JOURNAL OF MEMORY AND LANGUAGE **36,** 461–483 (1997)
ARTICLE NO. ML972501

Priming Patterns Are Different in Masked Word Identification and Word Fragment Completion

COLIN M. MACLEOD

University of Toronto at Scarborough, Canada

AND

MICHAEL E. J. MASSON

University of Victoria, Canada

Indirect tests of remembering have revealed two different patterns of priming following generation and reading tasks: (1) read words produce more priming than generated words, which produce little or no priming relative to new words, and (2) both read and generated words show reliable and equivalent priming. In a series of six experiments using both mixed and blocked presentation of encoding tasks, we confirmed that the word fragment completion task reliably produced the first pattern of results whereas we found that the masked word identification task almost always produced the second pattern of results. Only when three different tasks were presented in a blocked design during encoding did the identification task lead to less priming for generated than for read words. We conclude that the brief presentation of a whole word in the masked word identification task makes contact with an initial interpretive encoding that includes records of conceptual as well as perceptual operations performed during encoding.

In the early 1970s, a new conception of memory emerged. It began with the emphasis on processing inherent in the levels of processing framework (Craik & Lockhart, 1972) and with the encoding specificity principle (Tulving & Thomson, 1973) that stressed the interplay between encoding and retrieval. These fundamental ideas came together in the work of Kolers (1976; Kolers & Roediger, 1984) and in the idea of transfer appropriate processing (Morris, Bransford, & Franks, 1977). Under these views, memory contained not the products of processing but a record of the actual processes themselves. Remembering was successful to the extent that the same processes were applied at encoding and retrieval, resulting in successful transfer of processing.

This analysis of remembering gained even more currency as cognitive psychologists began to explore unconscious, implicit remembering in addition to conscious, explicit remembering. The implicit/explicit distinction (Graf & Schacter, 1985) is now entrenched in the memory literature and, in addition to fostering the emergence of many new memory tests, has had an increasing impact on theorizing about the operation of memory. Thus, for example, both Jacoby (1983b) and Roediger (1990) have maintained that the nature of the processing at the time of encoding strongly influences performance both on direct tests

This research was supported by Natural Sciences and Engineering Research Council of Canada Research Grants A7459 to Colin M. MacLeod and A7910 to Michael E. J. Masson. For their assistance in collecting the data, we thank Natasa Arbutina and Shelley Hodder in Toronto and Roberta Daye, Chris Gill, and Anne Kelly in Victoria. We also very much appreciate the helpful advice of Douglas Nelson and three anonymous reviewers on the initial version of this article. Address correspondence and reprint requests to Colin M. MacLeod, Division of Life Sciences, University of Toronto at Scarborough, Scarborough, Ontario, Canada M1C 1A4 (e-mail: macleod@lake.scar.utoronto.ca), or to Michael E. J. Masson, Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, British Columbia, Canada V8W 3P5 (e-mail: mmasson@uvic.ca).

0749-596X/97 $25.00

that tap conscious recollection and on indirect tests that tap unconscious remembering.

Data-Driven and Conceptually Driven Processing

For Roediger, Weldon, and Challis (1989), most indirect tests that measure memory implicitly, without requiring awareness, are especially sensitive to data level or perceptual encoding operations. Thus, changes in stimulus form or modality from study to test have a strong impact on tests such as word fragment completion (see Roediger & McDermott, 1993, for a review). In contrast, traditional direct tests that measure memory explicitly, tests such as recall and recognition, are tuned primarily to the conceptual level or meaningful encoding operations. Elaborations that focus attention on meaning during encoding especially benefit conscious recollection. In this way, Roediger has modified transfer appropriate processing to capture both ways of remembering.

One of the definitive experimental manipulations used in the service of this processing distinction is the read versus generate procedure. Here, subjects either read words presented in isolation (e.g., ''horse'') or produce them from a clue (e.g., ''the animal that a cowboy rides - h?''). It is well established that prior generation benefits direct tests dramatically relative to reading (Slamecka & Graf, 1978). What is intriguing is that the reverse is sometimes true for indirect tests, where reading can actually result in superior remembering (e.g., Jacoby, 1983b; Roediger, 1990). Typically, such implicit remembering is indexed by improvement due to repeated processing of a stimulus, often called *priming.*

In their writings, Roediger and his colleagues even suggested that the Read/Generate pattern be used as a sort of litmus test for when a memory test is data driven versus conceptually driven (see Roediger & Blaxton, 1987; Roediger et al., 1989). To capture this in the form of a simple rule, if Generate > Read, then conceptually driven; if Read > Generate, then data-driven.

Still, the majority of the results that the Roediger group has reported and emphasized in theory development correspond to the pattern set out by this rule. Roediger (1990) and Roediger and McDermott (1993) review many studies, especially those using the indirect word fragment completion test, where subjects must try to produce a target word that fits a set of missing letter constraints, as in the example ''s--at-r.'' The repeated finding is that having read the word ''sweater'' earlier leads to a considerably higher probability of completing the fragment than does having generated the word ''sweater,'' whereas the opposite is true on direct tests of recognition or recall. Priming is greater for read than for generated words on the word fragment completion test.

Initial Interpretive Encoding and Subsequent Elaborative Encoding

In earlier work, we observed a different pattern of results (Masson & MacLeod, 1992). The test we used was masked word identification (often called ''perceptual identification,'' though we see this as a misnomer for reasons that will become clear). Here, the subject sees a word presented very briefly, on the order of 30 ms. This presentation is immediately masked, and the subject's task is to identify the word by saying it aloud. Jacoby (1983b) showed that a reading encoding task produced better identification than did an antonym generation encoding task (hot-?; generate ''cold''), a result that we replicated when we used antonyms (Masson & MacLeod, 1992, Experiment 2). But we observed that the priming accruing to generated versus read words was usually identical for a wide range of generation rules, from definitional phrases to famous names to synonyms. This was contrary to the rule for indirect tests following a Read/Generate manipulation, and conflicted with the word fragment completion pattern.

