

Masson, M. E. J., & MacLeod, C. M. (1997).  
Episodic enhancement of processing fluency.  
In D. L. Medin (Ed.), *The Psychology of  
Learning and Motivation*, Vol. 37 (pp. 155-210).  
San Diego, CA: Academic Press.

## EPISODIC ENHANCEMENT OF PROCESSING FLUENCY

*Michael E. J. Masson*  
*Colin M. MacLeod*

### I. Introduction

About 25 years ago, the prevalent view of memory began to change. Prior to that time, the predominant theories were framed in terms of engrams or traces that were stored in memory at the time of encoding and recovered from memory at the time of retrieval, much as money is deposited in and withdrawn from a bank. This can be seen early on in the writings of Semon (1904/1921), but this style of theory prevailed into the 1960s (see textbooks of that time, e.g., Kintsch, 1970). Encoding processes laid down traces in memory and retrieval processes searched for them. Traces were products of encoding and objects of retrieval. Clearly, processes were involved in both encoding and retrieval, but what was stored in memory was the consequences of initial processing, not records of the processing itself. This trace view is well captured by the best known models of that period, the two-state or "buffer" models of memory (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965).

Then, in the early 1970s, a new perspective emerged. In large part, this stemmed from research carried out at the University of Toronto. It began with two key elements: the renewed emphasis on processing that formed the basis of the levels of processing framework ( Craik & Lockhart, 1972); and the increased stress on the interplay between encoding and retrieval

that culminated in the encoding specificity principle (Tulving & Thomson, 1973). These fundamental ideas came together in several places, notably in the work on transfer of skills (Kolers, 1976; Kolers & Roediger, 1984) and in the concept of transfer-appropriate processing (Morris, Bransford, & Franks, 1977). Under these views, memory did not record the products of processing; rather, memory was a chronicle of the processes themselves.

In this new theoretical framework, encoding does not extract and store away a trace based on analysis of the stimulus situation; instead, the analysis itself is recorded in memory. Kolers made this point especially forcefully: "Memory is not a place in the mind where things are stored, but a way of responding to events with skills acquired in previous encounters with like circumstances; recognition is not a matter of matching but a process of transferring skills across occasions" (1976, p. 564). Remembering, then, consists of processes occurring at the time of retrieval, and is successful to the extent that the same processes are applied at both points in time—encoding and retrieval—resulting in successful transfer of processing. Morris et al. put this quite simply: "the value of particular types of acquisition activities must be defined relative to the type of activities to be performed at the time of test" (1977, p. 531).

This analysis of remembering has come even more to the fore as cognitive psychologists have begun to explore not just conscious remembering, but also unconscious remembering. Almost certainly, we use memory without being aware of it a great deal of the time, probably more than we do for deliberate recollection of our past. As just one example, when we are working at our computers, there are a great many remembered skills and facts that we use without being aware that memory is being engaged. This conscious/unconscious distinction is now firmly entrenched in the memory literature as the explicit/implicit distinction (Graf & Schacter, 1985) and is increasingly influencing our theorizing about the operation of memory. As just two illustrations, both Jacoby (1983) and Roediger (1990) have argued that the nature of the processing at the time of encoding strongly influences performance both on direct tests of conscious recollection and on indirect tests of unconscious remembering. The essence of the argument is that different types of memory tests selectively map onto different processes executed at the time of encoding.

#### A. FLUENT REMEMBERING

These ideas about the connection between encoding and retrieval are intimately tied to the concept of fluency in remembering, a concept that Jacoby and his colleagues have championed for some time. In the words of Jacoby and Brooks (1984, p. 35), "When using the fluency heuristic, the person

infers that an item must have occurred before if it can be processed relatively more easily." When an earlier processing episode relates to a later processing episode, that subsequent processing will benefit because the processes can be rerun more easily the second time. This benefit may be quite unconscious. Alternatively, the subject may experience a sense of fluency, and this impression may lead to an attribution of "remembering." So, from this perspective, conscious remembering is an attribution contingent upon fluent reprocessing.

There is already considerable evidence in accord with this position. Jacoby and Dallas (1981, p. 333) invoked this account early on in contrasting performance on recognition (explicit) and masked word identification (implicit) tests of memory: "Subjects may base their recognition memory decision on judgments of the relative fluency of their own performance of a task." More recently, in their study of false recognition, Jacoby and Whitehouse (1989) referred to an "illusion of memory" wherein "fluency of processing can give rise to a feeling of familiarity" (p. 133) that in turn affects conscious recognition judgments. In their studies of reading words aloud and making judgments about those words, Whittlesea, Jacoby, and Girard (1990, p. 716) reinforced this idea: "fluent performance is unconsciously attributed to whatever source is apparent and . . . feelings of familiarity result when fluency is attributed . . . to past experience." As will become clear, the ideas of Jacoby and his colleagues have heavily influenced the account of memory we have been developing.

#### B. DATA-DRIVEN VERSUS CONCEPTUALLY DRIVEN PROCESSING

The dominant interpretation of implicit versus explicit remembering has grown out of the transfer-appropriate processing framework (Morris et al., 1977), initially through the work of Jacoby (1983) and since then through the work of Roediger (1990). On this view, the majority of indirect tests that measure memory implicitly, where an individual need not be aware that he or she is remembering, are "tuned" principally to data level or perceptual encoding operations. Thus, changes between study and test in stimulus form (e.g., between fonts, or from word to picture) or modality (e.g., from auditory to visual) reduce performance on tasks such as word fragment completion, where a stimulus such as *s\_\_at\_\_r* must be completed with a word such as "sweater" (see Roediger & McDermott, 1993, for a review). In contrast, standard direct tests that measure memory explicitly, tests such as recall and recognition, are tuned primarily to conceptual level or meaningful encoding operations. Elaborations that direct attention toward meaning during encoding, such as producing semantic associates of the study items, especially benefit conscious recollection. In this way,

transfer-appropriate processing has been modified to capture both ways of remembering, implicit and explicit.

One procedure that has seen extensive use in exploring this processing distinction is the read versus generate manipulation. During encoding, subjects either read isolated words (e.g., *umbrella*) or produce those words from semantic clues (e.g., "This keeps the rain off your head—*u*?"). It is well established that generated words are much better remembered than are read words on direct tests like recall or recognition (Slamecka & Graf, 1978). But the reverse is sometimes true for certain indirect tests: Reading can actually lead to better remembering than generating (e.g., Jacoby, 1983; Roediger, 1990). Of course, both types of study activity are superior to baseline performance for unstudied words. This improvement due to repeated processing of a stimulus has come to be called priming.

In developing their data-driven versus conceptually driven account, Roediger and his colleagues at one point suggested that the read/generate pattern be used as a benchmark for defining whether a particular test was data driven or conceptually driven (Roediger, Weldon, & Challis, 1989). Expressing this proposal as a conditional rule, we have: if read > generate, then data-driven; if generate > read, then conceptually driven. The majority of their results, particularly those emphasized in theory development, do fit this rule. Roediger and McDermott (1993; see also Roediger, 1990) review many of the relevant studies, especially studies using the indirect word fragment completion test. The repeated finding is that having read the critical words earlier leads to more priming than does having generated them, whereas the opposite is true on direct tests of recognition or recall.

### C. INITIAL INTERPRETIVE ENCODING AND SUBSEQUENT ELABORATIVE ENCODING

In our initial examination of these ideas regarding data-driven and conceptually driven processing, we used a different indirect test and obtained a different pattern of results (Masson & MacLeod, 1992). The test we used was masked word identification (also called perceptual identification), in which the subject must identify a word visually presented for about 30 ms and followed by a pattern mask. We were able to replicate the findings of Jacoby (1983) when we used the same sort of antonym generation scheme (*hot—c*; generate the opposite beginning with the specified letter). Recall that Jacoby's results were consistent with the rule: Read words were more often identified than generated words, despite the opposite pattern in recognition. But it turned out that antonyms were virtually unique in producing this pattern. For almost every other type of generation rule that we examined in a series of 11 experiments, from definitional phrases to famous

names to synonyms, we found a different pattern: Read and generated words produced similar amounts of priming in masked word identification. The overwhelming majority of results on our indirect test, therefore, did not fit the rule.

Our results led us to propose an alternative account of how encoding and retrieval operated. This account owes a considerable intellectual debt to an explanation first set out by Graf and Mandler (1984), who discriminated between integrative and elaborative encoding. We also suggested that subjects produced two encodings at the time of study. The initial interpretive encoding was made immediately upon perceiving the stimulus, and incorporated both conceptual and perceptual aspects of the stimulus as processed (and, indeed, any other available aspects). Reenacting this initial interpretive encoding was seen as underlying any priming on subsequent indirect tests, such as masked word identification, that involved identifying or in some other way producing a specific target stimulus. Thus, a studied word should show priming on such a test, regardless of whether it had been generated or read. An elaborate encoding could also be created, but this was subsequent to interpretive encoding and more subject to deliberate control. It was this "deeper" encoding, to use levels of processing language, that dominated performance on direct tests, explaining why generated words showed better explicit retention than did read words.

The general principle that we intended to emphasize by proposing the concept of an initial interpretive encoding is that the primary goal of encoding a stimulus is to interpret or classify the stimulus, so that appropriate knowledge about it can be called to mind. Adoption of that general principle had its roots in early work that demonstrated the significant influence that context has on encoding and retrieval (e.g., Masson & Sala, 1978; Meyer & Schvaneveldt, 1971; Neely, 1991; Rumelhart, 1977).

As an aside, we should note that our view actually is somewhat more radical than we have set out this far. In fact, we do not see the distinction between encoding and retrieval as particularly valuable any longer. Each encounter with a stimulus involves interpretation and potentially elaboration, whether that encounter is nominally the first one (encoding) or a subsequent one (retrieval). Furthermore, each encounter necessitates operations corresponding to what traditionally have been called encoding and retrieval. Identifying the operations is an important goal of research; maintaining an artificial dichotomy between encoding and retrieval operations would not appear to provide additional benefits. Processing overlap across episodes supports remembering whether that remembering is measured implicitly or explicitly. The "value added" of the encoding/retrieval distinction is unclear.

We return now to the concept of initial interpretation as involving both perceptual and conceptual aspects. More recent theoretical developments provide additional encouragement for our view that both aspects of a stimulus are collaboratively involved in the encoding operations that underlie repetition priming effects. Two articles are especially relevant in this regard. First, Strain, Patterson, and Seidenberg (1995) have shown that a conceptual variable, imageability of a word, influences speeded word reading even when words are presented in isolation. The important implication of this result is that even though the subject's task is to read aloud, with no inducement to engage in semantic elaboration of the word, conceptual knowledge associated with the word appears to be recruited.

Second, Becker, Moscovitch, and Behrmann (in press) have developed a connectionist model of semantic priming based on long-term changes to connection weights. One important prediction derived from this model is that semantic priming can have long-lasting effects. More important for our purposes, though, is the fact that the model constitutes an implementation of the idea that a specific processing episode can have a long-term effect on the fluency of word identification. Moreover, with a model such as theirs, that long-term influence involves not only perceptual aspects of a word's representation, but conceptual aspects of that representation as well. Although we have yet to explore this very promising development, we expect that a model of this class should be useful in deriving testable predictions regarding repetition priming effects.

In the remainder of this chapter, we describe several series of experiments that we have conducted in our laboratories and that reveal the principles determining the influence of specific processing episodes on subsequent word identification. We begin with a summary of the initial series of experiments that led to our adoption of the distinction between interpretive and elaborative encoding. Subsequent series of experiments to be described were aimed at (1) providing converging evidence that conceptual and perceptual processes jointly contribute to the priming observed in tasks such as masked word identification, and (2) testing hypotheses regarding how memory for prior episodes modifies word identification processes. All of these issues come together in the concepts of fluency and interpretive encoding.

## II. Experiment Series 1: Data-Driven and Conceptually Driven Encoding Tasks

The experiments reported by Jacoby (1983) provided evidence for two fundamentally important claims regarding episodically enhanced word

identification. First, improved identification of briefly exposed target words was dependent on prior perceptual processing of those words. Second, improved target identification was not a result of conscious recollection of the previously presented words. Jacoby used three encoding tasks, two of which are important for the current discussion: reading a target word in isolation and generating a target word from its antonym (e.g., *hot*—??). In the latter task, the target itself was not actually presented to the subject, but was experienced only in the act of generating it in response to its antonym. Subjects were then tested on the encoded words and a set of nonstudied words in one of two ways. One test was a yes/no recognition test in which subjects decided whether each target word had been presented earlier (either read or generated). The other test was the masked word identification task. Jacoby found that recognition memory was more accurate for generated than for read targets, but the reverse was true for performance on the masked word identification task. Moreover, little or no priming was found on masked word identification for targets that had been generated in the encoding phase.

The double dissociation between the recognition memory and masked word identification tests provided powerful evidence that conscious recollection was not responsible for the pattern of results observed in the masked word identification task. The finding that generation at encoding led to little or no enhancement on the masked word identification task indicated that performance on that task was affected almost exclusively by prior perceptual (data-driven) experiences. This outcome fits nicely with the transfer-appropriate processing framework (Morris et al., 1977; Roediger et al., 1989), by which the influence of prior processing episodes on subsequent tests is assumed to depend on the involvement of similar processing operations at study and test. The read encoding task was deemed to involve data-driven processing and the generation task was confined to conceptually driven processing. The fact that the masked word identification task showed substantial benefit only from the read encoding task implies that the masked word identification task should be classified as data driven.

Although the results obtained by Jacoby (1983) provided convincing support for the view that enhanced performance on the masked word identification task could be brought about only by perceptual processing experience, we were concerned about another factor that was at work in those experiments. Items encoded in the read task might have been placed at an advantage because the context in which they appeared during study and test was the same, namely, the word appeared by itself with no surrounding context. In the case of words encoded in the generation task, however, targets were processed in conjunction with a strongly associated context word (the antonym cue) but tested in isolation. The change in

contextual environment experienced with items in the generation encoding condition might have contributed to the lack of priming for those items.

To test this possibility, we designed an initial experiment in which the contextual information used to generate targets during study was present or absent at the time of test (Masson & MacLeod, 1992). A related study by Toth and Hunt (1990) provided preliminary support for this approach. They presented subjects with a semantic associate for each target at the time of study. The study tasks involved reading some targets and generating others from fragments in which one or two letters of the target were missing. The context word might be expected to play a larger role in the generate condition because some additional work had to be done to determine the target's identity. Subjects were then given a masked word identification test in which the target word was tested either in the presence or absence of its semantic associate. Toth and Hunt found that identification of items in the generate condition was improved by the presence of the relevant associate to a greater extent than was the case for items in the read condition. Although this result suggests that generated targets have the potential to benefit more from the reinstatement of contextual information provided during study, the generation task used by Toth and Hunt was not optimal for addressing the issue of concern here. Generating targets from word fragments with only one or two letters missing seems to involve a strong data-driven component because it was found that even when context was not reinstated at test, identification performance on generated targets was at least as high as performance on read items.

