) used features of the Collins and Loftus (1975) model and of a familiarity-based view of recognition memory (e.g., Atkinson & Juola, 1973, 1974) to derive predictions for a priming study. According to the latter type of model, items have a strength or familiarity value associated with them. This value is raised when the word is studied. On the average, studied items thus have a higher familiarity than nonstudied items. It is assumed that at test, if the strength of an item falls above an upper criterion, a yes response is made, whereas if the strength falls below a lower cutoff, a no response is made. If the familiarity of an item falls between the two decision criteria, additional list processing is conducted.

Given such a familiarity model, plus the additional assumption that the items underlying the strength distributions are themselves part of the overall network (cf. Atkinson, Herrmann, & Wescourt, 1974), Macht and O'Brien derived the following predictions for the effects of a prime. A prime raises the

¹ Ratcliff and McKoon (1981b) demonstrated that the onset of activation takes virtually no time to traverse a link, therefore ruling out the existence of a rather slow (50 ms per link, Anderson, 1976) "spread" of activation. Ratcliff and McKoon also showed, however, that the buildup of facilitation, which presumably reflects the increase in level of activation, proceeds at different rates and reaches different asymptotes depending on semantic distance. As shown by Anderson (1983a, 1983b) and Lorch (1982), the concept of spreading activation is therefore still tenable, bearing in mind that the rate and the absolute magnitude of the build-up are crucial, but not time of onset.
level of activation in the network and thereby also increases the familiarity of related items. The increase in familiarity is tantamount to moving both the old and the new item distribution of a given category upward on the strength continuum. More of the old item distribution should then fall above the upper criterion, thereby increasing the average speed of responding to a target probe. Similarly, fewer of the related new items should elicit a fast no response, because the average strength of all related items has been increased, thereby slowing responding to nonstudied critical items.

Macht and O'Brien (1980) obtained exactly these results when a semantically related priming condition was compared to an irrelevantly primed condition. Priming was achieved by requesting subjects to examine the truth value of sentences during the retention interval of a Peterson distractor task. In the related priming condition, the sentence would be relevant to the category of the three study items (e.g., “All trees have bark”) if the study items had been fir, willow, oak. In contrast, if the unrelated (control) condition the sentences would be irrelevant to the category (e.g., “All speeches are brief”). It was found that priming facilitated responding to critical items that had been studied but inhibited rejection of nonstudied items. The effects of priming thus interacted with list status of the critical probe, as would be expected from the Atkinson and Juola model.

The results of Macht and O'Brien must be contrasted with those of three other studies, reported by Taylor and Juola (1974), Neely, Schmidt, and Roediger (1983), and Johns (1985). The procedures and primes used in these three sets of experiments, as well as their results, were rather similar. The method of Taylor and Juola exemplifies that paradigm.

Taylor and Juola (1974) used recognition probes as primes. A probe serving as critical item on test trial N of a prememorized list experiment was preceded on trial N - 1 by a probe functioning as the prime, which was from the same or from a different category. Comparison of these two conditions allowed assessment of semantic priming. Orthogonal to that, probes on trials N or N - 1 could be lures or targets. Episodic contributions of priming were observed when the prime (trial N - 1) was a target as opposed to a lure. A third factor, of course, was the list status of the critical item.

As an example, consider a list made up of the items apple, pear, lion, and tiger. If at test the probes dog and lion followed each other, semantic priming is observed for a target (lion). The magnitude of the semantic priming is assessed in comparison to a sequence such as cherry-lion, because these two items do not share the same category. In both of these examples, no episodic priming is present because the primes (dog and cherry) had not been studied. To observe episodic priming, a sequence such as lion-peal is needed. Semantic priming would be absent because these two items do not share the same category. However, because the prime (lion) had been on the list, episodic priming would be present. The magnitude of the episodic priming is assessed in comparison with the sequence dog-pear, in which the prime was not on the list.

Note how type of priming and the type of processing required for the prime are independent: The type of priming (semantic vs. episodic) is solely determined by the relation between the prime and the critical item. If prime and critical item share a category, semantic priming can be observed. If the prime had occurred on the study list, episodic priming can be observed. Independent of that, the prime itself can require an episodic response (e.g., a recognition decision as in Taylor and Juola) or a semantic response (e.g., a sentence-verification as in Macht and O'Brien).

Taylor and Juola (1974) found that responding to the probes serving as critical items (regardless of whether they were lures or targets) was facilitated if the probe on trial N - 1 was from the same category (semantic priming). Furthermore, if the N - 1 probe was a target, responding to a target or a lure on trial N was also facilitated (episodic priming). The point of greatest interest is that in their experiment semantic priming facilitated rejection of lures, contrary to the findings of Macht and O'Brien (1980). Taylor and Juola explained the data by assuming that semantic and episodic priming affected only the subsequent encoding of a probe, not its familiarity.

Similar results were obtained by Johns (1985) and by Neely et al. (1983). These studies were aimed at a different question altogether (Johns) or used a somewhat different design (Neely et al.), thus requiring some replotting of the data. But, if the data from both studies are interpreted in the framework of the Taylor and Juola design, they replicated the main effects of semantic and episodic priming as found by Taylor and Juola. These data were again different from the results obtained by Macht and O'Brien (1980).

Neely et al. (1983) offered one possible explanation for the discrepancy between their own overall facilitation resulting from semantic priming and the interaction with the list status of the critical item obtained by Macht and O'Brien. Neely et al. suggested that the difference in retention interval between Macht and O'Brien's Peterson distractor technique (15 s) and their own study-test paradigm (150 s) was the probable cause, arguing that familiarity-based responding is only active at short retention intervals, whereas search is predominant at longer intervals “with a resulting facilitation on lure rejection from related primes” (p. 202).