On this basis, we argued for a different view of encoding and retrieval, more in keeping with Graf and Mandler's (1984) earlier account and other more recent accounts (e.g., Nelson, Schreiber, & Holley, 1992). Our claim was that subjects produced two encodings at the time of study. The first we called the *initial interpretive encoding,* recorded immediately upon encountering every stimulus, and incorporating conceptual as well as perceptual aspects of stimulus processing. We maintained that the re-enacting of this initial interpretive encoding was primarily responsible for priming on subsequent indirect tests such as masked word identification. Thus, whether generated or read, a word would show priming on such a test. However, the second encoding, the *elaborative encoding* that could also be created given time or incentive, dominated performance on direct tests, explaining why generated words were better recognized.

There is other evidence consistent with our account. As one illustration, Strain, Patterson, and Seidenberg (1995) have shown that reading a word aloud in isolation, a situation where conceptual aspects of the word might be thought to be irrelevant, nevertheless is sensitive to conceptual components of the word. In their study, the rated imageability of the word influenced the time to read the word aloud. Evidently, conceptual aspects of words are recruited quite routinely and unconsciously.

More recently, we have extended our account, arguing that reprocessing a word is enhanced by the greater ease with which that word comes to mind, not the increased ease with which it is perceived (Masson & MacLeod, 1996). To test this idea, we modified the masked word identification task. One modification used a forced choice identification test in which the target display was followed by two probe words, one of which matched the target. The other modification used a single probe word. By presenting the probe display after the target, we expected that subjects would emphasize evaluation of the probe item(s), abandoning their reliance on the ease with which a candidate word came to mind. Both modifications caused priming for studied targets to vanish. Yet when we inserted a delay between the target and single-probe display and required subjects to attempt to identify the target prior to the appearance of the probe, priming returned. Thus, consistent with our account, only when task demands induced subjects to rely on the ease with which a candidate word comes to mind did the effect of prior exposure materialize.

Fluent Remembering

The idea of a studied word easily coming to mind on a masked word identification test is intimately tied to the concept of fluency in remembering, a concept that Jacoby and his colleagues pioneered (e.g., Kelley & Jacoby, 1990). When prior processing is related to subsequent processing, that later processing will benefit from the enhanced ease with which the processes can be rerun. Furthermore, the fact that the previous episode comes to mind more easily may lead the subject to experience a sense of fluency, and that sense may in turn lead to an attribution of ‘‘remembering.’’

There is increasing support for this position. Jacoby and Dallas (1981, p. 333) expressed it early on in contrasting performance on recognition and masked word identification tasks: ‘‘Subjects may base their recognition memory decision on judgments of the relative fluency of their own performance.’’ In their studies of false recognition (Jacoby & Whitehouse, 1989) and of reading and making judgments about words (Whittlesea, Jacoby, & Girard, 1990), these investigators have reinforced this idea further. Put simply, conscious remembering is an attribution made on the heels of fluent reprocessing.

Conflicting Results in Masked Word Identification and Word Fragment Completion

Our primary goal in this article is to explore a conflicting data pattern in the literature. Recall that word fragment completion shows greater priming for read than for generated items whereas masked word identification typically shows equivalent priming for these two encoding conditions. Yet Weldon (1991, Experiment 1) has obtained in masked word identification the pattern characteristic of word fragment completion. Given that she used phrase definitions as her generation condition (very similar to materials we had used in Masson & MacLeod, 1992, Experiments 1 and 7), there is a fundamental and troubling inconsistency between our studies. The experiments to be reported in this article are aimed at locating the source of this discrepancy.

There is, of course, a more fundamental data conflict: Why do different priming patterns occur in masked word identification and word fragment completion? Our answer relies on the difference in the test stimulus presented to the subject. In word fragment completion, the subject sees only part of the word; in masked word identification, the subject sees the entire word, albeit only briefly. We propose that the partial word information in the word fragment completion test tends not to recruit prior conceptual encodings automatically, instead emphasizing perceptual problem solving in the form of letter insertion and lexical search (as suggested by Nelson, Keelean, & Negrao, 1989, Experiment 4). Thus, the Read condition, with its greater perceptual overlap between encoding and retrieval, results in more priming. In contrast, the complete word shown in the masked word identification task does make contact with both the conceptual and the perceptual encodings, which coalesce in the form of the initial interpretive encoding. The result is equivalent priming for both the Read and Generate conditions.

Even this explanation, however, cannot accommodate the Weldon (1991, Experiment 1) result where greater priming was observed for the Read condition than for the Generate condition in masked word identification. This conflict must be resolved. In the series of six experiments that follows, we work through the possible experimental reasons for the discrepancy between her results and ours using the masked word identification task and, in Experiment 3, the word fragment completion task. The factors to be examined across this series of experiments, both separately and interactively, are (1) materials and test instructions (Experiment 1), (2) between versus within subject designs (Experiments 2 and 3), (3) general laboratory differences (Experiment 3), (4) blocking versus mixing encoding tasks (Experiments 4–6), and (5) number of encoding tasks (Experiments 5 and 6). We will explain how and why we examined each of these factors as we introduce each experiment. To minimize redundancy, we begin with an overview of the general method, leaving only the important variations to be described within each experiment.

General Method

Subjects. The subjects were undergraduate students who volunteered to participate in return for extra credit in an introductory psychology course or for a $5.00 payment. Subjects in Experiments 1–4 were students at the University of Toronto at Scarborough; those in Experiments 5 and 6 were students at the University of Victoria.