In our initial experiment (Masson & MacLeod, 1992, Experiment 1), subjects generated targets from a brief definition and the first letter of the target word (e.g., "the orange vegetable a rabbit eats—c?"). By providing only the first letter of the target word, we avoided the strong data-driven processing that apparently occurred in Toth and Hunt (1990) study. For generation cues, we used definitions instead of antonyms, as Jacoby had done, because pilot testing revealed that when an antonym was presented as a context word during the masked word identification test, there was a strong propensity for subjects to respond with the context word's antonym even though that was not the target word. By using definitions as cues, we were able to select a representative word from the definition cue to use as a context word during the masked word identification test. The selected context words would not have sufficiently powerful associations with the targets to induce subjects to adopt the strategy we had observed with antonyms as context words.

The design of the experiment was a  $3 \times 2$  factorial within subjects design. Three encoding conditions were used: generate, read, and new (items not presented during the encoding phase). The second factor was test condition:

context and no context. In the study phase of the experiment, subjects were presented with two kinds of trials randomly intermixed. For targets assigned to the read condition, the target word appeared by itself at the center of a computer monitor until the subject read it aloud. On generation trials, a cue consisting of a brief definition and the first letter of the target appeared on the monitor and the subject attempted to name the intended target word. The encoding phase was followed by a test phase involving the masked word identification task. Three types of target were tested: generate, read, and new. Presentation of a target was preceded by a context stimulus consisting either of a row of Xs (the no context condition) or a relevant context word (a word taken from the definition cue for that target). In the case of generate targets, the context word had appeared during the study phase as part of the target's encoding episode. For the read and new targets, the context word was related to the target only by preexperimental experience. The context stimulus was in view for 250 ms. The computer monitor was then blank for 250 ms, before the target word was presented in lowercase letters. The target appeared for 33 ms at the center of the monitor and was immediately followed by a pattern mask. The subject then attempted to identify the target.

In the study phase, subjects typically had no difficulty in determining the correct target on generation trials. Across the experiments described in this chapter, subjects usually averaged over 90% correct when generating targets in the study phase. The results we report from the generate condition of these experiments are based only on generate items that were correctly generated during the study phase. To ensure that this conditionalized scoring of generate targets did not lead to item selection artifacts, we conducted parallel analyses that included all generate targets, regardless of whether they were correctly generated at study. These analyses almost invariably produced the same pattern of results found with the analyses of conditionalized data, indicating that item selection effects played little if any role in our comparison of generate and read encoding conditions.

The mean proportions of targets correctly identified as a function of study and test condition are shown in Fig. 1A. There was a significant effect of study condition, indicating that prior exposure led to higher probability of identification on the masked word identification task. There was also an effect of context, with better identification performance observed when targets were preceded by related context words. These two factors, however, did not interact. The lack of interaction means that the prior study episodes involving generation of target words from semantic cues provided no additional potency to the context words drawn from those cues. Rather, the benefit of the context words was derived entirely from preexperimental experience. This finding seriously compromised the attempt to test the

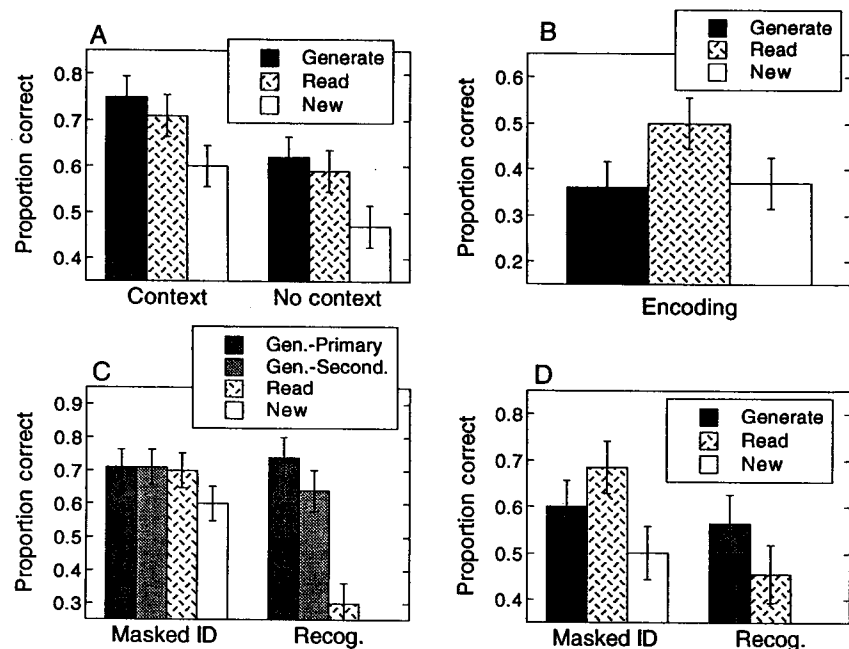


Fig. 1. Mean proportion correct on masked word identification and recognition memory tasks as a function of encoding task. (A) The generation cues were definitions and targets were tested in the context of a related word taken from the relevant definition or with no relevant context word. (B) The generation cues were antonyms. (C) The generation cues were definitions that specified either the dominant or the secondary meaning of the homograph target words. (D) The generation cues were idioms in which the target word served as the last word in the idiom. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from Masson and MacLeod (1992).

context reinstatement hypothesis. One other important but unexpected result emerged in this experiment. The amount of priming observed for generate targets was just as large as the priming found with read targets. Thus, unlike Jacoby (1983), we found no advantage on the masked word identification test for targets that had been encoded perceptually, relative to targets that had been encoded without being perceptually processed.

Given the unexpected anomalous finding of equal priming for generate and read targets, we examined the implications of that result rather than the question of context reinstatement. Our concern was that the relatively weak amount of priming found with the generate encoding task in the Jacoby (1983) study might not be a general result. It was important, therefore, to replicate the finding of greater priming in the read condition. To

do so, we turned to the materials used by Jacoby, which consisted of antonyms as generation cues. In our next experiment (Masson & MacLeod, 1992, Experiment 2), then, we adapted these materials by presenting a target word's antonym and first letter as generation cues (e.g., true—*f?*).

In the study phase, subjects read some targets and generated others from their antonym cues. The masked word identification task was then presented with three sets of target words: generate, read, and new. The mean proportion of identified words in each study condition is shown in Fig. 1B. The results clearly replicated Jacoby's (1983) finding of little or no priming when targets were previously generated from antonym cues. This result is important because it rules out uninteresting explanations of the difference between the results of our first experiment and the results obtained by Jacoby. The primary difference between these two studies is the nature of the materials.

In subsequent experiments in the Masson and MacLeod (1992) article, we tested various kinds of generation cues in an effort to discover the factor that determined whether generating targets would yield as much priming as reading them. Most of the generation cues we examined produced as much priming for generate words as for read words (e.g., synonym and associate cues, *hammer—n?*; famous last names as cues for first names, *M. Twain*). For instance, in Experiment 5, subjects generated homographs from phrases that specified either the primary or a secondary meaning of the homograph. To help constrain responses to the intended target word, additional letters besides the first letter of the target were provided (e.g., "the sprinter got ready for her *r\_c\_*"; "we are all part of the human *r\_c\_*"). Subjects read or generated from one of its possible cues each of a set of homograph target words. Another set of homographs was not presented during the study phase but appeared as new words in the test phase. Subjects then were given a masked word identification test in which studied and nonstudied targets were presented in isolation, as in the earlier experiments. Finally, subjects were given a yes/no recognition test in which each of the targets from the identification test was presented. The task was to indicate whether a target had been presented during the study phase. This recognition test was partly a source discrimination test because all of the new items had appeared on the immediately preceding masked word identification test (although not all had been identified).

The proportions of correctly identified words in each encoding condition of this experiment are shown in Fig. 1C. Once again, generated targets produced as much priming as read targets. Note also that it did not matter whether a homograph was generated in the context of a cue that specified its primary or secondary meaning. In either case, the same amount of priming was observed on the identification task. This lack of a dominance



effect contrasts with the results of the recognition test, which are also shown in Fig. 1C. On that test, subjects were more likely to correctly recognize a generated target when it had been generated from a cue that specified its primary meaning.

The dissociation between the masked word identification test and the recognition test revealed by the meaning dominance factor suggests that subjects were not relying to any substantial degree on conscious recollection when performing the identification test (see Experiment Series 3 for further discussion of this issue). Had conscious recollection played an important role on the identification test, we should have seen an effect of meaning dominance among the generated targets just as was found on the recognition test. Instead, the absence of the dominance effect on the identification test indicates that the brief presentation of a target is capable of recruiting a recent episode involving that word, regardless of which of the target's meanings was invoked by that episode. This possibility is consistent with the proposal that all meanings of a homograph are initially retrieved before one of its meanings is selected (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). We suggest that it is the early stages of word identification that are captured by the masked word identification task and that prior processing of either meaning of a homograph can be contacted by the brief presentation of the target word. In contrast to the masked word identification task, we assume that decisions on the recognition test are based for the most part on a selected meaning of the homograph and most typically the selected meaning is the dominant one. Thus, targets generated from a cue that specifies the dominant meaning are more likely to be recognized (for a related finding, see Light & Carter-Sobell, 1970).

The demonstration that various generation tasks produce as much priming as reading a target word forces the question as to why the antonym task fails to do so. In Masson and MacLeod (1992), we suggested that one of the factors governing the amount of priming created by generation encoding tasks is the extent to which the generation cue and the target word are integrated during encoding. Our hypothesis was that if cue and target are integrated at encoding, then the brief presentation of the target word on the masked word identification test would retrieve an integrated record of that encoding event, making it difficult to discriminate between the cue and the target word. In support of this proposal, we demonstrated that other generation cues that were assumed to lead to strong integration of cue and target behaved like antonym cues. That is, generation of target words from these cues yielded less priming on the masked word identification test than did the read encoding task.

An example of this result occurred in Experiment 6 of Masson and MacLeod (1992), in which generation cues were idiomatic expressions and

targets were the final words of those expressions (e.g., "kick the *b*\_\_\_\_\_"). This experiment used the same procedure as Experiment 5, including a yes/no recognition test following the masked word identification test. The mean proportion of identified targets in each encoding condition is shown in Fig. 1D. Although the generate task produced reliable priming compared to new targets, it was significantly less than the priming found in the read condition. The opposite pattern was obtained in the recognition test, in which a higher proportion of correct responses was given to generate than to read items. This crossed dissociation replicates the pattern found with the antonym generation task (Jacoby, 1983; Masson & MacLeod, 1992, Experiment 2).

Taken together, the series of experiments reported in Masson and MacLeod (1992) provided evidence to challenge the view that the masked word identification task reflects exclusively the effects of prior perceptual processing episodes. More generally, the experiments illustrated the plausibility of the claim that fluent word identification performance is sensitive to components of processing episodes that are responsible for constructing an interpretation of a presented word. These components include not only the perceptual aspects of the episode, but also the conceptual processes that are required for the construction of a meaningful interpretation of the stimulus.

### III. Experiment Series 2: Comparing Masked Word Identification and Word Fragment Completion

Given the masked word identification studies described in Experiment Series 1 and their apparent contrast to other studies using such implicit tests as word fragment completion (see, e.g., Roediger & McDermott, 1993), we suspected that different implicit tests could behave quite differently. Witherspoon and Moscovitch (1989) had already shown that word fragment completion and masked word identification tests were stochastically independent under a variety of conditions. On this basis, they concluded that it is the overlap in the processes used, as well as differences in the information that they summon from memory, that differentiates memory tests. But there was also a significant conflict in the literature in the form of a study by Weldon (1991, Experiment 1).

Weldon directly compared the masked word identification and fragment completion tests in her first experiment and found a gradient of priming for both tests, with the read condition showing more priming than the generate condition. Had her materials comprising the generation clues been antonyms, we would have expected this pattern (Jacoby, 1983; Masson &

MacLeod, 1992, Experiment 2). But she had used short definitional phrases, materials that we had shown (Masson & MacLeod, 1992, Experiments 1 and 7) produced equal priming for read and generate words in masked word identification. Therefore, we set out to determine the cause of these incongruent results.

In a set of experiments (MacLeod & Masson, 1997b) using the procedures described in Experiment Series 1, we set out to compare directly the masked word identification and fragment completion tests under similar conditions. In our first experiment, we used Weldon's (1991) materials to replicate Experiment 1 of Masson and MacLeod (1992). We noted that Weldon had required her subjects to respond to every test item in masked word identification (forced report), whereas we had permitted passing (free report), so we tested two groups of subjects, one with each type of report instruction. Our concern was that criterion differences might underlie the different patterns of results. But that was not the answer to the puzzle. As can be seen in Fig. 2A, both versions of the experiment produced results just like those of Masson and MacLeod's Experiment 1: Performance for the two encoding conditions was equivalent in the free-report and forced-report conditions. Thus, we replicated Masson and MacLeod's finding, but failed to replicate Weldon's.

A disconcerting possibility is that somehow subtle differences in the procedures being used in our respective laboratories were the source of the discrepancy in results. Because we also generally used different tasks, fragment completion in the case of Weldon and Roediger and masked word identification in our case, what looked like a task difference might actually be caused by some entirely different factor. In a sense, there was "their pattern," with read > generate in amount of priming, and "our pattern," with read = generate in amount of priming. The best way to discount this possibility was to carry out both tests in a single laboratory under identical conditions. When we did this (MacLeod & Masson, 1997b, Experiments 2 and 3), we obtained different patterns for the two tests. As indicated in Fig. 2B, masked word identification showed equal priming for read and generate encoding tasks, regardless of whether encoding task was manipulated within or between subjects. For fragment completion, shown in Fig. 2C, there was reliably more priming in the read condition than in the generate condition, both for within and between subject manipulations of encoding task. Thus, we consistently produced the read advantage for the fragment completion test and equivalent read and generate priming for the masked word identification test. Unfortunately, these results do not explain Weldon's (1991, Experiment 1) finding of a read advantage for priming in the masked word identification test.