Neely et al. thus correlated the occurrence of inhibitory priming with familiarity-based recognition responses, whereas they considered facilitation to be the result of memory search. Of crucial importance to their retention interval hypothesis, therefore, is to demonstrate that familiarity-based responding is predominant at short retention intervals but absent at longer retention intervals. The technique used to demonstrate the presence of familiarity-based responses entails the presentation of negative probes that the subject has seen quite recently. In comparison to novel lures, these repeated negatives are rejected more slowly (or less accurately), which is seen to result from the familiarity assessment. Contrary to the reasoning underlying Neely et al.'s hypothesis, such a disadvantage for recently presented lures can be observed no matter whether only a few seconds (Monsell, 1978), 20 s (Gorfine & Jacobson, 1973), approximately 10 min (Bennett, 1975), or up to 20 min (Atkinson & Juola, 1973) elapse between presentations of a repeated lure. Moreover, that disadvantage also appears to be quantitatively invariant across retention intervals. Responses to repeated lures were found to be slowed by approximately 75 ms, regardless of whether the interval between presentations was a few seconds (Monsell, 1978) or almost 20 min (Atkinson & Juola,
1973). Consequently, the crucial assumption underlying Neely et al.'s hypothesis appears untenable.

A difficulty also exists for the explanation Taylor and Juola (1974) offered for their own data, which ascribed the effects of priming to an encoding stage. One noteworthy aspect of the Johns (1985) data is that in her experiments processing of lures was facilitated through same-category priming only if that category had been represented on the study list. Put another way, semantic priming was contingent on the presence of an episodic relation (see also Herrmann & Harwood, 1980). Thus, the encoding hypothesis of Taylor and Juola would have to "predict" that a prime facilitates encoding of a subsequent probe only if the category they both share had been present at study. It is unclear how this could be done without ad hoc assumptions.

To summarize, prior research on the priming of recognition memory for categorized lists has produced two distinct outcomes. One outcome consists of a facilitation for the rejection of lures, and the other outcome involves an inhibition of lure processing. Regardless of how lures are affected, targets always appear to be facilitated. Neely et al. proposed an explanation based on a difference in the recognition processes that predominate at different retention intervals. According to their explanation, recognition decisions at short and long retention intervals are largely based on familiarity assessment and memory search, respectively. There is enough evidence, however, to reject that contention. Consequently, one must look elsewhere to resolve the empirical puzzle posed by the priming literature. In particular, there is one other crucial difference between the two conflicting sets of experiments: namely, the type of processing required on the primes. As will be recalled, Macht and O'Brien (1980) used a sentence-verification task to achieve priming of recognition, whereas the studies resulting in facilitation of lure processing measured priming as a function of a prior recognition test item. This difference in prime processing may turn out to be critical, considering that previous research has demonstrated how a change in the processing of the prime can substantially alter its effects.

For example, Smith (1979) discovered that a semantic priming facilitation in a letter-search task virtually disappears if little lexical processing is performed on the prime. In her experiments, no facilitation was observed when subjects only had to search for a letter in the prime. When subjects had to read the prime, on the other hand, the subsequent letter search involving the critical item was facilitated. Similar results were obtained by Smith, Theodor, and Franklin (1983) who continued to explore the phenomenon, this time using a lexical-decision task. When the depth of processing required for the prime was raised from shallow (letter search) to deep (semantic analysis), the corresponding net benefit of semantic priming was increased by more than 100 ms. Irwin and Lupker (1983) and Henik, Friedrich, and Kellogg (1983) also demonstrated that the magnitude of priming depends on the degree to which the prime is processed.

There is also at least one study that indicates that the qualitative effect of a prime may be reversible through a processing manipulation. Blaxton and Neely (1983) found that switching from reading of the primes to generating them, given their first letter (e.g., fruit-A), turned a semantic facilitation into an (albeit nonsignificant) inhibition.

It is clear, therefore, that the type of processing done on the prime determines the amount of priming in a variety of tasks. In addition, the direction of the priming effect (inhibition vs. facilitation) may sometimes be altered through changes in prime processing. Consequently, the hypothesis is put forward that the discrepant outcomes in primed recognition are explainable through differences in the processing of the primes. There are a number of differences between the Macht and O'Brien sentence-verification primes on one hand and the probes that functioned as primes in the studies observing facilitation on the other hand. Two of these differences appear to be of potential importance for the discrepant pattern of outcomes. First, inhibition and facilitation were obtained when the prime was semantically and episodically (Tulving, 1972, 1983) processed, respectively. Second, the tasks differed along the dimension of what information was necessary to process the prime. Inhibition was obtained when more associative information (Murdock, 1974) was required (to verify the sentences in Macht & O'Brien, 1980), and facilitation resulted from a response that relied predominantly on item information: namely, probe recognition.

This article examines whether the episodic versus semantic or the item versus associative dimension is more predictive of outcomes in primed recognition memory. In particular, it was first attempted to establish the type of prime processing as the critical determinant of the results (Experiments 1 and 2). The subsequent experiments (3 and 4) were aimed at delineating the conditions under which the type of prime processing is important. These experiments also generalized the results to different types of primes. All experiments used the study–test paradigm, in which subjects participate in several trials, each involving the presentation of a study list followed by a brief distractor activity and the recognition test list. Priming was implemented by interspersing primes in the probe sequence at test.