Materials. The materials used in Experiments 1–4 were taken from Weldon (1991), and were a subset of those used by Weldon and Roediger (1987, Experiment 2). They consisted of 60 target words, their definitions, and their fragment forms (used only in Experiment 3). Following Weldon (1991) and Masson and MacLeod (1992), a stimulus presented to elicit generation during encoding consisted of the definition plus the first letter of the target word (e.g., *an archer shoots a bow and - a?*). Because these materials are used so extensively here, and because they have not been presented elsewhere, we present the set in the Appendix.[1] The materials used in Experiments 5 and 6 were 80 items of the same form as the Weldon (1991) materials but were taken from Masson and MacLeod (1992, Experiment 1) with slightly modified definitions in some cases. In addition, 14 practice, 10 filler, and 50 pretest target words were selected for Experiments 5 and 6.

In Experiments 1–4, an independent random assignment of target items to encoding tasks was used for each subject, with 20 items in each of the Read, Generate, and New conditions. In Experiments 5 and 6, target items were divided into lists of 20 items each and the assignment of these lists to encoding tasks was counterbalanced across subjects so that each item appeared equally often in each encoding task.

Apparatus. Experiments 1–4 were controlled by an IBM-AT compatible microcom-

[1] We are grateful to Mary Susan Weldon for providing her materials and for giving us permission to reprint them in the Appendix.

puter with a color VGA monitor. The program was written in QuickBasic 4.5 and used the routines given by Graves and Bradley (1987, 1988) to achieve synchronization with the monitor refresh cycle. In Experiments 5 and 6, stimuli were presented on a Macintosh II microcomputer equipped with two monochrome monitors. Software was written to synchronize screen displays with the monitors' refresh cycle. One monitor was used to present stimuli to the subject. The other monitor, which could not be seen by the subject, was used to present the target stimulus to the experimenter. For both computer systems, stimuli appeared as black lowercase letters against a white background.

Procedure. In the study phase of each experiment, subjects engaged in one or more encoding tasks. For the Read task, a single word appeared either left-justified on the middle line of the monitor (Experiments 1–4) or at the center of the monitor (Experiments 5 and 6). The task was to read the word aloud. For the Generate task, a definition plus letter restrictor appeared on the monitor instead of a single word. The subject's task was to use the definition and first letter constraint to generate and say aloud the intended word. In Experiments 1–4, failures to produce the intended word were left uncorrected and there was a 500-ms blank between successive items. In Experiments 5 and 6, subjects were told the intended word when they failed to generate it. A third encoding task, the Associate task, was used only in Experiments 5 and 6. In that task, a target word appeared as in the Read task, and the subject read the target aloud and then said the first word that came to mind.

When subjects performed multiple encoding tasks, these were presented either in mixed or in blocked format. In the mixed format, trials from each task were randomly intermixed. In all but Experiment 6, it was clear either from the stimulus or by the use of blocked presentation of encoding tasks what task to perform on each trial. In Experiment 6, each trial began with a 1-s presentation of a single word that cued the subject of the appropriate task (GENERATE, READ, or ASSOCIATE). In Experiments 5 and 6, the first four trials in each encoding task were practice.

The masked word identification task was used in the test phase in all but Experiment 3, where the word fragment completion task was used. No mention was made of the relation between the encoding phase and the indirect test. In masked word identification, all 60 or 80 target words were presented in a random order for a fixed duration, followed by a pattern mask. Subjects attempted to identify each word. In all but the ''forced'' condition of Experiment 1, subjects could pass when they could not think of a response.

In Experiments 1, 2, and 4, each target appeared near the left edge of the middle line of the monitor for 28 ms and was followed by a mask consisting of a row of 14 ampersands. Time to produce a response on each trial was unlimited. Once the subject had responded, the experimenter pressed a key to indicate accuracy on that trial and to initiate the next trial, which began after a 500-ms blank interval. Two practice trials preceded the 60 experimental trials.

In Experiments 5 and 6, displays appeared at the center of the monitor. Each trial began with a 255-ms presentation of two hyphens separated by two more blank characters than the number of characters in the upcoming target. The target then appeared between the hyphens for 30 ms before the whole display was replaced by a mask. The character making up the mask string was ?, X, or &, depending on the subject's performance on a set of 50 pretest trials involving these masks. The pretest trials were run immediately after the study phase and consisted of five trials with a mask consisting of a row of '':'' characters, then 15 trials with each of the other three masks. The masking character that came closest to producing 50% correct target report was chosen for use in the critical trials. Ten filler items were randomly intermixed with the 80 critical targets used in Experiments 5 and 6 (60 from the study phase and 20 new), so that half of the targets would be new words.

In Experiment 3, a word fragment completion test was given instead of the masked word identification test. Each word fragment was presented at the left side of the middle line on the computer monitor until the subject responded, or until 12 s elapsed. As in Experiments 1, 2, and 4, there were two practice test trials before the 60 experimental test trials began.

After the indirect test was completed, subjects were given a direct test of memory. This test was included simply to assure that the standard generation advantage was present on a direct test of memory. In Experiments 1–4, this test was a recognition test. Each of the 60 targets appeared in a new random order, left justified on the center line of the monitor, until the subject orally responded YES or NO. Because all words on this test had appeared on the earlier indirect test, subjects were instructed to respond YES only to those items that had appeared during the study phase, ignoring the indirect test phase. In Experiments 5 and 6, the direct test was free recall. Subjects were required to write down as many words from the study phase as they could remember. They were cautioned against including words that had appeared only on the masked word identification test.

The experimenter kept track of the accuracy of each subject's oral encoding and test responses by pressing appropriate keys on the computer keyboard. In Experiments 1–4 accuracy was tracked by having the experimenter follow a printed protocol prepared prior to the testing of each subject. In Experiments 5 and 6, accuracy on each trial was determined by information presented on the experimenter's monitor.