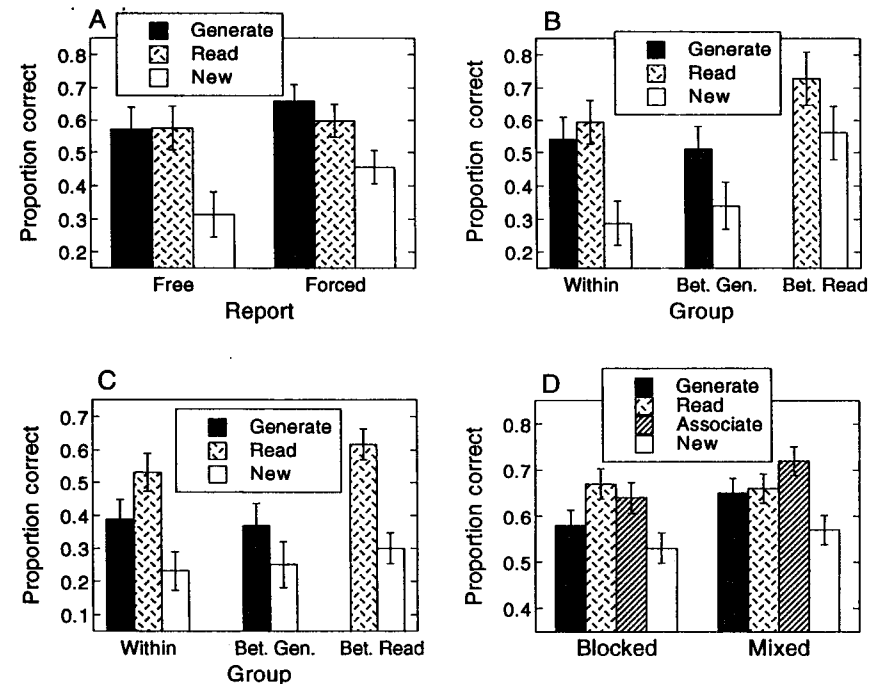


Fig. 2. Mean proportion correct on masked word identification and word fragment completion tasks as a function of encoding task. (A) Masked word identification performance under free and forced report instructions. (B) Masked word identification with encoding task manipulated within or between subjects. (C) Word fragment completion with encoding task manipulated within or between subjects. (D) Masked word identification with encoding tasks presented in blocked or mixed format. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from MacLeod and Masson (1997b).

We next carefully examined Weldon's (1991) procedure to see whether there was some element that discriminated her experimental approach from ours. We found two differences. First, Weldon used blocked presentation of read and generate trials during encoding, whereas we had used mixed presentation. Second, she used four encoding conditions (the additional ones being auditory presentation and pictorial presentation), whereas we had always used only two encoding conditions. It seemed most plausible to us that blocking was the potent factor, with mixing perhaps encouraging a kind of "leakage" in encoding between read trials and generate trials (e.g., subjects might have tried to form a mental image of a generated target's orthography). So, we conducted an experiment in which we blocked



the presentation of read and generate trials during encoding (MacLeod & Masson, 1997b, Experiment 4). In the subsequent masked word identification test, we once again found a pattern of equivalent priming for read (.28) and generate (.28) encoding conditions, almost perfectly reproducing the free-report condition in Experiment 1 of this series (priming in read = .26; generate = .27). These results showed that blocking alone was not the answer.

Although it was hard to accept, we decided that the critical manipulation might not be a main effect, but rather could be an interaction. Perhaps it was the combination of multiple encoding conditions with blocking that caused Weldon's (1991, Experiment 1) pattern for masked word identification to diverge from ours for the same task (Masson & MacLeod, 1992). In Experiments 5 and 6, we test this possibility directly. To the standard read and generate conditions, we added a third encoding condition: Read the word and produce an associate to it (e.g., for *watermelon*, the subject would say "watermelon" and then an associate such as "seeds"). In Experiment 5, the three encoding conditions were blocked; in Experiment 6, they were mixed. Figure 2D displays the results. When the three encoding conditions were blocked, there was more priming for the read condition (.14) than for the generate condition (.05); when they were mixed, priming was equal (read = .09; generate = 0.8). The answer, then, appears to be that the use of more than two encoding conditions in blocked presentation during study produces the pattern of priming that was observed in Weldon's (1991, Experiment 1) masked word identification study.

To summarize, the word fragment completion test reveals a highly consistent pattern, with more priming for words that were read than for words that were generated during encoding. It is important to realize, however, that generated words do produce priming quite reliably in word fragment completion; they simply produce less priming than read words. The masked word identification task also produces a very consistent pattern: Ordinarily, there is equivalent priming for read and for generated words, perhaps helping to explain why Witherspoon and Moscovitch (1989) found fragment completion and mask word identification to be independent. But there do appear to be exceptions in the case of masked word identification: When the generation rule involves antonyms (Jacoby, 1983; Masson & MacLeod, 1992, Experiment 2), there is little or no priming for generated words, and when there are more than two encoding conditions and these are blocked, there is a gradient of priming, with more for read words than for generated words.

Taken together, these results convince us that these two implicit memory tests, certainly two of the most widely used tests in the implicit memory literature, ordinarily behave differently. We hypothesize that this has to

do with the different form of the test stimulus confronting subjects in the two tests. The whole word is presented in masked word identification; only part of the word is presented in word fragment completion. A whole word can summon both the perceptual and the conceptual aspects of the initial interpretive encoding, whereas a partial word primarily calls forth the perceptual elements of that encoding. Thus, the read condition is favored in fragment completion, which invokes a strategy of letter insertion and lexical search (cf. Nelson, Keelean, & Negrao, 1989, Experiment 4), but not in masked word identification, where access to both form and meaning is involved.

Finally, we must admit that we do not have a compelling explanation for the two exceptions to the general finding of equivalent priming in masked word identification. We suggested in Masson and MacLeod (1992) that generation from antonyms may lead to an especially integrated encoding of the clue and the target words that then suffers at the time of test when the clue is no longer presented with the target in the generate condition. This decoupling of context for the generate condition impairs identification of generated words relative to read words. Such a contextual-integrative explanation probably could be developed for the multiple blocked encoding conditions exception as well, but it would be decidedly post hoc at this point.

Instead of trying to explain both exceptions in the same way, it is important to note that it is the generate condition that is impaired in the antonym case, whereas the read condition shows somewhat less priming in the mixed encoding task situation, relative to the blocked encoding case (Fig. 2D). Perhaps, as we suggested in the MacLeod and Masson (1997b) article, subjects treat the words in the read condition in a more cursory, less conceptual fashion when they occur in a mixture of trials that includes other encoding conditions that emphasize conceptual processing. The solution to explaining these discrepancies from the general pattern remains to be seen, but at least we do know the standard patterns now. Moreover, we can see that the two implicit tests show different patterns of priming. Like Witherspoon and Moscovitch (1989), we take this as evidence that the tests invoke different collections of processes, consistent with a transfer-appropriate processing account. We have begun to suggest what those differences are, emphasizing through the idea of initial interpretive encoding the role played by conceptual processing in such indirect tests.

#### IV. Experiment Series 3: The Question of Conscious Recollection

One of the most contentious issues associated with studies of automatic or unconscious influences of memory is the question of whether task perfor-

mance is mediated at least in part by conscious recollection. In the quest for pure measures of unconscious influences, this problem has been referred to by such terms as "conscious contamination." The concern is that even in indirect tests of memory that do not instruct subjects to recollect specific prior events, the ability to consciously remember those events may in some way influence current task performance. If conscious recollection contributes to performance, any putative unconscious influence of memory is assumed to be contaminated.

A widely used technique for addressing the problem of conscious influences on indirect tests of memory is to establish dissociations between indirect and direct tests of memory. A dissociation consists of demonstrating that an independent variable affects one of these two types of test, but not the other. Preferably, this dissociation is established under conditions in which the retrieval cues are the same for both the direct and indirect test. This definition of task dissociation is known as the retrieval intentionality criterion (Schacter, Bowers, & Booker 1989). The only difference between the two tasks is the instruction to the subjects: On the direct test, subjects are instructed to intentionally retrieve items from a prior study context, whereas on the indirect test no such instruction is given. For example, in an indirect test of memory involving word stem completion, subjects would be instructed to complete each stem with the first word that comes to mind; no instruction to remember previously presented words is given. On the direct version of this test, the same cues would be provided, but subjects would be instructed to use the cues consciously to remember the words that were studied earlier. If a factor such as modality of the presentation of the words during study affects the probability of producing target completions in one of these tests but not the other, the preferred conclusion is that different processes determined performance on the two tasks. In particular, such dissociations are used to argue that conscious recollection did not affect the indirect test.

Although the establishment of task dissociations has played a significant role in the study of conscious and unconscious influences of memory, two important challenges to the task dissociation logic have been put forward. First, Dunn and Kirsner (1988) have shown that dissociations in which a variable affects performance on one test but has no effect on another test, or even a crossed double dissociation, in which a variable affects two tests in opposite ways, do not logically rule out the possibility that performance on the two tests is determined by a single process. If a single process can account for task dissociations, it is difficult to argue that one task involves a conscious influence of memory while the other involves an unconscious influence.

Second, Jacoby and his colleagues (Jacoby, 1991; Jacoby, Lindsay, & Toth, 1992; Jacoby, Toth, & Yonelinas, 1993) have argued that direct and indirect tests of memory are not process pure with respect to conscious and unconscious influences on test performance, even when task dissociations are established. To demonstrate this point, and to provide a means of estimating the contribution of conscious and unconscious influences of memory to task performance, Jacoby (1991) proposed the method of process dissociation. In this method, it is assumed that controlled (conscious) and automatic (unconscious) memory processes make independent contributions to task performance. A crucial aspect of the procedure involves placing controlled and automatic processes in opposition to one another, as explained below. By using this procedure to obtain estimates of controlled and automatic influences, it can be shown that these estimates are differentially affected by certain independent variables (e.g., divided vs full attention), thereby establishing process dissociations and supporting the assumption of independence (Jacoby, Yonelinas, & Jennings, 1996).

The process dissociation procedure is of particular interest here because it has been used to estimate quantitatively the automatic and controlled influences of memory for generate and read encoding tasks. More specifically, Toth, Reingold, and Jacoby (1994) found that the generate encoding task had no automatic memory influence on a word stem completion task. This result is a striking contrast to the findings we have described so far in this chapter and deserves careful consideration.

Toth et al. (1994, Experiment 2) used short definitions as generation cues, just as Masson and MacLeod (1992, Experiments 1 and 5) and MacLeod and Masson (1997b) have done. During a study phase, subjects generated some target words from their definition cues and read aloud other target words that were presented in isolation. In a subsequent word stem completion test, stems were tested under two different instructional conditions. In the inclusion condition, subjects were instructed to use the word stem as a recall cue in an attempt to remember the word from the study list that fit the stem. If a subject was unable to recall a previously studied word, he or she was to complete the stem with the first word that came to mind. On the assumption that controlled and automatic influences of memory operate independently to determine which word will be produced to complete a stem, the probability of completing a stem with the studied target under inclusion instructions can be expressed as

$$I = C + (1 - C)A,$$

where  $C$  is the probability of consciously recollecting the target completion

(controlled responding) and  $A$  is the probability of the target completion coming to mind automatically.

In the second instructional condition, subjects were given exclusion instructions, whereby they were told to attempt to recall the studied word that would fit the stem, but not to give the word as a completion. Instead, subjects were to give a different word as the completion for that stem. Again, assuming independence between controlled and automatic influences of memory for the studied words, the probability of completing a stem with a studied word in the exclusion condition can be expressed as

$$E = (1 - C)A.$$

According to this equation, a target completion will be given only if it comes to mind through an automatic influence of memory and is not consciously remembered.

By obtaining the probability of completing word stems with target words in the inclusion and exclusion conditions, Toth et al. (1994) were able to compute estimates of controlled and automatic influences of memory by solving the two process dissociation equations. The mean proportion of word stems completed under inclusion and exclusion instructions and the mean estimate of controlled and automatic influences of memory reported by Toth et al. are shown in Fig. 3A.

The inclusion and exclusion task performance is of interest primarily for the sake of producing estimates of controlled and automatic influences of memory. But one aspect of the completion probability data is important for another reason. The completion probabilities for new items were not reliably different in the inclusion and exclusion conditions. This result is important because it was put forward by Toth et al. as evidence that subjects were not engaging in a generate-recognize strategy during stem completion, particularly in the exclusion condition. The validity of the process dissociation equations depends on the assumption that controlled and automatic influences of memory operate in the same way under inclusion and exclusion instructions. That is why, for example, subjects are instructed in the exclusion condition to attempt to recall studied words (then not to give those words as completions), just as they are instructed in the inclusion condition.

There is concern, however, that subjects might operate according to a different agenda in the exclusion condition. In particular, they may rely on a generate-recognize strategy in which potential completions come to mind without the subject necessarily trying to remember words from the study phase. If the subject fails to recognize that a potential completion is a studied word, then he or she uses it to complete the stem. This characterization of performance in the exclusion condition is contrary to what is assumed to

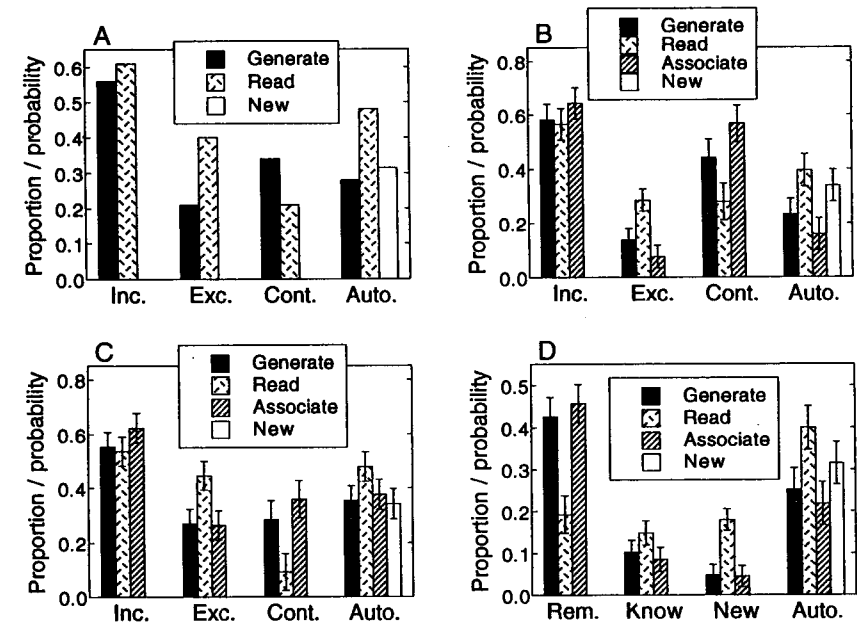


Fig. 3. Mean word stem completion and estimates of controlled and automatic influences of memory based on the process dissociation procedure as a function of encoding task. (A) Data are adapted from Toth et al. (1994, Experiment 2). (B) Replication of Toth et al. adding the associate encoding task. (C) Second replication of Toth et al. using exclusion task instructions intended to induce a high criterion for excluding previously studied targets. (D) Application of the independence remember/know procedure to estimate controlled and automatic influences of memory. Error bars indicate the 95% within-subject confidence intervals for the means. Data in (B), (C), and (D) are adapted from Bodner et al. (1997). Inc. = inclusion; Exc. = exclusion; Cont. = controlled; Auto. = automatic; Rem. = remember.

happen by the equation for exclusion task performance given earlier. Thus, it is important to show that subjects did not engage in this approach to the exclusion task.