**Experiment 1**

The purpose of the first experiment was to reproduce the pattern obtained by Taylor and Juola (1974) and Neely et al. (1983) in a standard study–test procedure. Experiment 1 also attempted to reduce a slight confounding that was present in these earlier studies. These experiments had found that responding to a critical item was facilitated through prior processing of an intralist item as opposed to a prior lure. One has to note that this episodic priming effect is confounded with response sequences: If a target is primed episodically, two yes responses in a row are required, whereas if a critical target is preceded by a lure prime, a no-yes sequence is required. Because of the usual arrangement of response keys in a computer-controlled experiment, the imbalance in the response sequence also

---

2 It must be noted that McKoon, Ratcliff, and Dell (1985) provided evidence to the contrary. They observed priming between lures drawn from an extralist category. Although the McKoon et al. data thus call into question the generality of the Johns (1985) and Herrmann and Harwood (1980) results, they also reduce the viability of the encoding hypothesis even further. That hypothesis must now "predict" why a prime sometimes facilitates subsequent encoding of a lure from an extralist category, and why it sometimes does not.
implies the presence of possible response repetition facilitation for episodically primed targets which is absent for the unprimed situation. A similar, albeit reversed imbalance exists for critical lures. Although it is virtually impossible to remove this confounding entirely, an attempt was made here to at least minimize it by assigning responses to primes and critical items to separate hands.

In line with the prior results, the following predictions were made: Responding to a probe should be facilitated through semantic priming, regardless of whether the critical probe was a target or a lure. Similarly, episodic priming was expected to occur. Again, list status of the critical item was expected to be of no relevance for episodic priming.

Method

Subjects and apparatus. The 32 subjects tested in Experiment 1 were volunteers from the University of Toronto downtown campus community. They participated either for a $4 per hour remuneration or in order to fulfill course requirements. All experiments were conducted by using an IBM Personal Computer with a BMC graphics monitor. During testing, subjects were seated in an Industrial Acoustics sound-attenuating chamber. The same subject pool and apparatus was used in all subsequent experiments as well.

Experimental design. The experimental design was a $2 \times 2 \times 2$ completely within-subjects design. The first factor was the list status (extra- vs. intralist) of the probe that served as a prime. This manipulated episodic priming. The second factor referred to the list status of the critical item, which could either be a target or a lure. The third and last factor was the relation between prime and critical item. Potential semantic priming was observed when both the prime and the critical probe shared the same category. The magnitude of the priming effect was measured relative to the condition in which prime and critical probe came from two different categories. The last two of these experimental variables were present in all subsequent experiments as well.

Materials. For this and the subsequent experiments, stimuli were taken from the Toronto Categorized Word Pool (Murdoch, 1976). The study list was composed of four categories, with eight words drawn from each category. Subjects participated in eight study-test trials in the single experimental session. The recognition test list consisted of all old items from the study list and an equal number of new items also drawn from the same studied categories. No items from extralist categories were included in this or in any of the subsequent experiments.

Words in the study lists were presented one at a time but blocked by category. The name of each category (e.g., DISEASES) was followed by instances of that category (e.g., INFLUENZA, AMNESIA), and then another category was shown in the same fashion. The block of categories assigned to a particular study-test trial and their order within a trial were randomly determined for each subject. Furthermore, items that served as list items for one half of the subjects served as lures for the other half of the subjects and vice versa. The construction of the study list proceeded in the same manner for all subsequent experiments as well.

Experiment 1 used a total of 16 words per category as study items and lures. These have been drawn from the possible 32 words of the Categorized Word Pool in a quasirandom fashion; uncommon items of very low frequency were excluded.

Construction of the test list (64 items total: half targets, half lures) proceeded as follows: Recognition probes were arranged into 32 pairs, in which each pair constituted a replication of one of the eight experimental conditions. There were thus 4 observations per cell per trial, or a total of 32 observations for a subject contributed to each condition. Table 1 presents an example for a study and test list used in Experiment 1. A pair of probes consisted of the prime probe followed by the critical item; each category contributed one critical probe to each condition. The corresponding prime was taken from the same category for the semantic priming conditions, and it was taken from the category immediately following or immediately preceding the particular category in the study list for the unrelated priming conditions. This scheme of combining categories for the different-category condition was used in all subsequent experiments as well. Primes themselves could be either targets or lures. This manipulation was orthogonal to the category status of the primes and implemented episodic priming. The number of items intervening between prime and critical item at study for episodically primed targets was identical for both the same- and the different-category condition. Primes preceded episodically primed critical items in the study list as often as critical probes preceded primes. Furthermore, the probability of an item following a probe to be from the same category was equal to the probability that the item was taken from the one different category randomly yoked with that prime.

Procedure. Subjects were tested individually. Instructions were given in writing and had to be recalled by the subjects in order to ensure their understanding. Instructions emphasized both accuracy and latency. Subjects were told that the purpose of the experiment was to investigate motor aspects of recognition memory performance.

Each trial's study list was presented for 1.6 s per item with a 400-ms interitem interval. The interval between each new category was used to display the category label, which then stayed on the screen until the first instance of that category had been presented. Following study list presentation was a 10-s distractor task consisting of counting backwards from a random 3- or 4-digit number by threes, which was in turn followed by the test list. A warning signal (a beep on the computer's speaker) preceded the beginning of the test list. During the test list, all stimuli remained on the screen until a response was made. Trials were separated by a short break period (less than 1 min), during which subjects were informed of their cumulative accuracy and reaction times for each condition involved.
critical items. These presentation parameters were used in all subsequent experiments as well.

During the presentation of the test list, a fixation signal (a plus sign) signalled the onset of a test pair. After 700 ms, the first member of the pair (prime) was presented. For half the subjects this prime was shown on the left side of the screen, and for the other half of the subjects it was shown on the right. After responding, the primes were followed 400 ms later by the critical probe on the opposite side of the screen. In order to eliminate possible motor priming effects as far as possible, participants responded with that hand on whose side the item was displayed. For each response hand, two keys were chosen to score yes and no responses. These keys were the backslash (\) and Z keys for the left hand, and the period and slash (/) keys for the right hand. For both hands the outer response key was used for positive responses ("Yes, I have seen this item before"), whereas the inner key served to indicate rejection of a probe. Thus, regardless of episodic priming condition, primes were responded to with a different hand than critical probes. Following responding to the critical item there was a 500-ms blank interval prior to onset of the next pair.