Data analysis. Each experiment in this article was analyzed in a similar way, with the Type I error rate set at .05 for all statistical tests. Analyses of performance on the indirect and direct tests are reported separately. Data from at most a few subjects in each experiment were discarded when performance on the masked word identification task was at ceiling or floor. Including such subjects would have reduced the likelihood of finding differences between conditions.

Words that were not correctly generated in the Generate encoding task were excluded when scoring performance on the memory tasks. Because these targets were not encoded by subjects in Experiments 1–4, they were functionally new items at the time of test so it was deemed best not to include them. This exclusion of missed Generate items raises the concern that item selection effects may lead to an overestimate of performance in the Generate condition. To allay this concern, we reanalyzed the data from each experiment including the missed Generate items. In all cases but one to be discussed below, this treatment resulted in the same pattern of significant effects as that found when missed Generate items were excluded. In general, we take this outcome to mean that item selection effects did not play a role in the results we report. Our previous work (Masson & MacLeod, 1992), also found that conditionalized analyses produced the same pattern as unconditionalized analyses, probably because the probability of correct generation is usually high (i.e., greater than .80).

We also note that, in the context of Experiments 1–4, including missed Generate items in the analyses constitutes a strict test of the possibility that exclusion of such items produced an item selection effect. Thus, although it is reassuring to find that including these items leaves the pattern of results unaltered, it is less clear what to make of a case where inclusion of missed Generate items causes the pattern to change. Such a change could be due either to item selection or to the fact that missed Generate items were never encoded by subjects and therefore functioned just like new items.

Experiment 1

Given the contrasting patterns of results of Masson and MacLeod (1992) and Weldon (1991, Experiment 1) in the masked word identification task, our first goal was to make comparison of the two studies more straightforward. Could the choice of materials have been crucial? This seemed unlikely given that Masson and MacLeod used definition-like items in their Experiments 1 and 7, materials very much like those of Weldon. Nevertheless, to be careful and thorough, we began

TABLE 1

EXPERIMENT 1: MEAN PROPORTIONS OF CORRECT RESPONSES IN MASKED WORD IDENTIFICATION AND OF YES RESPONSES IN RECOGNITION AS A FUNCTION OF ENCODING TASK AND TEST INSTRUCTION

	Masked word identification			Recognition		
Test instruction	Generate	Read	New	Generate	Read	New
Free report	.58	.57	.31	.98	.69	.14
	(.30)	(.30)	(.23)	(.06)	(.13)	(.16)
Forced report	.60	.66	.46	.95	.63	.16
	(.25)	(.26)	(.26)	(.05)	(.20)	(.12)

Note. Standard deviations are shown in parentheses.

by replicating our procedure using Weldon's materials.

We thought that a more likely candidate for explaining the different patterns was the instructions given to subjects at the time of test. Weldon required her subjects to provide a response on every trial, with no omissions permitted. We had permitted our subjects to pass when no response came to mind. We refer to these as the *forced* and *free* response conditions, respectively. Could the different patterns of results have been caused by different response output criteria due to this instructional difference? To determine the answer, we carried out the experiment both ways.

Method

Thirty-six subjects were divided evenly over the two instructional conditions. An additional four subjects were discarded for ceiling or floor performance. Subjects in the ''free'' condition were to try to identify the word aloud, passing if they wished; subjects in the ''forced'' condition were required to give a written response on every trial, with passing not permitted.

Results and Discussion

The mean proportions of items in the Generate task that were not correctly produced during the encoding phase were .16 in the forced group and .09 in the free group. There were no errors on Read items for either group. The mean proportions of correctly identified items on the masked word identification test are shown at the left side of Table 1. A 2 × 3 analysis of variance (ANOVA) with instructional group (free, forced) and encoding task (Generate, Read, New) as factors indicated a reliable encoding task effect, $F(2,68) = 36.23$, $MS_e = 0.016$. There was no effect of instruction nor did instruction and encoding task interact, both Fs < 2.0. The encoding task effect was probed by two further ANOVAs that included instructional group as a factor. Comparison of the Generate and New conditions revealed reliable priming, $F(1,34) = 41.48$, $MS_e = 0.017$. The main effect of instruction, however, was not reliable, although the interaction between instruction and encoding task approached significance, $F(1,34) = 3.46$, $p < .08$. The second ANOVA compared performance in the Generate and Read conditions: There were no reliable effects in this analysis, all Fs < 1.3. Power to detect a difference of .09 between these two conditions (an effect size based on that obtained by Weldon, 1991, Experiment 1, as well as in Experiment 5 of the present article), was estimated to be greater than .85. Thus, we are quite confident in the equivalence of the priming in the Generate and Read conditions.

The mean proportions of hits and false alarms in the recognition test are shown at the right side of Table 1. An ANOVA applied to performance on just the studied items (hits), with the variables of instructional condition (free, forced) and encoding task (Generate, Read), indicated that subjects were reliably more successful in recognizing words from the Generate as compared to the Read condition, $F(1,34) = 108.00$, $MS_e = 0.015$. There was no reliable effect of instruction nor was there an interaction between instruction and encoding task, both *F*s < 2.0. The standard Generate advantage for direct tests replicated here.

Using Weldon's (1991, Experiment 1) definitional materials and our procedure (Masson & MacLeod, 1992) produced the same pattern of results in masked word identification as we had observed in our prior experiments, not the pattern Weldon had found. Furthermore, this was true for her stringent identification instructions as well as for our lenient identification instructions. Thus, the difference between our previous studies was not a consequence of either materials or instructions. Indeed, our pattern—equivalent priming for the Generate and Read conditions—apparently is very robust in the face of material and instruction changes. We will have to look elsewhere to explain why Weldon found less priming for generated than for read words in her study.