Evidence that can be used in this cause consists of demonstrating that the probability of completing stems whose target completions were not studied with those nonstudied (new) targets is the same in the inclusion and exclusion conditions. The rationale is that if subjects had applied a generate-recognize strategy in the exclusion condition, they should be less likely to complete stems that correspond to new targets because of instances of false recognition. That is, some target completions in the new condition will falsely be classified as old and consequently withheld by the subject. This recognition check would not be invoked in the inclusion condition.

Finding lower completion performance in the exclusion condition than in the inclusion condition, then, would constitute evidence for the operation of a generate-recognize strategy in the exclusion condition. On the other hand, finding no difference between the completion rates for new items in these two conditions implies that subjects did not use the generate-recognize strategy. The completion rate for new items was not reliably different in the inclusion and exclusion conditions in the Toth et al. (1994) study, suggesting that subjects did not rely on a generate-recognize strategy.

The results shown in Fig. 3A indicate that the generate task led to greater control in responding than did the read task, as subjects were much more sensitive to the effect of inclusion/exclusion instructions in the case of generate items. This differential sensitivity translates into a higher estimate of conscious influences of memory in the generate condition, which is computed as

$$C = I - E.$$

The results in Fig. 3A clearly show that the conscious contribution of memory was estimated to be much larger following the generate encoding task than following the read task. In contrast, the estimates of automatic influences of memory showed the opposite pattern as a function of encoding task. The read encoding task produced a substantial estimate of automatic influences of memory, over and above the preexperimental influences as assessed by completion probabilities in the new condition. The generate condition, however, led to an automatic estimate that was no different from the completion probability for new items, indicating that there was no automatic influence of memory for the generate encoding episode.

Toth et al. (1994) concluded that the automatic influence of memory on the stem completion task operated through data-driven processes. Because the generate task did not engage data-driven processing of targets, no automatic influence of those encoding episodes was observed. Toth et al. also suggested that demonstrations of enhanced word identification performance on tasks such as masked word identification following generate encoding tasks (e.g., Masson & MacLeod, 1992) were due to conscious recollection of target words and not to automatic influences of memory for the encoding episodes.

We have two concerns regarding the conclusions reached by Toth et al. (1994). First, the generalization of their conclusion regarding the lack of automatic influences of memory for generate encoding episodes to word identification tasks other than stem completion is questionable. In the version of the stem completion task used by Toth et al., only the first three letters of a target word were provided. This kind of retrieval cue may

not provide enough constraints to permit the recruitment of conceptual knowledge about a target word, including episodic memory for conceptual processing operations performed during the generation task. In the masked word identification task, however, the target word is in full view, although for only a brief time. Evidence suggests that even under these conditions, the brief exposure of the entire word is capable of making contact with conceptual knowledge associated with that word. For example, semantic or associative priming effects have been found with brief, masked presentations of primes (Carr & Dagenbach, 1990; Dagenbach, Carr, & Wilhelmsen, 1989; de Groot, 1983; Lukatela & Turvey, 1994). Even if it is the case that stem completion does not present an opportunity for memory for conceptually driven processing episodes to influence task performance automatically, the masked word identification task may do so.

Our second concern is that the estimate of automatic influences of memory following the generate encoding task, as determined by the process dissociation procedure, may not be accurate. Various criticisms of the procedure and the estimates it yields have been raised, particularly with respect to the assumption that automatic and controlled contributions of memory operate independently (e.g., Curran & Hintzman, 1995; Joordens & Merikle, 1993). Our concern is not with the question of independent processes and we are willing not to question that assumption for the time being. Rather, we suspect that the estimate of automatic influences of memory is underestimated by the process dissociation procedure, particularly when assessing the effect of conceptually driven encoding tasks, because of what Richardson-Klavehn and colleagues have referred to as involuntary conscious memory (e.g., Richardson-Klavehn & Gardiner, 1995; Richardson-Klavehn, Gardiner, & Java, 1996).

Involuntary conscious memory involves two components. The first is an automatic influence of memory, perhaps causing a potential word stem completion to come to mind. The second component is a subsequent awareness that that word had appeared in the study phase of the experiment. Should this process occur during word stem completion under exclusion instructions, subjects very likely will withhold that word and seek an alternative completion. Notice that by the usual meaning of an automatic influence of memory, generation of a completion under these circumstances is a valid instance of an automatic influence of memory. The subsequent involuntary awareness of that word's prior occurrence, however, is alleged to mask that influence. The result is an underestimate of the true automatic influence of memory. Moreover, the effect of involuntary conscious memory is likely to be stronger for words encoded with conceptually driven study tasks because these tasks most strongly promote conscious recollection.

Although Toth, Reingold, and Jacoby (1995; Reingold & Toth, 1996) have argued that there is no clear evidence that involuntary conscious memory has a substantial influence on estimates of automatic influences of memory obtained with the process dissociation procedure, we have recently conducted a series of experiments that justifies this concern (Bodner, Caldwell, & Masson, 1997). Our first experiment was a replication and extension of the Toth et al. (1994, Experiment 2) study in which generate and read encoding tasks were compared using a word stem completion task. We extended that original study by including a third encoding task that was expected to be revealing with respect to the purported influence of involuntary conscious memory. In that condition, called the associate condition, subjects read aloud a target word just as in the read task, but then also reported the first word that came to mind upon reading the target word.

It was intended for the associate task to provide subjects with a data-driven processing experience identical to that in the read condition. Furthermore, the production of an additional, probably related, word was expected to add a conceptually driven processing component to the encoding task. Our expectation was that if process dissociation procedure estimates are unaffected by involuntary conscious memory, then the associate and read encoding tasks should lead to equivalent estimates of automatic influences of memory. On the other hand, if the estimate of automatic influences can be artifactually lowered by involuntary conscious memory, because of its conceptual component the associate condition would be more strongly subject to that influence than would the read condition. As a result, the associate condition should produce a lower estimate of automatic influences, perhaps as low as that found in the generate condition whose estimated automatic influence also is assumed to be powerfully affected by involuntary conscious memory.

Although in our first experiment we used inclusion and exclusion instructions similar to those used by Toth et al. (1994), we made one modification to the exclusion instructions. Whereas the exclusion task instructions used by Toth et al. appear to be vague with respect to whether the subject is permitted knowingly to use a studied word to complete a stem, we directly informed subjects that they were not to use previously studied words as completions. The mean proportion of stems completed with target words under inclusion and exclusion instructions and the corresponding estimates of controlled and automatic influences of memory are shown in Fig. 3B. Performance under inclusion instructions for items in the generate and read encoding conditions was similar to that reported by Toth et al., and items in the associate condition led to the highest proportion of target completions. This result is consistent with the idea that the associate condi-

tion bestowed the advantages of both conceptually driven and data-driven components upon the encoding episode.

Somewhat to our surprise, however, we were unable to replicate the results reported by Toth et al. (1994, Experiment 2) under exclusion instructions. Instead, we found that subjects rarely used target words as stem completions, even in the read and associate conditions, which should have produced a strong automatic influence of memory. The success of subjects in our experiment regarding exclusion of target completions might be taken as evidence that they used a generate-recognize strategy when under exclusion instructions. Were that the case, application of the process dissociation equations would be rendered invalid. The evidence on this question is suggestive. First, the completion rate for new items was somewhat different in the inclusion (.36) and exclusion (.32) conditions, a difference that approached statistical significance. Second, the estimates of automatic influence of memory in the generate and associate conditions were both reliably lower than performance in the new condition. That latter result is anomalous in the sense that valid estimates of automatic influences should not drop below baseline. These anomalous estimates suggest that subjects were relying on a generate-recognize strategy.

A similar result was obtained by Richardson-Klavehn, Gardiner, and Ramponi (1996) in a replication of a different experiment from Toth et al. (1994), in which level of processing in the study phase was manipulated. Richardson-Klavehn et al. used exclusion instructions in which subjects were instructed to use the stems to recollect studied words but not to give those words as completions for the stems. Subjects were also clearly told that they should give no response in the exclusion condition if only a studied word completion came to mind. Richardson-Klavehn et al. found that stem completion performance under exclusion instructions was very low, just as did Bodner et al. (1997). They concluded that the higher completion rate observed by Toth et al. (1994, Experiment 1) under exclusion instructions was a result of subjects not consistently following instructions.

In a second experiment, we persevered in an attempt to replicate the Toth et al. (1994, Experiment 2) results, while including the associate condition. We hypothesized that subjects in the Toth et al. study may not have been as thorough in withholding studied words as completions under exclusion instructions because they had set a higher criterion for deciding that a potential completion was an old word. Further, we supposed that under exclusion instructions, subjects would be unlikely to seriously attempt direct retrieval of old words using word stems as cues. Rather, we considered it likely that subjects would generally follow a generate-recognize strategy in which potential completions would be rejected if they seemed to be old words. Therefore, in Experiment 2 of Bodner et al. (1997), we tested the

possibility that the results reported by Toth et al. were the product of a generate–recognize strategy in the exclusion condition, combined with a high criterion for deciding that a completion was an old word. We used exclusion instructions that told subjects to provide the first completion that came to mind, unless they were sure that word was an old one. If a potential completion was considered to be an old word, a different completion was to be provided.

We expected that if the Toth et al. (1994, Experiment 2) results were due to undetected use of a generate–recognize strategy, our Experiment 2 should replicate the pattern of data obtained by Toth et al. A successful replication, however, could also arise if, despite providing instructions that encourage a generate–recognize strategy, subjects fail to adopt that strategy. The associate encoding condition plays an important role in resolving this potential ambiguity. If subjects operate as assumed by the process dissociation equations, rejecting under exclusion instructions only those completions obtained through direct retrieval, the estimate of automatic influences of memory should be the same for the associate and read conditions. In contrast, if subjects actually use a generate–recognize strategy with a high recognition criterion, the estimate of automatic influences in the associate condition should be reduced relative to the read condition. At the same time, the other markers of the generate–recognize strategy would be absent (i.e., differential completion of new items across inclusion and exclusion instructions, automatic estimates falling below baseline) because the high recognition criterion would prevent exclusion scores from approaching zero and would make rejection of completions for new items unlikely. This outcome would suggest that the Toth et al. results were also due to an undetected influence of involuntary conscious memory operating under exclusion instructions.

The results of our second experiment are shown in Fig. 3C. Performance under inclusion and exclusion instructions in the generate and read encoding conditions was very similar to that reported by Toth et al. (1994, Experiment 2). The estimates of controlled and automatic influences of memory also replicated the pattern found by Toth et al. There was a higher estimate of controlled processing for the generate than for the read condition, but the reverse was true for the automatic estimate. Despite the use of generate–recognize instructions, the data from the generate and read conditions showed no evidence that the boundary conditions for applying the process dissociation equations had been violated. The stem completion rates for new items under inclusion and exclusion instructions were not significantly different from each other and the automatic estimate for the generate condition was not significantly different from the completion rate for new items.

Although the generate and read encoding conditions replicated the Toth et al. (1994) results, the associate condition revealed a set of estimates that clearly indicates subjects were influenced by involuntary conscious memory. The controlled estimate for the associate condition was higher than for either generate or read, indicating that subjects consciously remembered a substantial number of words encoded in that condition. More important, the automatic estimate for the associate condition was significantly lower than the estimate for the read condition, and not different from the generate estimate or from the completion rate for new items.

Thus, despite applying the same data-driven encoding operations as in the read condition, the associate condition failed to produce evidence for an automatic influence of memory. This result supports the hypothesis that encoding tasks that involve substantial conceptually driven processing, such as the generate and associate tasks used here, are subject to much larger effects of involuntary conscious recollection on later indirect tests of memory than are encoding tasks that do not involve such processing. Even though a subject may experience a word completion as coming to mind automatically, without deliberate recollection, he or she may subsequently recognize that word as having appeared in the study phase and therefore withhold it when completing its stem under exclusion instructions. This situation creates a serious problem when applying the process dissociation procedure and equations to obtain estimates of automatic influences of memory. In particular, automatic influences can be greatly underestimated for encoding tasks that include a strong conceptually driven processing component.

The results of the second experiment of Bodner et al. (1997) may be criticized on the grounds that subjects were induced to use a generate–recognize strategy in the exclusion condition and therefore it was inappropriate to apply the process dissociation equations to obtain estimates of automatic and controlled influences of memory. Our counterargument is that although the instructions we used encouraged such a strategy, none of the behavioral markers of that strategy was apparent in the data, just as was the case in Toth et al. (1994). We suggest that the Toth et al. results may have arisen from undetected application of a generate–recognize strategy. Consequently, the parameter estimates they obtained are no more valid than those obtained in our second experiment.

We sought to reinforce our argument in a third experiment by using a procedure that does not involve testing subjects under exclusion instructions. Jacoby et al. (1996) developed an alternative technique for estimating the independent contributions of automatic and controlled processes in remembering prior events based on the remember/know distinction developed by Tulving (1985) and Gardiner (1988). In its original formulation,



remember/know judgments were applied during a recognition memory test in which subjects were asked to classify into two categories items they recognized as having been studied previously: those for which some aspect of the original encoding episode could be remembered and those for which no such memory was available but the subject simply "knew" that the item had been studied.

In the procedure developed by Jacoby et al. (1996), called the independence remember/know (IRK) procedure, subjects are given direct retrieval instructions, like those used in the inclusion component of the standard process dissociation procedure. Subjects are told that if they cannot remember a word from the study list, they are to complete the stem with the first word that comes to mind. Once a completion is provided, the subject then classifies the completion into one of three categories: remember (R), know (K), and new (N). These categories are defined as in the remember/know procedure. To obtain estimates of controlled and automatic influences of memory, the proportion of remember responses is treated as a pure estimate of controlled influences of memory (i.e.,  $C = R$ ). The estimate of automatic influences is derived by combining K and N responses, and dividing by  $1 - R$  (i.e.,  $A = [K + N]/[1 - R]$ ). The relation between these equations and the usual process dissociation equations is readily seen if one takes the combination of all three categories as an approximation to inclusion performance ( $R + K + N$ ) and the combination of know and new categories as an approximation to exclusion performance ( $K + N$ ).

In the third experiment of Bodner et al. (1997), subjects were tested using the IRK procedure following a study phase in which words were encoded using the generate, read, and associate tasks as before. We expected that the estimates of controlled and automatic influences of memory would follow the pattern seen in the second experiment, in which generate and associate encoding conditions would yield high controlled estimates but very low automatic estimates that are not different from the baseline stem completion rate. The proportions of stems that were completed and subsequently assigned to each of the three response categories are shown in Fig. 3D. The generate and associate conditions produced a similar profile of responding across the three categories, with a much higher proportion of stems completed and assigned to the remember category than was the case for the read condition. In contrast, items in the read condition were more likely to be completed and assigned to either the know or new categories than was the case for items in the generate and associate conditions.

The corresponding estimates of controlled and automatic influences of memory are also shown in Fig. 3D. The pattern of these estimates replicated the results from our second experiment. The controlled estimates for the generate and associate conditions were substantially higher than for the

read condition, but the reverse was true for the automatic estimate. The automatic estimate for the associate condition was even significantly lower than the completion rate for new items. This outcome was primarily due to a few subjects assigning all completions in the associate condition to the remember category, thereby producing an estimate of zero for the automatic parameter. Exclusion of these subjects from the analysis produced an automatic estimate of .24, which was not significantly different from the completion rate for new items.