Subjects received several practice trials on the response key arrangement prior to the experiment. During that practice, test items were mimicked by the commands yes and no, to which subjects had to respond using the above response keys. Immediate feedback was given after every response.

**Results**

**Accuracy.** Accuracy was calculated in two ways: First, the number of errors committed on critical items was recorded regardless of the response to the preceding prime. These errors are reported below as unconditional errors. In addition, conditional accuracy was computed, which required correct responding to both the prime and the critical item.

Overall conditional accuracy was 78%. A 2 x 2 x 2 analysis of variance (ANOVA) on conditional errors did not detect any differences among conditions. Unconditional error rates were found to be the same for all critical targets (.10), regardless of priming condition. The same was true for lures, although their average error rate was slightly higher (.13). No further analyses were performed on this measure. Both types of error rates are shown in Table 2.

**Response latency.** Mean reaction times for responses to critical probes were computed in the following way for all experiments: Means for each subject included conditionally correct responses only. These means were formed by logarithmically transforming the raw scores (to eliminate an apparent skew in the distribution) and averaging these transformed scores for each experimental condition. Prior to entering that mean into the final ANOVA, the antilog was taken in order to arrive at numeric values that were directly interpretable. In computing each subject's mean, latencies longer than 2,000 ms were deleted from Experiment 1, resulting in a loss of 1% of the observations.

All analyses in this paper were carried out with a Type I error level of .05. By using that criterion, three significant main effects on reaction time were discovered. First, the list status of the critical probe (target vs. lure) was significant, with $F(1, 31) = 54.05$, $MS_e = 10,741$. The corresponding means were 836.3 ms and 741.1 ms for lures and targets, respectively. Second, there was a main effect for semantic priming. Same category primes resulted in faster (769.5 ms) responses to critical probes than different category primes (807.9 ms), regardless of whether the critical probe was a target or a lure. With $F(1, 31) = 32.57$, $MS_e = 2,907$, the difference in means was highly significant. Similarly, the list status of the prime affected responding to the critical probe, $F(1, 31) = 15.74$, $MS_e = 4,056$, with intralist primes (772.9 ms) resulting in faster responses than extralist primes (804.5 ms). Figure 1 and Table 2 present the means for all conditions of Experiment 1.

As is apparent from Figure 1, episodic priming was less effective for critical lures than it was for critical targets. This was confirmed through the presence of an interaction between list status of prime and the list status of the critical probe, $F(1, 31) = 4.23$, $MS_e = 2,415$. No other two-way or three-way interactions approached significance.

**Discussion**

The finding of a main effect of semantic priming corroborates the results of Taylor and Juola (1974), Johns (1985), and Neely et al. (1983). The experiment also showed a main effect for episodic priming, again in line with some of the previous results. Considering the present arrangement of the response keys, the contribution of motor repetition to this effect should have been minimized. Yet, it may not have been eliminated completely, as shown by the fact that episodic priming was greater for critical targets than lures.

Brief consideration must be given to an artifact in experi-
Table 2  
Average Reaction Times (in Milliseconds) and Accuracies for Critical Probes in Experiment 1 as a Function of Category Status and of List Membership of the Primes

<table>
<thead>
<tr>
<th>Category of prime</th>
<th>Target: Extra*</th>
<th>Target: Intra*</th>
<th>Lure: Extra*</th>
<th>Lure: Intra*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>740</td>
<td>696</td>
<td>825</td>
<td>817</td>
</tr>
<tr>
<td></td>
<td>.77</td>
<td>.79</td>
<td>.77</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>.90</td>
<td>.90</td>
<td>.87</td>
<td>.86</td>
</tr>
<tr>
<td>Different</td>
<td>786</td>
<td>742</td>
<td>866</td>
<td>837</td>
</tr>
<tr>
<td></td>
<td>.79</td>
<td>.78</td>
<td>.75</td>
<td>.74</td>
</tr>
<tr>
<td></td>
<td>.90</td>
<td>.90</td>
<td>.86</td>
<td>.86</td>
</tr>
</tbody>
</table>

Note: Only correct responses to both prime and critical probe are included in the RT measure, and individual observations underlying each subject's mean are log-transformed, with values exceeding 2,000 ms not entering in the analysis. Means across subjects are based on antilogs of each individual's mean. Trimmed observations count as conditional errors, but are not part of the unconditional error measure.

Figure 1. Obtained reaction times for all conditions in Experiment 1.
as a prime, subjects in Experiment 2 had to perform a category judgment on two items, because this made construction of items requiring a no response easier. As an example, subjects had to respond yes to a pair such as DOG-CAT (because both items were members of the same category). The appropriate response to a pair such as DOG-TABLE, on the other hand, was no. This task is clearly semantic, because it can be performed without any prior study phase. It is also associative, because it involves interrelating two items. Consequently, the primes in Experiment 2 differed from the ones used in Experiment 1 on those two dimensions that had been identified at the outset as plausible causes for the discrepant patterns. In addition, Experiment 2 used the same retention interval as Experiment 1, thus controlling the variable that Neely et al. thought to be responsible for the discrepant outcomes.

Method

Subjects and arrangement of materials. Sixteen subjects participated in Experiment 2. Unlike the first experiment, it was necessary to use all 32 items from each of the categories. Each category was split into two sets of 16 words each, which served as lures and targets, respectively, for every second subject.

Experimental design. The experimental design was a 2 × 2 completely within-subjects arrangement. The first factor was the list status of the critical item (lure vs. target), whereas the other factor determined the category relation between prime and critical probe (same vs. different). Because the primes required a semantic judgment and were not studied, the third factor used in Experiment 1 (list status of primes) was not included here. All subsequent experiments also used this design.