EXPERIMENT 2

The next step was to examine the generality of the pattern we had observed in masked word identification both in our previous work (Masson & MacLeod, 1992) and in Experiment 1. We decided to manipulate whether the Read/Generate manipulation took place within or between subjects. Weldon had blocked the encoding tasks, such that subjects always completed all of the trials of one encoding task before beginning the next one. In contrast, we had always randomized the encoding trials. Switching to a between subjects design from our standard within subject design seemed like an even more extreme design change. Would a priming difference emerge between the Generate and Read conditions in a between subjects design?

Certainly, design effects can be powerful in memory studies, as shown by Begg and his colleagues (Begg & Roe, 1988; Begg & Snider, 1987; Begg, Snider, Foley, & Goddard, 1989) in the context of a Read versus Generate manipulation. They found an advantage for Generate over Read on a recognition memory test when encoding task was manipulated within subject (a single group studied a mixed list), but not when encoding task was manipulated between subjects (different groups studied pure lists). They concluded that generating a subset of the target words inhibited processing of target words in the Read condition when the two tasks were mixed, making subjects ''lazy readers.'' Thus, we undertook to compare these two types of designs. For consistency, we continued to use Weldon's definitional materials and our ''free'' response test instructions.

Method

The 48 subjects were divided equally over the three encoding conditions: Mixed, Between-Generate, and Between-Read. An additional four subjects were discarded for ceiling or floor performance. The procedure for the Mixed study condition was identical to that in Experiment 1. For the two between subjects conditions, the only change from Experiment 1 was that subjects did only one task during encoding, either generating 40 words or reading 40 words. Subjects in the two between subjects conditions were yoked, such that corresponding subjects studied exactly the same words in exactly the same order. The 20 extra studied words for the subjects in each of the between subjects conditions were not included in the analyses.[2]

Results and Discussion

The mean proportions of words that subjects failed to generate correctly in the encoding phase were .09 in the Within condition and .09 in the Between-Generate condition. There were no errors in reading words during the encoding phase. The mean proportions of correctly identified words in the masked word identification test, conditionalized on correct responding in the encoding phase, are shown at the left hand side of Table 2. These means

[2] The pattern of results for the extra 20 words that were either read or generated in the two between-subjects conditions was very similar to that for the yoked critical items, as expected. The same was true in Experiment 3, so, in the interest of avoiding redundancy, we have not presented analyses of those data.

TABLE 2

EXPERIMENT 2: MEAN PROPORTIONS OF CORRECT RESPONSES IN MASKED WORD IDENTIFICATION AND OF YES RESPONSES IN RECOGNITION AS A FUNCTION OF ENCODING TASK AND DESIGN

	Masked word identification			Recognition		
Design	Generate	Read	New	Generate	Read	New
Within	.54	.59	.29	.97	.77	.12
	(.20)	(.21)	(.22)	(.07)	(.17)	(.14)
Between-generate	.51	—	.34	.97	—	.09
	(.31)		(.30)	(.04)		(.14)
Between-read	—	.73	.56	—	.82	.33
		(.25)	(.30)		(.13)	(.19)

Note. Standard deviations are shown in parentheses.

indicate that there was a similar amount of priming for words in the Generate and Read conditions in both designs.

Priming effects were tested using a set of ANOVAs. First, an ANOVA on proportion correct in the within subject condition found a reliable difference among the three encoding conditions, $F(2,30) = 24.36$, $MS_e = 0.018$. Planned comparisons showed that there was a reliable priming effect in the Generate condition when compared to the New condition, $F(1,15) = 19.12$, $MS_e = 0.027$, but that the difference between the Generate and the Read conditions was not significant, $F(1,15) = 2.39$, $MS_e = 0.009$. Using the .09 effect size described in Experiment 1, power to detect such a difference between these two conditions was estimated to be greater than .75.

Second, a mixed factor ANOVA with encoding task (Generate, Read) as a between subjects factor and prior exposure (Generate/Read, New) as a repeated measures factor was applied to the data from the between subjects design. This analysis indicated that, averaging across old and new items, proportion of correct identification was higher in the Read group than in the Generate group, $F(1,30) = 7.75$, $MS_e = 0.142$, and that more old items (Generate/Read) were identified than New items, $F(1,30) = 44.35$, $MS_e = 0.016$. The interaction between encoding task and prior exposure was not significant, $F(1,30) = 1.61$, $p > .20$, indicating that priming was similar in the Generate and Read groups. Because this was a between subjects comparison, power to detect a difference of .09 between these two conditions was lower, estimated at .45.

Two matters regarding the indirect test data in Experiment 2 warrant further comment. First, it was in the within subject data of Experiment 2 that we observed the only instance of a difference between the analyses conditionalized on correct generation and the corresponding unconditionalized analyses. The three unconditionalized means were .59 for Read, .50 for Generate, and .29 for New. This time there was a reliable difference in priming favoring Read over Generate, $F(1,15) = 7.17$, $MS_e = 0.010$. Because this apparent Read advantage did not arise in the unconditional analyses of the 11 experiments reported by Masson and MacLeod (1992) nor in any of the other experiments reported in the present article, we are inclined to see it as uninformative, particularly given the arguments made in the General Method section.

The second matter that warrants comment is the considerably better performance of the Between-Read group in contrast to both the Between-Generate and the Within groups. Note that this was true for both the studied items and the new, unstudied items. Although obviously speculative, we offer a possible explanation. It may be that the uninterrupted practice at reading words on the screen during study provided a general advantage to all items on the indirect masked word identification test. After all, masked word identification is a sort of difficult reading test. The difficulty of reading the masked words may have been reduced when the reading practice during study was twice as long as in prior studies and uninterrupted by Generate trials.