The experiments using the process dissociation procedure were successful in replicating the original Toth et al. (1994, Experiment 2) result, which has been taken as evidence that the generate encoding task does not lead to any automatic influence of memory on stem completion. But these experiments also call into question the validity of that conclusion because they demonstrate that another encoding task, one that has both conceptually driven and data-driven processing components, shows the same pattern of estimates of automatic influences as the generation task. This result is paradoxical on the view that automatic influences of memory on the stem completion task are based on data-driven processes. The paradox is readily resolved, however, by the observation that estimates of automatic influences produced by the process dissociation procedure can be seriously underestimated because of involuntary conscious recollection.

We argue that when operating under exclusion instructions or when assigning stem completions to remember, know, or new categories, subjects often have the experience of a completion automatically coming to mind, followed by awareness that the completion had appeared in the study phase. This involuntary conscious remembering of prior episodes is systematically excluded from estimates of automatic influences of memory by the process dissociation procedure. The exclusion of such instances from consideration as examples of automatic influences of memory leads to the observed underestimate of those influences. The estimates of automatic influences of memory derived from the process dissociation procedure are more properly considered estimates of involuntary memory influences that are not accompanied by awareness. Encoding tasks that vary in the amount of conceptually driven processing they require will quite naturally lead to differential estimates of this subclass of automatic influence; greater conceptually driven processing at encoding is going to reduce the likelihood that subsequent involuntary influences of memory will go forward without themselves prompting awareness of the relevant prior experience.

Awareness of prior occurrence that arises after the automatic or involuntary influence of memory has occurred is most often not regarded as grounds for ascribing that influence to conscious recollection. Rather, by most accounts, these instances of involuntary conscious memory are considered

valid examples of an automatic influence of memory on task performance. The difficulty encountered in applying the process dissociation procedure is that these instances are inadvertently attributed to conscious recollection. Application of the procedure therefore creates the likely possibility that the automatic influence of memory for conceptually driven encoding episodes, such as those created by the generation task, are substantially underestimated.

Although the pattern of estimates obtained by Toth et al. (1994) fits well with the transfer-appropriate processing view, at least under the assumption that the stem completion task relies exclusively on data-driven processing, the Bodner et al. (1997) results cast doubt on those estimates. Those results do not, however, cause a problem for the general transfer-appropriate processing framework. Rather, we see the results as fitting with that framework and as constituting a warning about the dangers of classifying memory tests as exclusively or even predominantly determined by either data-driven or conceptually driven processes. We suggest that in many indirect tests of memory that involve identification of a target stimulus, both types of process are invoked, and that prior memory episodes involving either process in the course of identifying a target can enhance subsequent identification of that same item.

#### V. Experiment Series 4: Speeded Word Reading As an Indirect Measure of Memory

Aside from our work with the process dissociation procedure, we have pursued another approach to address the question of whether repetition priming effects we have found on word identification tasks following a generation episode could be the result of conscious recollection. Consider first how conscious recollection could affect performance in a task such as masked word identification. When subjects have difficulty identifying a masked word, they may resort to trying to remember words from the study phase. Given the much better conscious recollection of generated words as compared to read words (Slamecka & Graf, 1978), subjects are more likely to succeed in recollecting generated words, which then augment performance in the generate condition. It is this augmentation that leads to equal amounts of priming in the read and generate conditions. On this account, were it not for that extraneous influence, substantially more priming would be observed in the read condition than in the generate condition; indeed, the generate condition might produce no priming at all (e.g., Toth et al., 1994).

In Experiment Series 4 (MacLeod & Masson, 1997a), we tested this proposal by moving to a word identification task that would be very unlikely to engage conscious recollection processes. Our goal was to select a task for which interpretive encoding would be the dominant process, just as we propose is the case for the masked word identification task, then to use this task to confirm the equivalent priming we saw in masked word identification for read versus generated words. In so doing, we would both generalize the findings we described in Experiment Series 1 and 2 and help to dispel concerns that the pattern we observed hinged on conscious recollection selectively assisting the generate condition. Ideally, this new task would be one in which the test word would be presented without any form of degradation, and where its processing would be rapid and seemingly effortless.

A near perfect candidate to meet these criteria is speeded oral word reading, a task often called "naming" in the literature (e.g., Balota & Chumbley, 1984; Forster & Chambers, 1973). The subject's task is simply to read aloud as quickly as possible a clearly presented word, and the dependent measure is latency to read. There is no problem solving, and the task is certainly not difficult. In fact, based on the interference that occurs in color naming in the Stroop (1935) task when the subject must say "red" to the stimulus word *green* printed in red, it is frequently held that word reading is automatic, in the sense that it requires little attention (see MacLeod, 1991, for a review of the Stroop literature). Word reading thus seems to be an ideal implicit measure of memory, a claim supported by existing research (e.g., Carr, Brown, & Charalambous, 1989; Logan, 1990). Word reading should not promote conscious recollection both because the task is extremely easy and because it is performed very quickly.

With this logic in mind, we set out to extend our studies of masked word identification to speeded oral word reading (MacLeod & Masson, 1997a). These new experiments adopted the same study procedure as in the masked word identification experiments, but substituted a word reading test as the indirect test of memory. Note that because the measure is now latency to read words, priming will now appear in the form of shorter bars in the studied conditions. The first experiment in this series used the definition materials that Weldon (1991) had used, and that we had borrowed for our Experiment Series 2 (MacLeod & Masson, 1997b). As can be seen in Fig. 4A, there was reliable priming for both read and generate words, and that priming did not reliably differ across those conditions, precisely the pattern we have seen previously for masked word identification.

When we switched to antonyms as our materials in Experiment 2, we also found equivalent priming for read words and generated words that was similar in magnitude to the priming observed with definitions (see

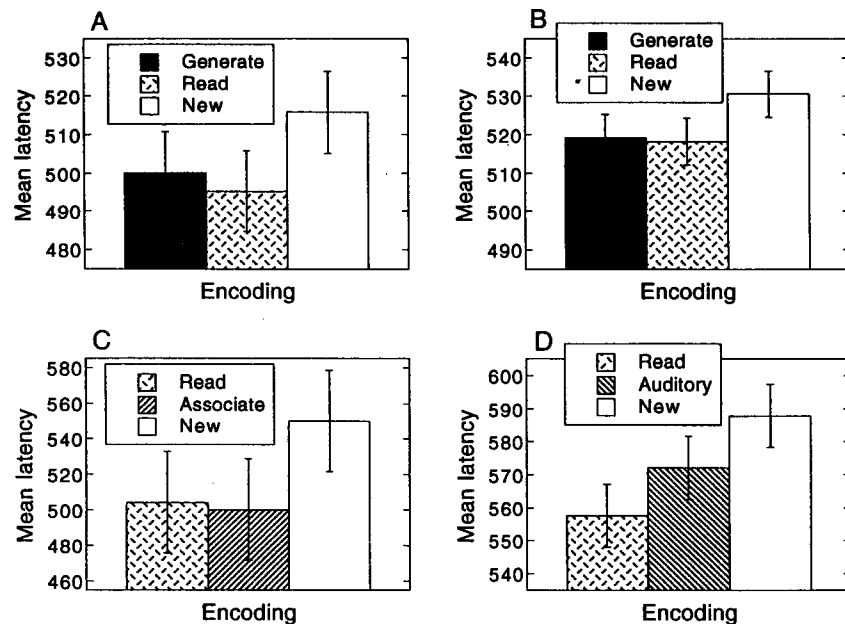


Fig. 4. Mean speeded word reading time as a function of encoding task. The generation cues were (A) definitions or (B) antonyms. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from MacLeod and Masson (1997a).

Fig. 4B). Intriguingly, then, even antonyms, which showed no priming for generated words in masked word identification, now showed equivalent priming whether they had been read or generated during encoding. In Experiment Series 1, where masked word identification was the test (Masson & MacLeod, 1992), we explained the absence of priming for words generated from antonym clues as possibly being due to the integrated encoding of clue and target, which hampered the target when it stood alone at the time of test. Perhaps the automaticity of speeded word reading minimizes the relevance of this integration at encoding because now retrieval is trivially easy because the target word is clearly in view.

In the third experiment in the speeded word reading series, we removed the generate condition and replaced it with a condition we had used in Experiment Series 2 and 3: the associate condition, in which subjects read the target then generated an associate. The rationale for making this substitution was that now the subject would have the same perceptual experience in both encoding conditions, but more apparent emphasis on conceptual encoding in the associate condition. If we are correct, however, that both

perceptual and conceptual processing routinely contribute to the initial interpretive encoding of each studied word anyway, then this change should not alter the equivalence of priming. Indeed, as shown in Fig. 4C, we obtained similar amounts of priming for the associate condition (50 ms) and the read only condition (46 ms), further generalizing the finding.

Given the consistency with which these experiments produced equal amounts of priming across various study conditions, one might doubt that speeded word reading is a good measure of priming, in that it perhaps resonates consistently to any prior experience with an item. To address this possibility, we turned to one of the manipulations that has been most reliable in producing differential repetition priming effects, that of modality (e.g., Bassili, Smith, & MacLeod, 1989; Roediger & Blaxton, 1987; see Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993, for reviews). It has been shown repeatedly that changing the modality (e.g., from auditory to visual) between study and test reduces priming sharply and sometimes eliminates it. This effect may have to do with the decreased contribution of perceptual aspects of the encoding when modality shifts. In any event, modality change virtually always decreases priming.

We chose to manipulate modality in our paradigm to determine whether the speeded word reading measure would be sensitive to a manipulation that routinely affects other implicit tests. Subjects either saw a word or heard it during study, with the auditory and visual encoding conditions blocked for ease of administration. As Fig. 4D shows, the modality manipulation was successful in producing differential priming. With the word reading test involving visual presentation of the targets, words studied in the visual modality produced significantly more (about twice as much) priming as those studied in the auditory modality, though auditorily presented words still produced reliable priming (as in many previous studies, such as Bassili et al., 1989). Word reading is, therefore, a sensitive measure of implicit remembering.

More recently, one further result has been obtained to complement the findings of Experiment Series 4. MacLeod and Daniels (1997) manipulated two variables at encoding prior to a speeded word reading implicit test. One of these was the now familiar read versus generate manipulation. Factorially combined with this was a manipulation of whether each word was followed by an instruction to remember the word for a later memory test or to forget it. This latter manipulation, called directed forgetting, is known to have a powerful effect on explicit remembering (for a review, see MacLeod, in press). The materials were the Weldon (1991) definition items of Series 2. The question was how these two variables would affect explicit remembering (free recall) as opposed to implicit remembering

(speeded word reading). Figures 5A and 5B provide the answer. Figure 5A shows that directed forgetting had a strong effect on recall of words read during study, but not on words generated during study. Apparently directed forgetting cannot affect an item that is already richly encoded. But for present purposes, the more crucial result concerns the speeded word reading implicit measure. Once again, Fig. 5B shows that all studied conditions led to priming, but that there was no differential priming as a function of how the word was encoded. Neither manipulation, despite its renowned strong impact on explicit measures, had any impact on priming in speeded word reading.

Experiment Series 4 (MacLeod & Masson, 1997a) plus the MacLeod and Daniels (1997) study show us that the masked word identification pattern is a general one. At least one other implicit measure, speeded word reading, shows the same equivalence in priming for read and generated words as does masked word identification. Furthermore, speeded word reading would appear to be a very good implicit measure in that the task does not foster a problem-solving set that in turn could lead to conscious recollection in the face of failure to identify a particular target word. Instead, the task is easily and quickly performed, perhaps automatically. It is even conceivable that conscious recollection would interfere with the smooth performance of word reading, slowing it down. Yet, latencies in our experiments are very similar to those seen in typical word reading studies, and certainly do not seem unusually elevated. These results make arguing for intrusion of conscious recollection into performance under the generate condition much less feasible.

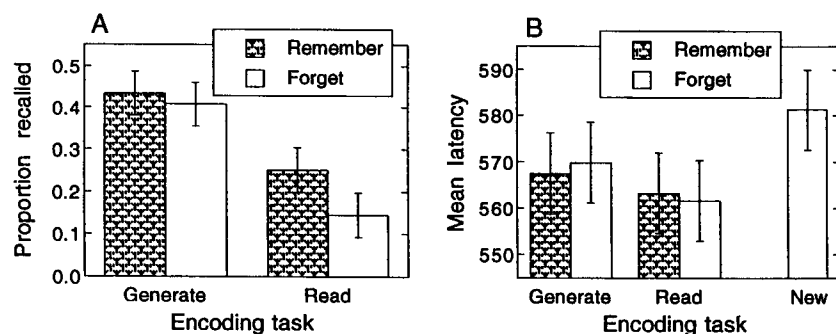


Fig. 5. Mean proportion of words correctly recalled and mean speeded word reading latency following remember and forget instructions in a directed forgetting paradigm. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from MacLeod and Daniels (1977).

## VI. Experiment Series 5: Color Naming versus Word Reading and the Specificity of Priming

Our view of the encoding–retrieval interaction that constitutes remembering is rooted in the idea of transfer-appropriate processing (Morris et al., 1977; Roediger, 1990). For transfer to be appropriate, there must be at least some correspondence between the processes engaged by the initial episode and those engaged by the subsequent episode. We are reminded of Osgood's (1949) "transfer surface" and the notion that transfer varies as a function of a number of factors relating to the stimuli and responses. In our case, the "surface" of transfer would be dictated not by the stimuli and responses, however, but by the processes involved. An obvious question concerns just how related the processes on the two occasions must be for transfer to occur.

In the series of experiments to be described in this section (MacLeod, 1996), two different implicit tests were compared. The first test was speeded word reading, familiar from Experiment Series 4. The second was color naming, modeled after the classic Stroop (1935) task. Color naming has previously been used to measure semantic priming in studies by Warren (1972) and others (for a review, see MacLeod, 1991, pp. 173–174). Warren showed subjects three words from a category (e.g., aunt, uncle, cousin) and then had them name the print color of a word that could be from either the presented set (e.g., uncle), their superordinate (e.g., relative), or an unrelated control (e.g., horse). He showed that subjects were slowed in naming the color of print when the meaning of the target word printed in color had been primed by immediately prior study.

Initially, the goal of the experiments in this section (reported by MacLeod, 1996) was to try to develop an implicit measure in which the priming from prior study of a word was expressed as interference rather than facilitation. Although the results turned out rather differently than Warren's (1972) work suggested they would, they are nevertheless informative about the specificity of priming.