An example for a study and test list used in Experiment 2 is given in Table 3. As can be seen from the table, the test list consisted of a sequence of filler–prime–critical probe triplets that were constructed as follows: First, probes were chosen at random from each category to serve as fillers. Each category contributed four list items as target fillers and four others as lure fillers. The remaining four study items from a category served as critical probes, together with four other lures from the same category.

Each prime, on the other hand, consisted of a pair of items that were taken either from the same category or from two different categories. Subjects had to make a judgment as to whether both pair members were from the same category. There were 16 prime pairs of each kind, and none of the items serving as members of a pair appeared elsewhere in the experiment. Only when both members were from the same category could primed responses be observed. These primes (e.g., CAT-DOG) are therefore referred to as critical or experimental primes. The other mixed-category primes (e.g., CAT-TABLE) were needed only to equalize probabilities of yes and no responses. Of the 16 experimental primes, one from each category was assigned to one of the experimental conditions. Of the 16 mixed-category primes, half were assigned to be followed by targets and half to be followed by lures. These targets and lures were taken from one of the categories present in the mixed-category prime.

Each category also provided one critical probe for each of the four experimental conditions, yielding four observations per trial and subject, or a total of 32 per subject per condition. Thus, there were four critical lures per trial that were preceded by a prime pair whose members both shared the probe's category. There also were four lures preceded by prime pairs whose members were from a different category. The same was the case for critical targets. All sequence probabilities were thus held equal. Given a prime such as CAT-DOG, for instance, the following item could either be a lure or a target with equal probability, and orthogonal to the list-status dimension the probe could either be from the same category (MAMMALS) or from the one other category which had been randomly yoked with the prime's category for that trial (e.g., FURNITURE). The same sequence characteristics were true for mixed-category primes such as CAT-TABLE, although recognition responses following these primes were of no interest.

The test list consisted of a random sequence of 32 such triplets altogether. The order of conditions (i.e., of filler-prime-critical probe triplets) was randomized separately for each subject and each trial.

Procedure. Instructions emphasized accuracy over latency, especially for the category membership judgments, because pilot work had shown that to be required to maintain performance at a high level.

The study list and the distractor task were presented in exactly the same fashion as in Experiment 1. Thereafter, each triplet of the test list was presented in the following way. A word (the filler) appeared in the center of the screen, and subjects responded by pressing the slash (/) key (lower right of the keyboard) for a yes response if they recognized the filler probe as old or the backslash (\) key (lower left) for a no response. These response keys were also used in Experiments 3 and 4. After a 400-ms interval, the prime was shown in the same location as the preceding probe and also required a yes–no decision (“Yes, these items belong to the same category”, or “No, they do not”) by using the same response keys. Critical probes were displayed 400 ms after the response to the prime. Critical probes were identical to filler probes as far as the subjects were concerned, and all intervals— intra- as well as intertriplet—were equal.

During practice, fillers and critical probes were replaced by the words yes or no (randomly chosen), and the primes were mimicked through pairs of numbers between 1 and 9 which either were (subjects had to respond yes), or were not (respond no) both odd or both even. Response keys and type of print were identical to those used later in the test sequence.

Results

Accuracy. No effects for conditional accuracy were observed in the overall 2 (list status of probe) × 2 (prime category)
TABLE 4
Average Reaction Times (in Milliseconds) and Accuracies for Critical Probes in Experiment 2 as a Function of the Category Status of the Prime

<table>
<thead>
<tr>
<th>Category of prime</th>
<th>Target</th>
<th>Lure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean RT</td>
<td>722</td>
<td>948</td>
</tr>
<tr>
<td>Conditional accuracy</td>
<td>0.81</td>
<td>0.73</td>
</tr>
<tr>
<td>Unconditional accuracy</td>
<td>0.93</td>
<td>0.86</td>
</tr>
<tr>
<td>Different</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean RT</td>
<td>771</td>
<td>905</td>
</tr>
<tr>
<td>Conditional accuracy</td>
<td>0.80</td>
<td>0.77</td>
</tr>
<tr>
<td>Unconditional accuracy</td>
<td>0.90</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Note. Only correct responses to both prime and critical item are included in the RT measure, and individual observations underlying each subject's mean are log-transformed, with values exceeding 2,000 ms not entering in the analysis. Means across subjects are based on antilogs of each individual's mean. Trimmed observations count as conditional errors, but are not part of the unconditional error measure.

Experiment 2 also provided the beginning of an explanation for these discrepant outcomes, based on the type of processing required for the prime. The task used to achieve priming has been clearly established as a critical determinant of what happens to processing of lures: A semantic and associative task resulted in inhibition (Experiment 2), whereas an episodic task (Experiment 1) resulted in facilitation. It is not clear yet, however, whether the priming task must be associative in nature, in addition to being semantic, to produce an inhibition of lure processing. Experiments 3 and 4 were designed to investigate the effects of an item-predicated yet semantic priming task.

EXPERIMENTS 3 AND 4

Experiments 3 and 4 observed priming as a function of an item-predicated semantic task, by requiring a lexical-decision task to be performed on the prime. The lexical-decision task corresponds most closely to a single item episodic recognition task as it was used in Experiment 1. The only difference between Experiments 3 and 4, in turn, was the length of the study list.

EXPERIMENT 3: METHOD

Subjects and arrangement of materials. Twenty-two subjects participated in Experiment 3. Each category in the Categorized Word Pool was randomly split into two sets (A and B) of 16 items each. Identity of Sets A and B was alternated for every second subject. Eight items were randomly sampled from the respective Set A to form one category on the study list, and eight items were selected from Set B to provide lures for that category on the test list. Categories were randomly allocated to trials.