Consider now recognition test performance, which is summarized as mean hit and false alarm rates on the right side of Table 2. As usual, in both designs, subjects were more likely to recognize generated than read words. For the within subject design, an ANOVA found the Read and Generate hit rates to be reliably different, $F(1,15) = 20.44$, $MS_e = 0.016$. For the between subjects design, encoding task (Generate, Read) was a between subjects factor and prior exposure (old, new) was a within subject factor. New words were included in the analysis to take into account different false alarm rates in the two encoding groups; the difference between old and new words was taken as a measure of accuracy. There was no main effect of encoding task, but there were reliably more ‘‘yes’’ responses to old than to new items, $F(1,30) = 416.44$, $MS_e = 0.018$. The interaction between encoding task and prior exposure was significant, $F(1,30) = 31.29$, indicating that the difference between hits and false alarms was reliably greater in the Generate than in the Read condition.

The results of the masked word identification task indicate that manipulation of encoding task as a between subjects or a within subject variable does not alter the relative amount of priming brought about by Read and Generate tasks. This outcome is different from that of Begg and his colleagues (Begg & Roe, 1988; Begg & Snider, 1987; Begg et al., 1989), who found an advantage for Generate over Read on a recognition memory test for mixed lists (within subject) but not for pure lists (between subjects). Given this result, one might expect that the Read condition would yield lower priming in the within subject condition of Experiment 2. On the other hand, it is not surprising that a different result occurred here with the masked word identification task because identification and recognition tests must involve at least some different processes.

Another contrast between Begg’s studies and ours is that we used conceptual generation cues at the time of encoding, whereas Begg used word fragments in his experiments. This difference may be crucial because even our recognition test results did not look like Begg’s: Recognition accuracy in the present experiment was greater for Generate than for Read words, even when this comparison was made between subjects. The combination of Read and Generate tasks used here, then, does not appear to produce the inhibited reading process described by Begg and his colleagues.

Experiment 2 demonstrates that our pattern in masked word identification, in which Generate and Read show equivalent priming, occurs under both of the standard design conditions. Using definitions as generation cues, we have yet to observe the Read > Generate > New pattern that Weldon (1991) reported for both word fragment completion and masked word identification.

Experiment 3

Experiment 3 was an exact replication of Experiment 2 except for one change: Here, word fragment completion replaced masked word identification. If the pattern for word fragment completion is different from that for masked word identification, we should now see Weldon’s pattern. There are a great many studies to suggest differential priming favoring Read over Generate in word fragment completion (see Roediger & McDermott, 1993, for a review), so we fully expected to see this pattern when we switched to that test, indicating that the tests do differ, and that it is not some feature of how these experiments are done in the two laboratories that is critical.

Method

As in Experiment 2, 16 subjects were assigned to each of the encoding conditions: Mixed, Between-Generate, and Between-Read. An additional subject was discarded for floor performance. The procedure was identical to that in Experiment 2 except that the indirect test was changed from masked word identification to word fragment completion, using the word fragments shown in the Appendix.

TABLE 3

EXPERIMENT 3: MEAN PROPORTIONS OF CORRECT RESPONSES IN WORD FRAGMENT COMPLETION AND OF YES RESPONSES IN RECOGNITION AS A FUNCTION OF ENCODING TASK AND DESIGN

	Word fragment completion			Recognition		
Design	Generate	Read	New	Generate	Read	New
Within	.39	.53	.23	.95	.67	.12
	(.18)	(.15)	(.08)	(.09)	(.17)	(.10)
Between-generate	.37	—	.25	.94	—	.13
	(.13)		(.12)	(.07)		(.11)
Between-read	—	.62	.30	—	.82	.21
		(.13)	(.13)		(.10)	(.17)

Note. Standard deviations are shown in parentheses.

Results and Discussion

The mean proportions of words generated incorrectly in the encoding phase were .11 in the within subject case and .07 in the Between-Generate case. One subject made a single error in the Between-Read condition. The mean proportions of fragments correctly completed on the word fragment completion test, conditionalized on correct responding during encoding, are shown at the left of Table 3. The pattern of means was the same for the within subject and the between subjects designs: Although priming was obtained in the Generate condition, considerably greater priming was observed in the Read condition. This is the familiar pattern for word fragment completion, in sharp contrast to what we have been observing for masked word identification.

An ANOVA on proportion correct in the within subject condition found a reliable difference among the three encoding conditions, $F(2,30) = 27.68$, $MS_e = 0.013$. Subsequent planned comparisons indicated that there was reliable priming in the Generate condition when compared to the New condition, $F(1,15) = 10.20$, $MS_e = 0.016$, and that the completion rate was reliably greater in the Read than in the Generate condition, $F(1,15) = 18.06$, $MS_e = 0.011$. A mixed factor ANOVA with encoding task (Generate, Read) as a between subjects factor and prior exposure (Generate/Read, New) as a repeated measures factor was used to compare priming effects in the between subjects condition. This analysis indicated that, averaging across old and new items, successful fragment completion was more likely in the Read group than in the Generate group, $F(1,30) = 18.32$, $MS_e = 0.020$, and that old items (generated or read) were completed more frequently than new items, $F(1,30) = 60.75$, $MS_e = 0.012$. There was also an interaction between encoding task and prior exposure, indicating that the effect of prior exposure (priming) was reliably greater in the Read group, $F(1,30) = 12.92$. Given that there was more priming in the Read group than in the Generate group, an additional test for priming in the Generate group was conducted. This test found a reliable priming effect, $F(1,15) = 6.42$, $MS_e = 0.017$.