In the first experiment, subjects studied 48 high-frequency words under instructions to remember them for a later test, characterized as an explicit test. Following study, subjects were tested on 48 words in a speeded word reading test and 48 words in a color naming test, with test order counterbalanced. On each 48-word test, there were 24 studied words and 24 new words, with no overlap in materials between the two tests. For both tests, the words were presented in color, with subjects instructed to ignore the color in the word reading test and to ignore the word in the color naming test. To discourage the use of conscious recollection during these tests, subjects were told that the tests were filler activities before the real memory test.

As shown on the left side of Fig. 6A, the word reading test revealed reliable facilitation, with the studied words responded to on average about 9 ms faster than the new words. There was, however, no corresponding enhancement of interference: The colors of studied words were, if anything, named slightly faster than those of new words. The same pattern was evident in a second experiment using low-frequency words, which generally show more priming (e.g., MacLeod & Kampe, 1996). Word reading showed facilitation of about 15 ms (right side of Fig. 6A), but color naming again showed no interference.

An experiment not included in the MacLeod (1996) article extended this result by repeating Experiment 2 but with the addition of some standard incongruent Stroop trials (e.g., *red* printed in the color green, say "green"). Because such trials always cause substantial interference, the idea of includ-

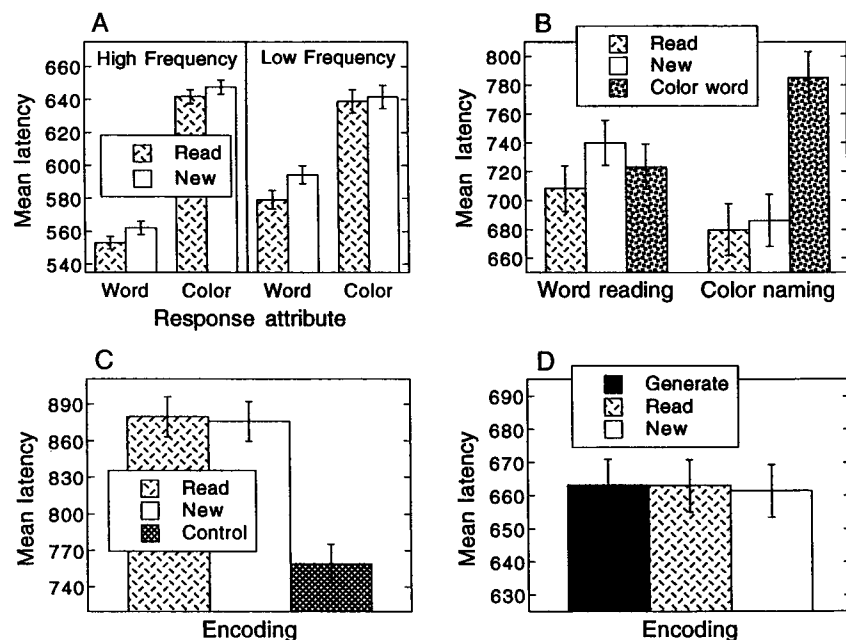


Fig. 6. Mean response latency on speeded word reading and color naming tasks with word stimuli printed in color as a function of prior exposure. (A) Word reading and color naming tasks with high- and low-frequency words. (B) Word reading and color naming tasks, including a set of color words as targets. (C) Color naming task, including control items consisting of rows of Xs. (D) Color naming task. Error bars indicate the 95% within-subject confidence intervals for the means. Data in (A) and (C) are adapted from MacLeod (1996), data in (D) are adapted from MacLeod and Masson (1997a).

ing them was to highlight the presence of words, perhaps drawing more attention to the noncolor words as well. As shown in Fig. 6B, however, although color naming responses on the incongruent Stroop trials were about 100 ms slower than on the other trials, indicating the usual Stroop effect, the studied words still did not interfere more than the new words. This was true despite the strong facilitation of word reading for studied words as compared to new words.

The third experiment in the published series examined only color naming. This time, the number of intervening items as well as the time between the study of a word and its test were reduced by using multiple "minilists" of nine studied words and nine color naming test items. Of the nine test items in each minilist, three were studied words, three were new words, and three were control items (i.e., XXXX). The purpose of these modifications was to give priming more of a chance to express itself by moving the study and test of individual items closer together, and to determine whether noncolor words did in fact interfere with color naming relative to a neutral baseline. The color naming results shown in Fig. 6C indicate that although both studied words and new words did interfere substantially relative to the control, once again there was no enhancement of interference for studied words.

To bridge the experiments in the previous series with those in this series, we reintroduced our standard read versus generate manipulation and followed our standard procedure. This experiment is Experiment 2 in the MacLeod and Masson (1997a) set, for which the word reading data have already been described in the previous section. Recall that subjects showed 12 ms of priming for generated words and a virtually identical 15 ms of priming for read words, relative to new words. But these subjects also did a color naming task, with the order of the two tests counterbalanced. As shown in Fig. 6D, there was no interference for either generated words or read words relative to new words. Once again, word reading showed facilitation and color naming showed no effect due to priming.

Taken together, then, the MacLeod (1996) experiments and the related experiments just described show that primed words facilitate speeded word reading but do not interfere with color naming. The failure to find increased interference in color naming for studied words is not an isolated one, as Burt (1994) has reported similar findings. These findings have a number of implications, but only one is especially relevant here. Apparently, studying a word in all of these experiments primed that word sufficiently to produce facilitation in word reading. Of course, word reading at test is essentially the same task as word reading at study, so transfer would be expected. But color naming at test has little in common with word reading at study in the way of shared processes. Thus, it would appear that the priming attached

to the words during study is specific to reading those words, and does not spill over onto a color naming task where word reading is irrelevant and indeed discouraged by the instructions to subjects.

The explanation offered by MacLeod (1996) was that priming is process specific. Priming will ordinarily be expressed as facilitation on only a very similar processing task, and should not appear as interference with another task, even one that involves the same studied materials. It is not the materials that are primed, but rather the processes that were applied to those materials on initial contact. This argument is supported by another related set of data. MacLeod, Grant, and Ridout (1996) report three experiments in which they had subjects study a list of words and then take part in a flanker task (see Eriksen, 1995, for a review of this paradigm). In this task, there were always two stimuli on each test trial, one above the other. The target, varying over trials from top to bottom, was enclosed in arrowheads; the subject's task was to decide whether the target was a word. This target could be either a studied word, a new word, or a nonword. The adjacent flanker could be either a studied word, a new word, a nonword, or a neutral control row of Xs. Although targets benefited from prior study (by about 50 ms), flankers were totally unaffected. Priming a flanker did not enhance the interference it caused with the decision about the target.

These two studies both suggest that priming is specific not to the studied item (or its representation), but to the processes in play when that item was analyzed upon its initial occurrence. To the extent that these processes are recruited at test, as when a word that was read during study must be read again at test or must have its lexical status judged (which necessitates reading it), primed words will cause facilitation. But when attention at test is directed to a dimension of the stimulus display other than the primed one, for example, to the color of print of a studied word or to an adjacent target, priming is irrelevant. Indeed, in such cases, priming would have a negative impact, so it is probably good that priming does not routinely affect elements of the stimulus display other than the target.

Our goal in introducing this set of experiments that contrasts color naming and word reading was to show the specificity of the processes involved in priming. These experiments nicely illustrate that it is not words or their representations that are primed, but rather the operations that were applied to those words during the previous encounter. If the words were read at study and must be read again at test, then facilitation will appear because of the repetition of the reading operations. If, however, something else must be done to the words upon second encounter, then the prior processing may not be relevant and so priming will not be evident. We certainly intend to explore this (hyper)specificity more, but for the present only wish to note that it is entirely consistent with the explanation we have been elucidat-

ing here. The initial interpretation records how the stimulus was processed. When this encoding is summoned later because of similar processing of the same stimulus, priming will appear in the form of facilitation. That priming will not leak through to disrupt processing on another stimulus dimension, thereby protecting us from undesirable interference when a stimulus recurs in a different setting that requires different types of processing.

## VII. Experiment Series 6: Sources of Priming in Masked Word Identification

In a recent special issue of the *Canadian Journal of Experimental Psychology* devoted to implicit memory, we (Masson & MacLeod, 1996) tackled a question that had become a growing concern to us in our studies of priming in masked word identification. Put simply, the question is this: Is the priming we have observed repeatedly for studied words in masked word identification the result of increased efficiency in extracting perceptual information from studied words (i.e., sensitivity in signal detection theory terms), or does that priming result from some kind of bias favoring studied words?

Our interest in this question was spurred by the recent work of Ratcliff and McKoon (1993; Ratcliff, McKoon, & Verwoerd, 1989). They had subjects decide which of two probe items was identical to a just-presented masked target. Using this two-alternative forced choice (2-AFC) test, they found that prior exposure to the masked target item did not increase accuracy of recognition, arguing against the view that priming is the result of increased perceptual clarity for studied targets. Instead, Ratcliff et al. (1989) suggested that a perceptual bias was operating. They based this argument on the fact that, when tested with orthographically similar test pairs (e.g., *lied*, *died*), one of which was old (i.e., had been previously seen), subjects tended to choose the old alternative regardless of whether the masked target word was old or new.

Ratcliff et al. (1989) went on to propose that this perceptual bias happened early in perception, and cited two main reasons for this claim. First, response times were very fast, leaving little room for strategic processing. Second, subjects said that the old words seemed to "pop out." Faced with a degraded stimulus as in masked word identification, subjects prefer to select an old word when the possible responses are similar. That this is a bias was underscored by the finding in their Experiment 5 that there was no bias in favor of old words when the two alternatives were not orthographically alike (e.g., *died*, *love*).



We wondered whether the free report procedure that had been used in our previous experiments might be harboring such a bias toward old, studied words or whether the old words were actually seen better, as argued by Reinitz and Alexander (1996). To address this question, we conducted a series of seven experiments (Masson & MacLeod, 1996). In all of these, subjects first studied a set of words by reading them; the generation encoding task was not included here because our question was about why priming occurs in masked word identification, and did not require comparing two encoding conditions. Subjects then were given one or two forms of masked word identification test. There were three different forms of test in all. One form of the test was the "free report" version of identification that we have described in prior sections of this chapter: Subjects saw the masked word and then tried to say what it was. Regardless of whether priming results from sensitivity or bias, free report should lead to better performance for studied words. To discriminate these two explanations, we used two other testing procedures following the masked word. One of these was a 2-AFC test and the other was a single probe test. Incorporation of the single probe test allowed us to examine whether any bias observed on the 2-AFC test stemmed from strategic use of discriminative information in the two choices, a strategy that could not operate in the case of a single probe.

We first describe the pattern of results over the entire series of experiments, then return to what those results tell us about sensitivity versus bias. Experiment 1 was essentially a replication of the Ratcliff et al. (1989) study with one key change. Except for a baseline condition, Ratcliff et al. had one old and one new alternative on every trial, whether the masked target was old or new. This was what allowed them to detect the bias favoring old items. We chose instead to prevent this particular bias from operating by always having both alternatives on the recognition test have the same study status, either both old or both new. Like Ratcliff et al., our test pairs were orthographically similar, differing by only one internal letter (e.g., *cage, cave*). Subjects studied 80 words (consisting of 40 pairs of orthographically similar words) one at a time in random order. They then took part either in a mixture of free report and 2-AFC masked word identification trials, each using 20 studied targets and 20 nonstudied targets, or a set of test trials using the single probe procedure (with 40 studied and 40 nonstudied targets). For the 2-AFC and single probe tests, studied words that were not tested as targets were used as foils.

Figure 7A shows that for Experiment 1, free report produced a strong advantage for studied words, consistent with our masked word identification studies (i.e., MacLeod & Masson, 1997b; Masson & MacLeod, 1992). But this advantage disappeared on the 2-AFC test. When, in Experiment 2, we switched from related probe pairs to using pairs of words on the 2-AFC

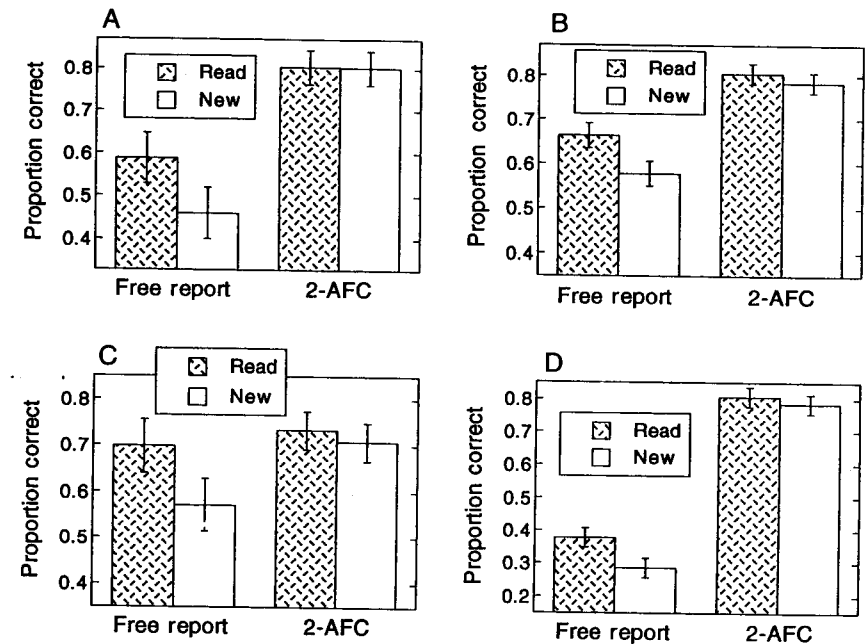


Fig. 7. Mean proportion correct on free report and two-alternative forced choice (2-AFC) versions of masked word identification as a function of prior exposure. (A) Probe pairs on the 2-AFC test were orthographically similar. (B) Probe pairs on the 2-AFC test were not similar in any way. (C) A 1-s pause was inserted between the masked target and presentation of the probes on the 2-AFC test. (D) A 2-s pause was inserted between the masked target and presentation of the probes on the 2-AFC test, during which subjects attempted to report the target word; these report attempts constitute the free report data. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from Masson and MacLeod (1996).

test that were not orthographically related, we obtained the same result (see Fig. 7B). In Experiment 4, we inserted a 1-s delay between the target and the test probes on the 2-AFC trials. We expected that subjects would attempt to identify the target during that interval before the probes appeared. If subjects could more successfully identify studied as compared to nonstudied targets under these conditions, the advantage for studied targets ought to be revealed on the 2-AFC test. That effect did not appear, as can be seen by examining Fig. 7C.

In Experiment 6, a further attempt to produce a priming effect on the 2-AFC test was made by inserting a 2-s delay between the masked target and the probes, and by instructing subjects to attempt to identify the target

before the two alternatives appeared. To increase the power of this experiment, all targets were tested using the 2-AFC method and there were no free report trials. Nevertheless, this experiment also failed to produce a priming effect on the 2-AFC test, despite the fact that the free report attempts were more often correct for studied than for nonstudied targets. That advantage of prior study, however, was not robust enough to produce an effect on the 2-AFC test, probably because correct free report occurred on only about one third of the trials. The relatively low free report performance was a result of using display parameters that were sufficiently demanding that 2-AFC performance did not reach ceiling. Thus, subjects were often unsuccessful in their preliminary free report attempt and therefore frequently waited for the probes to appear before making a response to the masked target word.