For the test list, four nonstudied items from Set A were taken to function as word (experimental) primes for that given category. Pseudoword primes were obtained by changing the spelling of four members of Set B of that category such that they became pseudowords. A subject would never see items from Set B and their pseudoword derivations together in the experiment. These pseudowords obeyed English spelling rules and were pronounceable. They also maintained a resemblance to items

Figure 2. Obtained reaction times for all conditions in Experiment 2.
Table 5
Examples of a Study List and a Test List Used in Experiments 3 and 4

<table>
<thead>
<tr>
<th>TOOLS</th>
<th>SCIENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>JIGSAW</td>
<td>ARCHITECT</td>
</tr>
<tr>
<td>BOLTS</td>
<td>PROFESSIONS</td>
</tr>
<tr>
<td>SCREWS</td>
<td>PSYCHOLOGY</td>
</tr>
<tr>
<td>SCREW</td>
<td>TOPOGRAPHY</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test list pair</th>
<th>Category of prime</th>
<th>List status of critical probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>word ? HAMMER</td>
<td>same</td>
<td>target</td>
</tr>
<tr>
<td>BOLTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>word ? WRENCH</td>
<td>different</td>
<td>target</td>
</tr>
<tr>
<td>PSYCHOLOGY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>word ? GEOGRAPHY</td>
<td>same</td>
<td>lure</td>
</tr>
<tr>
<td>ASTRONOMY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>word ? AGRONOMY</td>
<td>different</td>
<td>lure</td>
</tr>
<tr>
<td>SCREWS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: One example is given for each experimental condition. Filler probes and nonexperimental observations (i.e., pseudoword primes) are omitted. Right-hand columns specify levels of experimental variables for prime–critical probe pair on the left.

from Set B of the category, which pilot work had shown to be necessary in order to obtain a robust priming effect. For example, pseudowords for the category PROFESSIONS were items such as MURCHENT, ARCHITECT, JANETUR, and the like. Highly similar pseudowords have been used before in priming research (cf. Antos, 1979; Lupsier, 1984), and they present an extension of other findings (Shulman & Davison, 1977; Smith et al., 1983) that priming effects are magnified if pseudowords are used that are more similar to real words. For the subsequent discussion, refer to Table 5 for an example of the kind of stimulus arrangement used in Experiments 3 and 4.

Recognition test items were distributed as follows: Half of all study items of a given category and half of the corresponding set of lures were randomly designated to serve as filler probes. A filler was the first member of each of the test triplets, similar to Experiment 2. Assignment of filler items to position on the test list was random. Half of the remaining recognition test items (two lures, two targets from each category) were assigned to follow word (experimental) primes such that each experimental condition was represented once for each category. Critical responses were observed only for those probes that followed word primes. As in Experiments 1 and 2, critical probes (lures or targets) were taken either from the same category as the prime, or from the one different category randomly yoked to it. Consequently, a given category item presented as a prime was equally likely to be followed by a probe from the same or from the different category. Similarly, because half of the critical items from a given category were targets and the others were lures, probability of a yes versus a no response following a word prime was also held equal. Finally, nonexperimental primes (i.e., pseudowords) were followed by recognition test items whose properties were the same as those following experimental primes. Recognition responses to probes following pseudowords were of no interest. Primes and the probes following them were distributed throughout the test list in a random order.

In summary, each category provided four critical probes (one in each condition) and also four word primes, yielding four observations per condition per trial per subject, or a total of 32 per subject for each condition.

Procedure. Instructions emphasized both accuracy and latency. The study list was presented in the same fashion as in the preceding experiments. Each triplet in the test sequence was shown as follows: The filler (a recognition test item) appeared on the screen and was, following responding, replaced 400 ms later by the prime (a lexical-decision item). The prime was printed in inverse video (i.e., dark green on a light green background) with the prompt "word?" to the left of it. Upon performing the yes/no lexical decision by using the usual response keys, the prime disappeared and was replaced 400 ms later by the critical probe, which was printed in normal video (i.e., bright green on a dark background). Again, following responding, the next triplet was presented 400 ms later.

Prior to onset of the experimental procedure, subjects received a practice trial consisting of 10 study items (the same across all subjects, taken from the Toronto Word Pool; Friendly, Franklin, Hoffman, & Rubin, 1982), followed by a test sequence of 20 triplets similar in setup to the experimental trials.

Experiment 3: Results

Accuracy. The $2 \times 2$ within-subjects analysis of conditional accuracy showed that there was a significant main effect of list status of the critical probe (targets: .80 vs. lures: .67), with $F(1, 21) = 40.7, M_S = .0094$. In addition, the analysis uncovered a slight trend toward more accurate responding in the same category condition than in the different category condition, $F(1, 21) = 3.13, p < .10$. Overall conditional accuracy level was 73%, which was slightly lower than in the preceding experiments. The unconditional error rates also differed between targets (.15) and lures (.29), but were identical for each response class between priming conditions. Table 6 displays the accuracy means for all conditions and their RT counterparts.

Response latency. Trimming used a cutoff of 3,500 ms. This excluded 1% of the observations from the RT analysis of variance. Paralleling the accuracy measures, critical lures were responded to more slowly (1239.5 ms) than targets (904.5 ms), with $F(1, 21) = 108.83, M_S = 22.696$. The second potential main effect in the analysis (category of prime) was far from significant, with an $F(1, 21) < 1$, whereas the interaction between priming and list status of the critical probe was again significant, $F(1, 21) = 5.95, M_S = 5.059$. The facilitation observed for critical targets fell 2 ms short of the critical significant difference for the .05 level (Duncan’s test). The inhibition for lures was also nonsignificant, but came close to attaining statistical robustness as well.