The mean hit and false alarm rates on the recognition test for each design are shown at the right in Table 3. These means indicate that subjects were more likely to recognize generated than read words in both designs, as usual. This was confirmed for the within subject design by an ANOVA comparing Read and Generate scores, $F(1,15) = 55.61$, $MS_e = 0.011$. For the between subjects design, an ANOVA with encoding task (Generate, Read) and prior exposure (old, new) was used to assess recognition performance, following the logic of Experiment 2. The ANOVA found no overall difference due to encoding task, $F < 1$, but there were reliably more ‘‘yes’’ responses to old than to new items, $F(1,30) = 672.43$, $MS_e = 0.012$. There also was an interaction between encoding task and prior exposure, indicating that the difference between hits and false alarms was reliably greater in the Generate than in the Read condition, $F(1,30) = 12.50$.

The results of Experiment 3 stand in clear contrast to those of Experiment 2 and most of our prior studies using the masked word identification test. We had absolutely no difficulty producing the gradient pattern typically seen in word fragment completion, with more priming for words that subjects had read than for words that they had generated, although there was still reliable priming even for the generated words. Thus, Weldon's different pattern in masked word identification is not simply the result of different laboratories producing different results.

EXPERIMENT 4

In Experiments 1 and 2, as in our prior work using masked word identification (Masson & MacLeod, 1992), we observed equivalent priming for read and for generated words. However, in all of those experiments, we had mixed the two tasks randomly during study. Weldon (1991) blocked her study tasks, having subjects do all of the trials of one task before any of the trials of another task. Could blocking account for the different patterns in her data and in ours?

Blocking certainly does have powerful effects elsewhere in the memory literature. As a recent illustration, Thapar and Greene (1994) showed that the advantage of semantic processing over nonsemantic processing was greater when the conditions were blocked than when they were mixed at the time of encoding on several indirect measures of memory, including word fragment completion (Experiment 1) and masked word identification (Experiment 2). This design manipulation, however, had no effect on direct measures such as recognition (Experiment 4) and recall (Experiments 5 and 6). The present experiment directly examined this design issue by blocking the Read and Generate tasks during study prior to masked word identification.

Method

Eighteen subjects took part in the study. The data from two additional subjects were discarded due to performance near the ceiling or the floor. The 20 Read and 20 Generate trials were now blocked during study, with 9 subjects receiving each block order during study.

Results and Discussion

The mean proportion of items not correctly produced during the encoding phase in the Generate task was .12; there were no errors on Read items. On the left side, Table 4 displays the mean proportions of correctly identified items in the masked word identification test. A repeated measures ANOVA on encoding task (Generate, Read, New) indicated a reliable effect, $F(2,34) = 22.69$, $MS_e = 0.020$. To unpack this overall effect, we did two further ANOVAs. The first compared the Generate and New conditions, confirming reliable priming in the Generate condition, $F(1,17) = 26.51$, $MS_e = 0.026$. The second ANOVA compared performance in the Generate and Read conditions, confirming that they did not differ, $F < 1$. Power to detect a difference of .09 between these two conditions was estimated to be .59.

We also did an analysis of the masked word identification data including block order as a variable (2 blocks by 3 test conditions). The point of this analysis was to consider the possibility of processing ''leakage'' between the two blocks, particularly when Read precedes Generate. When the first block involves reading, subjects may, during the second block (generation), image the word as if it had been read; this would be unlikely to happen when the Generate block came first because there would be no experience with reading in the experiment. If such contamination were occurring during generation in the Read then Generate order, this could help to compensate for the ''true'' smaller priming in the Generate condition, according to a view such as that of Roediger (1990). The upshot is that we found no evidence of ''leakage'': The main effect of block order was unreliable, $F < 1$, as was

TABLE 4

EXPERIMENT 4: MEAN PROPORTIONS OF CORRECT RESPONSES IN MASKED WORD IDENTIFICATION AND OF YES RESPONSES IN RECOGNITION AS A FUNCTION OF ENCODING TASK

Masked word identification			Recognition		
Generate	Read	New	Generate	Read	New
.62	.62	.34	.92	.62	.13
(.26)	(.29)	(.24)	(.12)	(.20)	(.10)

Note. Standard deviations are shown in parentheses.

its interaction with test condition, $F(2,32) = 2.17$, $MS_e = 0.019$, $p = .13$. If anything, the trend in the interaction suggested that the task done first at study benefitted more at test.

Turning now to the recognition test data, the mean proportions of hits and false alarms are shown on the right side of Table 4. An ANOVA applied to performance on just the studied items (hits) indicated that subjects were reliably better at recognizing words from the Generate as opposed to the Read condition, $F(1,17) = 29.73$, $MS_e = 0.027$.

Experiment 4 makes clear that the priming effects observed both here and in our prior work (Masson & MacLeod, 1992) do not hinge on whether the encoding tasks are intermingled or separated. Blocking encoding did not alter our basic pattern of equivalent priming in masked word identification for the Read and Generate conditions. Unlike the case of type of processing effects (Thapar & Greene, 1994), blocking versus mixing these two conditions is not the critical difference between Weldon's (1991) finding and ours. What, then, does make the difference?

EXPERIMENT 5

Experiments 1–4 showed that basic design differences, materials, and reporting instructions are not responsible for the differences in masked word identification between our results and those of Weldon (1991). An oversight on our part became apparent, however, in a further experiment that we had originally conducted for other reasons. In Experiment 5, for the first time, *three* different encoding tasks are included. Two of the tasks were the same as those in our previous experiments: generate from a brief definition and read aloud. The third task required subjects to read a word aloud and then to say the first word that came to mind (i.e., provide an associate). We refer to this task as the Associate task. Given that we were using three different encoding tasks, two of which involved presentation of isolated words, we decided to present the tasks in a blocked format to reduce the possibility of confusion regarding which task ought to be performed on each trial.