The remaining experiments involved the single-probe test, where a single word followed the masked target and the subject had to indicate whether that word was identical to the target. The study status of the probe was always the same as that of the target, to prevent the Ratcliff et al. (1989) type of bias from operating. The left side of Fig. 8 shows the data from Experiment 3, where the probe word immediately followed the mask. Accuracy on the probe task was measured using  $A'$ , a measure of sensitivity that is not subject to the assumptions of normally shaped and equal-variance signal and noise distributions (e.g., Grier, 1971). Clearly, there was no evidence of priming. The same thing happened in Experiment 5 with a 2-s delay inserted between the masked target and the probe (see the middle section of Fig. 8). Again, we had expected that subjects would attempt to

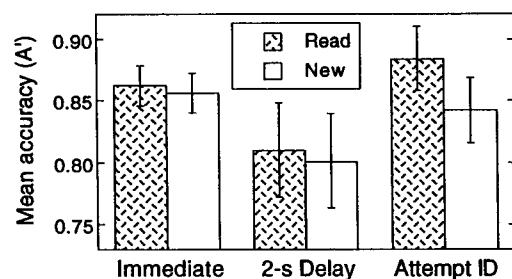


Fig. 8. Mean accuracy on the single-probe version of the masked word identification task as a function of prior exposure. Dependent measure is  $A'$ , a signal detection measure of sensitivity. *Left section:* the probe immediately followed the masked target. *Middle section:* there was a 2-s pause between the masked target and the probe. *Right section:* there was a 2-s pause between the masked target and the probe, during which subjects attempted to report the target word. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from Masson and MacLeod (1996).

identify the target during that delay interval and that greater success at identifying a studied target compared to a nonstudied target would carry over to the single-probe decision.

In Experiment 7, however, when we combined a 2-s delay with instructions to attempt identification of the target prior to the appearance of the probe, we now observed priming in the probe task, as shown in the right section of Fig. 8. The single-probe procedure permitted use of less demanding display conditions than was the case with the 2-AFC test. Therefore, subjects were more frequently able in Experiment 7 to identify targets during the 2-s delay (about 45% of the time). We suspect that this greater level of preliminary report success led subjects' decisions about the subsequent probe to be significantly affected by what was experienced during the preliminary report attempt.

This series of experiments presents a quite simple picture. The better performance for studied words than for nonstudied words seen in the free report data for masked word identification both in this series and in others (MacLeod & Masson, in press; Masson & MacLeod, 1996) does indeed reflect a kind of bias. The advantage for studied words disappears in 2-AFC and single-probe versions of the testing procedure unless, as in Experiment 7, sufficient time and adequate display conditions are provided to allow frequent identification of the masked target word before the probe word appears. The important remaining task is to make sense of this pattern of results.

On the basis of their results, Ratcliff et al. (1989) suggested that performance differences favoring studied items were not being driven by heightened perceptual sensitivity to those items. We agree with this conclusion. Ratcliff et al. went on to propose a type of early perceptual bias in favor of studied words. It is not clear that this form of bias can account for our results, however, given that we prevented that bias from operating by always having test pairs and probes match the target with respect to study status. Instead, we argued in the Masson and MacLeod (1996) article for a fluency heuristic, as outlined in the introduction to this chapter. Our suggestion is that a masked target word that was studied earlier comes to mind more easily than one that is new, and that this is especially likely to be observed under impoverished conditions such as masking. A studied word comes to mind more easily because of the automatic retrieval of the previous encoding episode involving that word, cued by the masked target presentation. Because the studied word comes to mind more fluently, subjects are more likely to conclude that it was the presented target and to output it in free report identification. On a 2-AFC or a single-probe test, however, subjects are instead led to use discriminative information in comparing test items to the masked target word. As a result, the processing fluency generated

by studied words is ignored. Only when fluency is forced back into the probe task by adding a delay and urging subjects to attempt to identify the masked target prior to the probe's appearance do subjects again revert to relying on fluency and again show priming.

There is a tendency to think of a bias explanation as a disappointing one, the interpretation being that the situation under study does not represent a very interesting psychological phenomenon. We would argue quite vigorously against this view. The fluency heuristic may indeed be a very fundamental aspect of remembering, helping us to repeat actions more smoothly each time and even permitting us to make attributions about whether a situation or stimulus has been encountered previously. Surely these are the primary functions of memory. The experiments in this series tell us that priming in masked word identification is a memory phenomenon, not a perceptual one. Consequently, we must disagree with the conclusion reached by Reinitz and Alexander (1996) that priming occurs via the increased rate of gain of visual information from studied items. The evidence presented in this section in large part underlies why we prefer not to call the masked word identification task "perceptual identification." The lesson is that we should avoid yoking our theories to our task designations because the tasks ordinarily outlive the initial theories proposed to account for performance in them.

### VIII. Experiment Series 7: Episodic Effects on Perceptual Judgments

We have argued that a fluency heuristic is responsible for the repetition priming observed in the masked word identification task. Moreover, this heuristic is assumed to operate similarly following data-driven and conceptually driven encoding tasks. The consequence of either type of prior encoding episode is that a target word comes to mind more fluently when that word is later presented under perceptually challenging conditions. In this final series of experiments, we report evidence to support the claim that this fluency can be experienced by subjects as a perceptual phenomenon regardless of whether its source is an earlier perceptual encoding, as in the read task, or a conceptual encoding with no direct perceptual processing, as in the generate task.

The experiments in this series build on results reported by Witherspoon and Allan (1985) and others (Jacoby, Allan, Collins, & Larwill, 1988; Whittlesea et al., 1990). These researchers have shown that subjects can be led to attribute the processing fluency arising from prior exposure of target items to the perceptual conditions under which target items are

tested rather than to prior exposure. In the Witherspoon and Allan study, for example, subjects' estimates of the exposure duration of briefly presented, masked target words were increased if the targets had appeared in an earlier study list. Similarly, Jacoby et al. found that when sentences were presented auditorily in a background of white noise varying in loudness, ratings of the loudness of the noise were lower if sentences had been heard in an earlier study phase. In general, it appears that attributions of processing fluency to perceptual characteristics of stimuli can be induced if subjects are asked to report on those characteristics.

We attempted to capitalize on this phenomenon in a comparison of the influence of read and generate encoding episodes. The question we addressed was whether both of these types of encoding are capable of affecting subjects' perceptual judgments. If there is a qualitative difference in what subjects experience as a result of a data-driven encoding task as opposed to a conceptually driven task, we might expect that perceptual judgments would be influenced only (or at least primarily) by the former type of task. On the other hand, if both classes of encoding task create the potential for processing fluency that is qualitatively similar, both kinds of task should affect perceptual judgments about subsequent target displays in the same way.

These ideas were tested in a series of experiments by Caldwell and Masson (1997) that adopted the general procedure developed by Witherspoon and Allan (1985). In the first experiment, subjects read or generated from definitions a series of target words. Next, subjects were shown a series of filler words, each displayed for 30, 60, or 90 ms and followed by a pattern mask. Subjects were informed that three different durations were being used and the task was to rate the duration of each display on a three-point scale. After the filler trials, critical targets began to appear, some of which had been read or generated in the study phase, whereas others were new words. At this point, however, unknown to the subjects, only 30-ms and 60-ms exposure durations were used. It was expected that prior exposure to targets would lead subjects to experience greater processing fluency and to attribute that experience to longer exposure duration, as had been observed by Witherspoon and Allan. That effect would be expressed as longer duration ratings for previously studied targets. By withholding 90-ms exposure trials, we hoped to increase the likelihood that subjects would be induced by experienced fluency to assign the highest duration rating to trials on which old targets were shown for only 60 ms (or even 30 ms). Of particular interest was the possibility that both read and generate encoding tasks would produce a similar effect on the duration ratings.

On a three-point scale, the mean duration ratings for generate, read, and new critical targets presented in the 60-ms condition showed little variation

(2.25–2.27) and did not reliably differ as a function of encoding condition. In the 30-ms condition, however, there were significant differences between the mean ratings for the three conditions. In particular, the generate and read targets produced mean ratings that were not reliably different (1.55 and 1.60, respectively), although each of these conditions was significantly higher than the average for new items (1.46).

The results of the first experiment were encouraging if not fully convincing. The failure to find an effect of prior presentation among the 60-ms targets may have been a result of the fact that none of the trials in the critical part of the test phase used the longest exposure duration of 90 ms. Therefore, subjects may have been reluctant to use the highest rating. An examination of the rating values shows this was not the case. For critical targets shown in the 60-ms condition, nearly 40% were assigned the maximum duration rating of three but these assignments were evenly spread across old and new targets alike.

Another possibility is that in the duration judgment task subjects may not have processed the target words in a manner that would consistently lead them to experience the fluency produced by prior encoding episodes. Given that the only requirement was to rate the exposure duration of the words, subjects often may not have engaged in attempts to identify the words. On our account, it is the attempt to identify a target that will induce processing fluency and lead to the subjective experience of longer exposure duration. To test this idea, a second experiment was conducted in which subjects again judged exposure duration, but they did so only after first attempting to identify the target word. Another change from the first experiment involved the exposure durations used in the test phase. For the filler words presented to acquaint subjects with the various exposure durations, the durations were 15, 45, and 75 ms. When the critical target words began to appear, mixed in with some filler targets, the critical items all were shown at a single duration midway between two of the practice durations. Again, this change was not announced to the subjects. The set of filler items was shown at the third duration. For half of the subjects, the critical targets were shown for 30 ms and the fillers were shown for 75 ms. For the other half of the subjects, the critical targets were shown for 60 ms and the fillers were shown for 15 ms.

Thus, we created a between-subjects replication of the critical exposure durations used in the first experiment. In this experiment, however, the critical target duration was situated between two previously experienced durations. By making the actual exposure duration (30 or 60 ms) “ambiguous” with respect to the durations with which the subjects initially had been trained, the expectation was that subjects would tend to assign new

targets to the lower duration category and old targets to the higher duration category that flanked the actual exposure duration.

The mean proportion of correctly identified targets and the mean duration ratings are shown in Figures 9A and 9B, respectively. Subjects in the 60-ms group correctly identified nearly all of the critical target words, leading to little difference in accuracy as a function of prior exposure. In the 30-ms group, subjects identified about half of the critical targets; the generate and read conditions led to very similar identification performance, which was substantially better than identification accuracy in the new condition (our typical pattern). These different patterns of identification accuracy led to a highly reliable interaction between encoding task and target duration.

Despite the difference in the pattern of identification performance in the two duration groups, the profile of duration ratings was remarkably consistent across duration. Duration ratings were, of course, much higher for the 60-ms group than for the 30-ms group. The effect of prior exposure was robust, with higher duration ratings for previously studied items, regardless of whether the encoding episode involved generation or reading. There was no interaction between encoding task and duration, indicating that the pattern of effects on duration ratings was the same for both groups. Planned comparisons indicated that the duration ratings did not differ between the generate and read encoding tasks, but the average rating across these two prior exposure conditions was reliably higher than for the new condition. This pattern held even when the 60-ms group was considered alone.

Prior exposure to target words led to a small but highly reliable increase in rated exposure duration, regardless of whether subjects engaged in per-

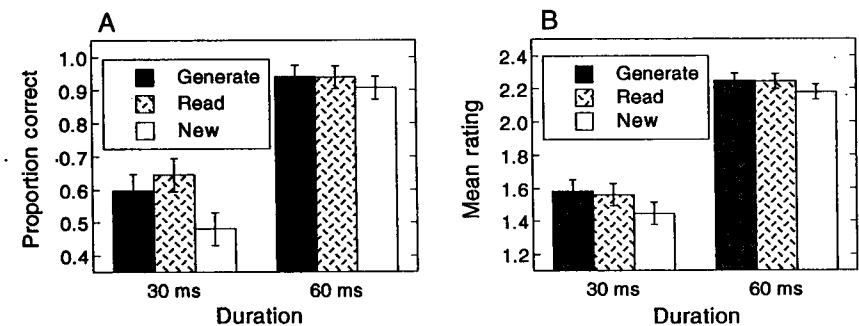


Fig. 9. (A) Mean proportion of correctly identified targets and (B) mean target duration ratings as a function of encoding task. Error bars indicate the 95% within-subject confidence intervals for the means. Data are adapted from Caldwell and Masson (1997).

ceptual processing of the targets during the study phase. The importance of inducing subjects to identify the targets is apparent from the fact that the effect of prior exposure was more robust here than in the first experiment of the series, when subjects were not instructed to identify the targets. In that first experiment, we failed to find an effect of prior exposure on duration ratings when the actual exposure duration was 60 ms. As expected, having subjects attempt to identify the targets made the effect of prior exposure on duration ratings more robust. We suggest that this occurred because the influence of prior exposure on duration ratings operates through fluency experienced when viewing and attempting to identify the target words.

One might argue, however, that the effect of prior exposure on duration ratings was a result of either (1) conscious recollection of the words' prior occurrence or (2) a strategy whereby subjects systematically gave higher duration ratings to words they successfully identified. Various aspects of the results of the second experiment in this series are inconsistent with these possibilities. First, if subjects were relying on conscious recollection to determine duration ratings, we should have observed a stronger effect of prior exposure in the generate than in the read condition. Many previous studies have shown that awareness of prior occurrence as measured by direct tests of memory is much more likely to follow the generate encoding task than the read encoding task (e.g., Jacoby, 1983; Masson & MacLeod, 1992). Instead, the generate and read conditions consistently produced similar duration ratings as well as similar identification success. Even in the 60-ms group, where nearly all targets were correctly identified, the mean duration ratings were nearly identical for targets in the generate and read conditions.

On the question of strategically based ratings, tied to successful identification, we note that a reliable effect of prior exposure on duration ratings was found in the 60-ms group, even though identification accuracy in that group was uniformly high so that accuracy was unaffected by prior exposure. This dissociation between identification accuracy and duration ratings suggests that differential ratings for studied and nonstudied targets was not simply a result of successful identification. Rather, the pattern of ratings indicates that subjects were sensitive to differences in identification fluency, not merely to the success or failure of their identification attempts.

Finally, it is particularly important that the generate and read encoding conditions produced similar results in both identification accuracy and duration ratings. We take these results as evidence that the processing fluency experienced by subjects during the masked presentation of target words is qualitatively similar for items that were generated or read at the time of encoding. Thus, prior perceptual experience of a target word was not required to produce enhancement of either identification or exposure dura-

tion ratings. Instead, an episode involving conceptually driven processing of a target word (i.e., generation from a semantic cue) was sufficient to produce both enhanced identification and inflated duration judgments. These results lend further support to our argument that skilled word identification recruits relevant prior experiences that are either perceptually or conceptually relevant to the current identification attempt. To the extent that especially relevant prior episodes are recruited, a target word will be experienced as more readily perceived, and that experience may be attributed to whatever dimension the subject is asked to judge.