Experiment 4: Method

Experiment 4 was identical to Experiment 3 with the following exceptions: Subjects were presented with 16 trials of 16 study items (drawn from two categories) each, rather than eight trials of 32 study items each. Consequently, test list length was cut into half as well, with 48 rather than 96 responses required. Breaks between trials were shortened somewhat in order to keep the overall duration of a session at less than 1 hr. Twenty subjects participated in Experiment 4.

Experiment 4: Results

Accuracy. As in Experiment 3, the only effect on conditional accuracy was that negative probes were responded to less accur-
Average Reaction Times (in Milliseconds) and Accuracies for Critical Probes in Experiment 3 as a Function of the Category Status of the Prime

<table>
<thead>
<tr>
<th>Category of prime</th>
<th>List status of critical probe</th>
<th>Target</th>
<th>Lure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same</td>
<td>Mean RT</td>
<td>884</td>
<td>1256</td>
</tr>
<tr>
<td></td>
<td>Conditional accuracy</td>
<td>.78</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>Unconditional accuracy</td>
<td>.93</td>
<td>.78</td>
</tr>
<tr>
<td>Different</td>
<td>Mean RT</td>
<td>925</td>
<td>1223</td>
</tr>
<tr>
<td></td>
<td>Conditional accuracy</td>
<td>.82</td>
<td>.68</td>
</tr>
<tr>
<td></td>
<td>Unconditional accuracy</td>
<td>.94</td>
<td>.81</td>
</tr>
</tbody>
</table>

Note. Only correct responses to both prime and critical probe are included in the RT measure, and individual observations underlying each subject's mean are log-transformed, with values exceeding 3,500 ms not entering in the analysis. Means across subjects are based on antilogs of each individual's mean. Trimmed observations count as conditional errors, but are not part of the unconditional error measure.

Discussion

Both Experiment 3 and Experiment 4 produced an interaction between the list status of the critical probe and the priming condition. Target verification was facilitated by a prior lexical decision on a word from the same category, whereas processing of lures was inhibited. Although neither experiment on its own was characterized by a great statistical robustness, the pattern of the interaction was replicated across both studies and both retention intervals. The experiments showed that priming through a semantic task, even if that task is not associative, inhibits lure processing.

This effect could not have been the result of the mere presence of pseudowords in Experiments 3 and 4, because Experiment 2 did not contain any pseudowords and yet showed the same pattern of results. Similarly, an explanation for the results of Experiments 3 and 4 based on work by Balota and Chumbley (1984) can be ruled out. Balota and Chumbley suggested that lexical decision may also involve an assessment of the familiarity of an item, similar in spirit to the Atkinson and Juola (1973, 1974) recognition model. Thus, in Experiments 3 and 4, subjects could have generalized such a familiarity-based strategy from the lexical-decision primes to the recognition probes. Such a strategy would of course predict inhibition of lure processing. But this explanation can be ruled out, given that Experiment 2 involved a category membership judgment that, again according to Balota and Chumbley, does not involve a familiarity check.

General Discussion

Before one can discuss the implications of the results of the four present experiments, several potential causes for concern and alternative explanations must be dealt with. First, it must be ensured that performance on the primes themselves did not covary with performance on the critical items. Ideally, given proper randomization, reaction time and accuracy of responses to primes should not differ at all between conditions within an experiment. Indeed, the observed variations between means were slight: For Experiments 1–4, the largest difference in reaction time (computed with the appropriate trimming scheme) between primes for the different conditions was 20, 23, 16, and 20 ms, respectively. The corresponding maximum differences per experiment in accuracy were 1%, 5%, 3%, and 3%, respectively. Of the 32 $F$ tests in ANOVAs that were performed on the prime data for all experiments, only one was found to be unexpectedly significant. (In addition, lure primes in Experiment 1 were responded to more slowly than target primes, but this is
what is often found in a recognition experiment, and is therefore not unexpected). Consequently, differences in performance on the primes could not have contributed to the results.

Second, the overall performance level on the primes across experiments must also be close to invariant in order to ensure that primes were of roughly equal difficulty. Indeed, the grand means of accuracy did not vary much across studies, with values of 86%, 80%, 85%, and 86% for Experiments 1-4, respectively. Note how primes in Experiment 1 were responded to with precisely the same accuracy as those used in Experiment 4. Yet, those two experiments uncovered widely different patterns of results.

Third, in all experiments that showed an inhibition for the rejection of lures, subjects had to switch from one task (category judgment or lexical decision) to another (recognition) when responding to prime and critical item, respectively. Contrary to that, in Experiment 1, which obtained facilitation, subjects performed the same task on the prime and on the critical item. It is thus conceivable that this difference between experiments was responsible for the different outcomes. However, this appears unlikely, because responding to targets was facilitated through semantic priming regardless of whether the priming task was the same as the recognition task. It is implausible to assume that task switching was responsible for drastically altering responses to critical lures, given that it had no effect on target verification. Furthermore, I have conducted several other unpublished experiments that required task switching between prime and critical probe, but that nevertheless demonstrated a facilitation for lures. Taken together, these two points should allay any fears that the present results were an artifact of the task switching required in Experiments 2-4 but not in Experiment 1.

The fourth and last issue concerns another procedural difference between Experiment 1 on one hand and Experiments 2-4 on the other hand. None of the primes used in the last three experiments had been presented to subjects at study. Contrary to that, half the primes used in Experiment 1 had been studied. The reason for this procedural difference between experiments lies in the essence of the task dimension that was manipulated. An episodic task can only be performed after a prior study phase, whereas a semantic task can be (and usually is) performed on novel items that subjects have not seen previously in the same experiment. However, it is, in principle, possible to perform a semantic judgment on an item that has been studied previously (e.g., Neely & Durgunoglu, 1985). The present set of experiments cannot serve to elucidate that special case of episodically studied primes that require semantic responses. Thus, for the summary below, it must be kept in mind that primes requiring episodic and semantic responses always involved studied and novel items, respectively.