## IX. Conclusion

We have attempted to build a case for the view that memory for prior encoding episodes increases the fluency with which the current encoding of a stimulus goes forward. In tasks involving the identification of a target word, either through a brief, masked presentation, or through completion of a word fragment or stem, prior experiences with the target word that share aspects of the current processing task will be recruited and used in the construction of an interpretation of the target stimulus. The availability of such prior episodes will enable a more fluent construction of an interpretive encoding. On this account, prior episodes do not contribute to more efficient perceptual processing of the target's physical features (cf. Reinitz & Alexander, 1996). Rather, the recruitment of a relevant prior encoding episode contributes to the fluent interpretation of those features. For this reason, prior episodes that emphasize a word's meaning also have the potential to improve identification of that word when it is presented under degraded conditions on a subsequent test. To the extent that the form in which a word is presented supports the recruitment of conceptual aspects of prior processing, even prior episodes that involved no direct perceptual experience (such as the generation task) can enhance identification performance.

The general effect of this initial interpretive encoding of a target word is subjectively experienced as the word's identity coming to mind fluently. That experienced fluency is in turn subject to attribution on the part of the subject. For example, if the task is to identify the target word, experienced fluency is likely to increase the subject's confidence that the correct candidate has been selected. If the task is to make a judgment about some perceptual aspect of the target, such as the duration of its presentation, fluent reprocessing may be attributed to that aspect of the target (e.g., Experiment Series 7). Finally, if the task is to determine whether the target word had occurred earlier in the context of the experiment, fluent identifi-

cation provides evidence of such an occurrence. Thus, artificially created processing fluency can create illusions of prior occurrence (e.g., Whittlesea et al., 1990). On this account, conscious awareness of prior occurrence results from an evaluation of the current encoding episode (see Whittlesea's chapter, this volume), rather than contributing to its success.

More generally, we see the influence of prior encoding episodes as the basis for the development of skilled word reading (e.g., Masson, 1986). Although models of word reading typically ascribe skilled performance to activation of stable, abstract word representations such as logogens (e.g., Morton, 1969), there is mounting evidence to suggest that specific episodes powerfully affect word identification even among highly accomplished readers. Some of the earliest research demonstrating the effects of repetition on word identification is remarkable for this very reason (e.g., Scarborough, Cortese, & Scarborough, 1977). In our view, the surprisingly large benefit of a single additional exposure to a word derives from the substantial similarity in contextually bound episodes that constitute the initial and repeated presentations of a word in the repetition priming paradigm. That strong similarity gives the first presentation of the target in the context of the experiment a privileged role relative to the many hundreds or thousands of previous experiences with that target word. Thus, an episode with strong contextual relevance produces a much more substantial influence on performance than would be expected from a simple increment in the number of generic encounters with a word (Kirsner & Spelman, 1996). On this view, although skill can be generalized, it is at the same time subject to contextual specificity in the same way and for the same reason as other memory phenomena.

#### A. CONSCIOUS RECOLLECTION

On the account advocated here, conscious recollection of a target's prior occurrence is an attribution that arises in part from the fluency experienced during the encoding of the target. Further grounds for deciding that a target had been studied earlier are likely to be present if that study experience involved conceptually driven processing. Elements of such encoding episodes are typically well remembered, particularly when the target word that was the core of the episode is available as a cue.

An alternative view of the role played by conscious recollection is that this type of remembering can inflate or contaminate performance on indirect tests of memory. That is, conscious recollection is assumed to have the potential to contribute to successful generation of a target word on tasks such as masked word identification or stem completion. The influence of conceptually driven encoding tasks on subsequent word identification

tests that are assumed to be dominated by data-driven processes is therefore ascribed to conscious recollection, rather than to an automatic influence of memory for conceptually driven processing (Toth et al., 1994). On this view, our demonstrations of similar amounts of repetition priming following generate and read encoding tasks are the result of conscious recollection artifactually enhancing priming effects among targets encoded with the generate task.

The most clear statement of how this contamination could come about has been developed by Jacoby and his colleagues (Jacoby et al., 1992) in the context of the process dissociation procedure. In that approach, it is assumed that conscious recollection and automatic influences of memory independently contribute to responding on word identification tests. This position contrasts with the account we prefer, in which conscious recollection or awareness of prior occurrence is a constructed attribution that follows successful word identification. In Experiment Series 3, we showed how the process dissociation procedure can yield underestimates of the automatic influence of memory for conceptually driven encoding episodes because of this attribution process. Moreover, by assessing the effects of generate and read encoding tasks on a speeded word reading test in Experiment Series 4, we were able to create a situation that should have minimized if not prevented any useful contribution of conscious recollection to test performance. Those experiments replicated our findings with the masked word identification test, in which the generate encoding task led to reliable repetition priming that was not significantly less than that observed for targets that had been encoded in the read task. We take these results as strong evidence that conscious recollection is not responsible for producing the repetition priming on word identification tasks that follows generate encoding episodes. In fact, it is just the opposite: Successful identification of such word targets on these indirect tests of memory leads to conscious recollection of prior occurrence.

#### B. CHALLENGING THE TASK PURITY ASSUMPTION

In our research on the influence of processing episodes on word identification, we have emphasized the notion of interpretive encoding. We have done so in an effort to highlight the important role played by conceptual processes in constructing an interpretation of a stimulus. Our view is that this construction is a crucial part of encoding and identification. This approach challenges the assumption that certain word identification tasks are pure with respect to data-driven and conceptually driven processing operations. Earlier demonstrations of modality specificity and weak or nonexistent benefits of prior conceptual processing for tasks such as word



stem completion and masked word identification have encouraged the view that such tasks are perceptual in nature (e.g., Blaxton, 1989; Jacoby, 1983; Weldon, 1991).

Without denying the validity of modality-specific effects on these tasks, we maintain that contributions from conceptually driven processes are also genuine. We have argued that the transfer-appropriate processing framework should be extended to include the important role played by conceptual processes in the construction of interpretive encodings of stimuli (Masson & MacLeod, 1992). Demonstrations of greater enhancement produced by data-driven encoding experiences indicate that certain word identification tasks (e.g., word fragment completion) provide less opportunity for conceptual processes to take hold than do other tasks, such as masked word identification or speeded word reading. In these latter tasks, the availability of the entire word is presumed to enable greater conceptually based recruitment of prior episodes. Accordingly, our research has shown that there are word identification tasks in which conceptually driven encoding episodes make a substantial contribution to later processing fluency. More generally, then, we expect that the influence of conceptually driven encoding episodes on subsequent word identification performance will vary, depending on the extent of conceptual processing afforded by the cues available in the identification task.

#### ACKNOWLEDGMENT

Preparation of this chapter and the experiments reported here were supported by research grants from the Natural Science and Engineering Research Council of Canada to Michael Masson and to Colin MacLeod. We are grateful to the many graduate students and research assistants who have contributed to this project, particularly Natasa Arbutina, Glen Bodner, Graham Brown, Judy Caldwell, Carrie Hicks, Shelley Hodder, Anne Kelly, Alyssa Megale, and Tonya Stokes.

#### REFERENCES

- 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.
- Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340-357.
- Bassili, J. N., Smith, M. C., & MacLeod, C. M. (1989). Auditory and visual word-stem completion: Separating data-driven and conceptually driven processes. *Quarterly Journal of Experimental Psychology*, 41A, 439-453.

- Becker, S., Moscovitch, M., & Behrmann, M. (in press). Long-term semantic priming: A computational account and empirical evidence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 657-668.
- Bodner, G. E., Caldwell, J. I., & Masson, M. E. J. (1997). *Involuntary conscious memory affects parameter estimates in the process dissociation procedure*. Manuscript submitted for publication.
- Burt, J. S. (1994). Identity primes produce facilitation in a colour naming task. *Quarterly Journal of Experimental Psychology*, 47A, 957-1000.
- Caldwell, J. I., & Masson, M. E. J. (1997). *Conceptually driven encoding episodes create perceptual misattributions*. Manuscript in preparation.
- Carr, T. H., Brown, J. S., & Charalambous, A. (1989). Repetition and reading: Perceptual encoding mechanisms are very abstract but not very interactive. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 763-778.
- Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 341-350.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- 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.
- Dagenbach, D., Carr, T. H., & Wilhelmsen, A. (1989). Task-induced strategies and near-threshold priming: Conscious effects on unconscious perception. *Journal of Memory and Language*, 28, 412-443.
- de Groot, A. M. B. (1983). The range of automatic spreading activation in word priming. *Journal of Verbal Learning and Verbal Behavior*, 22, 417-436.
- Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review*, 95, 91-101.
- Eriksen, C. W. (1995). The flankers task and response competition: A useful tool for investigating a variety of cognitive problems. *Visual Cognition*, 2, 101-118.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, 12, 627-635.
- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, 16, 309-313.
- Graf, P., & Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior*, 23, 553-568.
- Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 501-518.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, 22, 485-508.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.
- Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 240-247.

- Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 1-47). New York: Academic Press.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306-340.
- Jacoby, L. L., Lindsay, D. S., & Toth, J. P. (1992). Unconscious influences revealed. *American Psychologist*, *47*, 802-809.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, *122*, 139-154.
- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, *118*, 126-135.
- Jacoby, L. L., Yonelinas, A. P., & Jennings, J. (1996). The relation between conscious and unconscious (automatic) influences: A declaration of independence. In J. Cohen & J. W. Schooler (Eds.), *Scientific approaches to the question of consciousness* (pp. 13-47). Mahwah, NJ: Erlbaum.
- Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General*, *122*, 462-467.
- Kintsch, W. (1970). *Learning, memory, and conceptual processes*. New York: Wiley.
- Kirsner, K., & Speelman, C. (1996). Skill acquisition and repetition priming: One principle, many processes? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 563-575.
- Kolers, P. A. (1976). Reading a year later. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 554-565.
- Kolers, P. A., & Roediger, H. L., III. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*, 425-449.
- Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 1-11.
- Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1-35.
- Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107-128.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163-203.
- MacLeod, C. M. (1996). How priming affects two speeded implicit tests of remembering: Naming colors versus reading words. *Consciousness & Cognition*, *5*, 73-90.
- MacLeod, C. M. (in press). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches*. Mahwah, NJ: Erlbaum.
- MacLeod, C. M., & Daniels, K. A. (1997). *Implicit vs. explicit remembering: Directed forgetting meets the generation effect*. Manuscript in preparation.
- MacLeod, C. M., Grant, S. C., & Ridout, S. M. (1996). *Priming flankers and targets in a lexical decision flanker task*. Manuscript submitted for publication.
- MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 132-142.
- MacLeod, C. M., & Masson, M. E. J. (1997a). *Fluent word reading as a measure of implicit remembering*. Manuscript in preparation.
- MacLeod, C. M., & Masson, M. E. J. (1997b). Priming patterns are different in masked word identification and word fragment completion. *Journal of Memory and Language*, *36*, 461-483.
- Masson, M. E. J. (1986). Identification of typographically transformed words: Instance-based skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 479-488.
- Masson, M. E. J., & MacLeod, C. M. (1992). Re-enacting the route to interpretation: Enhanced perceptual identification without prior perception. *Journal of Experimental Psychology: General*, *121*, 145-176.
- Masson, M. E. J., & MacLeod, C. M. (1996). Contributions of processing fluency to repetition effects in masked word identification. *Canadian Journal of Experimental Psychology*, *50*, 9-21.
- Masson, M. E. J., & Sala, L. S. (1978). Interactive processes in sentence comprehension and recognition. *Cognitive Psychology*, *10*, 244-270.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 227-234.
- Morris, C. D., Bransford, J. P., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519-533.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, *76*, 165-178.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.
- Nelson, D. L., Keelean, P. D., & Negrao, M. (1989). Word-fragment cuing: The lexical search hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 388-397.
- Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*, 132-143.
- Ratcliff, R., & McKoon, G. (1993, November). *Bias in implicit memory tests*. Paper presented at the annual meeting of the Psychonomic Society, Washington, DC.
- Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 378-387.
- Reingold, E. M., & Toth, J. P. (1996). Process dissociations versus task dissociations: A controversy in progress. In G. Underwood (Ed.), *Implicit cognition* (pp. 159-202). New York: Oxford University Press.
- Reinitz, M. T., & Alexander, R. (1996). Mechanisms of facilitation in primed perceptual identification. *Memory & Cognition*, *24*, 129-135.
- Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology*, *39*, 475-543.
- Richardson-Klavehn, A., & Gardiner, J. M. (1995). Retrieval volition and memorial awareness in stem completion: An empirical analysis. *Psychological Research*, *57*, 166-178.
- Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996). Memory: Task dissociations, process dissociations and dissociations of consciousness. In G. Underwood (Ed.), *Implicit cognition* (pp. 85-158). New York: Oxford University Press.
- Richardson-Klavehn, A., Gardiner, J. M., & Ramponi, C. (1996). *Does depth of processing influence involuntary retrieval in stem completion?* Manuscript submitted for publication.
- Roediger, H. L., III. (1990). Implicit memory: Retention without remembering. *American Psychologist*, *45*, 1043-1056.
- Roediger, H. L., III, & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word fragment completion. *Memory & Cognition*, *15*, 379-388.

- Roediger, H. L., III, & McDermott, K. B. (1993). Implicit memory in normal human subjects. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 63–131). Amsterdam: Elsevier.
- Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A process account. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–14). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.
- Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1–17.
- Schacter, D. L., Bowers, J., & Booker, J. (1989). Intention, awareness, and implicit memory: The retrieval intentionality criterion. In S. Lewandowsky, J. C. Dunn, & K. Kirsner (Eds.), *Implicit memory: Theoretical perspectives* (pp. 47–65). Hillsdale, NJ: Erlbaum.
- Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, 14, 489–537.
- Semon, R. (1921). *The mneme*. London: Allen & Unwin. (Originally published in 1904 as *Die Mneme*. Leipzig: Engelmann)
- Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 592–604.
- Strain, E., Patterson, K., & Seidenberg, M. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1140–1154.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18, 645–660.
- Toth, J. P., & Hunt, R. R. (1990). Effect of generation on a word-identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 993–1003.
- Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1994). Toward a redefinition of implicit memory: Process dissociation following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 290–303.
- Toth, J. P., Reingold, E. M., & Jacoby, L. L. (1995). A response to Graf and Komatsu's critique of the process dissociation procedure: When is caution necessary? *European Journal of Cognitive Psychology*, 7, 113–130.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, 26, 1–12.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.
- Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, 94, 90–100.
- Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review*, 72, 89–104.
- Weldon, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 526–541.
- Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, 29, 716–732.
- Witherspoon, D., & Allan, L. G. (1985). The effects of a prior presentation on temporal judgments in a perceptual identification task. *Memory & Cognition*, 13, 101–111.
- Witherspoon, D., & Moscovitch, M. (1989). Stochastic independence between two implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 22–30.