Summary of Results

The results from all four experiments can then be summarized as follows: If a purely semantic decision is required for the prime, subsequent processing of negative recognition probes is inhibited. This is the case regardless of whether the semantic judgment is one that is predominantly item predicated or one that requires associative information. If, on the other hand, a prime is used that requires an episodic judgment, processing of lures is facilitated. This facilitation is increased if the prime had been present on the study list.

The effect of primes on target probes, on the other hand, does not appear to vary as a function of the nature of the prime processing. That is, verification of targets is facilitated after an episodic as well as after a semantic priming task. Again, for a prime requiring an episodic judgment, that facilitation is greater when the prime had been studied.

This pattern is clearly inconsistent with the retention interval hypothesis of Neely et al. (1983). Inhibition of lure processing has now been obtained at retention intervals ranging from 15 s (Macht & O'Brien, 1980) through approximately 60 s (Experiment 4) to 120 s (Experiments 2 and 3). Conversely, facilitation has been obtained at retention intervals ranging from 80 s (Experiment 1) through 150 s (Neely et al., 1983) and several minutes (Johns, 1985) up to 24 h (Taylor & Juola, 1974). We now know that the type of prime processing, not retention interval or list length, is the critical predictor of outcomes in priming of recognition memory. The retention interval hypothesis of Neely et al. should thus be discarded unless future research can prove its usefulness. Beyond rejection of that hypothesis, however, how can the different effects of different priming tasks be explained?

From a purely heuristic point of view, the episodic/semantic dimension is the most useful one to capture the pattern of results. That is, we can now predict an outcome on the basis of knowledge of the type of processing that is required for the prime. A semantic task yields inhibition, whereas an episodic task results in facilitation. Beyond its predictive capabilities, the episodic/semantic dichotomy can only provide a rather vague framework for an explanation for the present data. According to that vague framework, the data may be interpreted as showing that priming may sometimes be restricted to episodic memory only, in which case the consequences of priming are different from the situation when all category items (studied or nonstudied) are primed. If episodic memory (i.e., the list) is primed only (through a recognition probe functioning as prime), subsequent episodic processing is facilitated. On the other hand, if the entire category is primed (through a prime requiring a semantic judgment), subsequent rejection of a lure is inhibited because the familiarity of all items in a category has been raised.

If this reasoning is put into the more detailed terms of an Atkinson and Juola type model, from which all research in recognition priming appears to depart, an account for the findings in the present paper can be proposed: (a) A prime that requires a semantic judgment raises the level of activation of all items in a category, studied or nonstudied. Familiarity-based responses to targets are thus facilitated, whereas these responses are inhibited for lures. This mechanism explains all inhibitory effects. (b) A prime that requires an episodic judgment, on the other hand, activates the studied list only. Consequently, targets are facilitated because the familiarity of old items from that category has been raised. To explain the effect on lures, one must assume that priming of the list also speeds up the search process (as suggested by Neely et al., 1983) for subsequent probes and that this has the same quantitative effect on lures that the rise in familiarity has on targets. (c) When the prime is not from the
same category, there is nevertheless some residual priming of the list, provided the prime requires an episodic judgment and is itself a target. This explains purely episodic priming for targets in Experiment 1. To account for the effect on lures, it must again be assumed that residual list priming also speeds up subsequent search.

This account is highly speculative and clearly post hoc, but components of it have been suggested previously. Macht and O'Brien (1980) derived their predictions by using point a from above. Neely et al. (1983) assumed a speeding of search to underlie facilitation for lures (see points b and c). What the previous points do, then, is to integrate elements from the previous literature and to specify the conditions under which each of these processes is relevant. Because those conditions (semantic vs. episodic tasks required on the primes) are defined independently and can be known a priori, the account can serve to derive predictions for other priming studies.

Conclusion

The present experiments have demonstrated that priming can facilitate or inhibit processing of lures even when list length and retention interval are controlled. The critical variable that determines whether facilitation or inhibition arises is the type of judgment that must be performed on the prime. If a semantic judgment is required (regardless of what type of information is involved) lure processing is inhibited. If the prime requires an episodic judgment, on the other hand, subsequent rejection of lures is facilitated.

These data can be explained if it is assumed that the list can be primed selectively. The idea of selective list priming is consistent with the view that episodic and semantic memory may constitute two different memory systems (Tulving, 1983). The idea is also consistent with the work of Wickens and his associates (Wickens, Moody, & Dow, 1981; Wickens, Moody, & Vidulich, 1985) who also demonstrated that lists of varying lengths are stored in memory as a well-defined and closely interlinked entity.

References


Received June 24, 1985
Revision received September 30, 1985

---

**Low Publication Prices for APA Members and Affiliates**

**Keeping You Up-to-Date**

All APA members (Fellows, Members, and Associates) receive — as part of their annual dues — subscriptions to the *American Psychologist*, the *APA Monitor*, and *Psychology Today*.

High School Teacher and Student Affiliates receive subscriptions to the *APA Monitor* and *Psychology Today*, and they can subscribe to the *American Psychologist* at a significantly reduced rate.

In addition, all members and affiliates are eligible for savings of up to 50% on other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the British Psychological Society, the American Sociological Association, and Human Sciences Press).

**Essential Resources**

APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the APA*, the *Master Lectures*, and APA's *Guide to Research Support*.

**Other Benefits of Membership**

Membership in APA also provides eligibility for low-cost insurance plans covering life: medical and income protection; hospital indemnity: accident and travel; Keogh retirement; office overhead: and student/school professional, and liability.

For more information, write to American Psychological Association, Membership Records, 1200 Seventeenth Street NW, Washington, DC 20036