Because this experiment originally was designed for another purpose, the materials and procedures were not identical to those in Experiments 1–4, although they were quite similar. The outcome of the experiment, however, turned out to be crucial for the issue of why Weldon (1991, Experiment 1) and Masson and MacLeod (1992) obtained different patterns of priming in masked word identification. As it happens, in addition to blocking her encoding conditions, Weldon also always used more than two encoding conditions, including in addition to generation and reading other encoding tasks involving auditory presentation and picture naming. Although unanticipated, her choice of this combination of number and blocking of encoding conditions was critical.

Method

Thirty-six subjects took part in the study. The data from six additional subjects were not included because performance was at ceiling or floor. Order of presentation of the three encoding task blocks was counterbalanced. The Generate and Read conditions were set

TABLE 5

EXPERIMENTS 5 AND 6: MEAN PROPORTIONS OF CORRECT RESPONSES IN MASKED WORD IDENTIFICATION AND IN RECALL AS A FUNCTION OF ENCODING TASK

	Masked word identification				Recall			
Design	Generate	Read	Assoc.	New	Generate	Read	Assoc.	New
Experiment 5								
Blocked	.58	.67	.64	.53	.26	.07	.29	.01
	(.24)	(.24)	(.24)	(.25)	(.12)	(.07)	(.14)	(.00)
Experiment 6								
Mixed	.65	.66	.72	.57	.28	.06	.22	.02
	(.22)	(.25)	(.24)	(.26)	(.14)	(.08)	(.10)	(.06)

Note. Standard deviations are shown in parentheses. ''Assoc.'' refers to the condition where subjects first read a word and then produced an associate to that word.

up as in the earlier experiments. Associate trials were set up similarly to Read trials, except that subjects said the first word that came to mind after reading aloud the target word.

Results and Discussion

The mean proportion of words that subjects failed to produce in the Generate encoding task was .04. The mean proportions of correctly identified targets in the masked word identification task are shown on the top left side of Table 5. An ANOVA with encoding task (Generate, Read, Associate, and New) as a within subject factor indicated that there were reliable differences among the four encoding conditions, $F(3,105) = 15.10$, $MS_e = 0.010$. Planned comparisons showed that the Generate condition produced reliable priming relative to the New condition, $F(1,35) = 4.70$, $MS_e = 0.010$, but that identification accuracy was higher in the Read than in the Generate condition, $F(1,35) = 13.55$, $MS_e = 0.011$. A final comparison showed that there was no reliable difference between the Read and Associate conditions, $F < 1.1$.

As in Experiment 4, the possibility that encoding task effects were influenced by the order of encoding task presentation was assessed with an additional ANOVA that included order of encoding task as a factor. There was no reliable effect of encoding task order, $F(2,33) = 1.48$, $MS_e = 0.197$, nor any interaction between order and encoding task, $F < 1$. Moreover, subjects who performed the Generate task after the Read task identified reliably more Read than Generate items, $F(1,11) = 6.36$, $MS_e = 0.013$. Thus, Experiment 5 provides no evidence that identification of Generate items benefits from leakage of processing operations from the Read to the Generate task.

The mean proportions of items reported on the recall test, now the direct test of memory, are shown at the top right side of Table 5. Targets reported in the New condition were considered intrusions because subjects were instructed to recall only items that had appeared in the encoding phase of the experiment. An ANOVA comparing performance in the three conditions involving studied items showed that they differed reliably, $F(2,70) = 57.78$, $MS_e = 0.009$. Pairwise comparisons indicated that both the Generate and the Associate items were better recalled than the Read items, $F(1,35) = 79.62$, $MS_e = 0.008$, and $F(1,35) = 80.65$, $MS_e = 0.011$, respectively. These results show that both of the conceptually driven encoding tasks led to better recall than did the reading task, the standard pattern for direct tests.

An implication of the finding that the Read and Associate conditions produced similar amounts of priming, whereas the Generate condition produced less priming, is that the conceptual basis for priming effects was weakened in Experiment 5, relative to the prior experiments. We have argued that the brief presentation of a whole word can make contact with memory for conceptual processing carried out during the study phase. The results of Experiment 5 suggest that there are circumstances under which such contact fails to occur, although the exact causes of that failure are unclear.

Experiment 5 is our first to obtain greater priming of masked word identification after reading items than after generating them from the type of cues used here. We used these same materials in an earlier study (Masson & MacLeod, 1992, Experiment 1), and similar materials in Experiments 1–4, all of which produced about as much or more priming after generation as after reading. What sets Experiment 5 apart from these earlier experiments is the inclusion of three different encoding tasks during study. This design is very similar to that used by Weldon (1991, Experiment 1), in which four different encoding tasks were presented in blocked format during study. Weldon also found reliably more priming in masked word identification in the Read condition than in the Generate condition. The similarity in these two outcomes raises the possibility that differential priming in the Read and Generate conditions may be associated with blocked presentation of at least three different encoding tasks. Experiment 6 examines this possibility.

Experiment 6

If better performance in the Read as compared with the Generate condition was due to the blocked presentation of at least three encoding tasks, then mixing the three encoding tasks used in Experiment 5 should result in the pattern we have typically found in the past: similar performance in the Generate and Read conditions. Therefore, Experiment 6 was a replication of Experiment 5 with the only change being mixed rather than blocked presentation of the three encoding tasks.

Method

Twenty-four subjects took part in this experiment. Two additional subjects were tested but their data were not included in the analyses because their performance levels were at ceiling or floor. In the encoding phase, subjects were now given instructions for all three encoding tasks before any trials began and were informed at the start of each trial which task to perform on that trial.

Results and Discussion

The proportion of Generate trials on which subjects failed to produce the correct target during encoding was .03. The mean proportions of correct responses on the masked word identification and recall tasks are shown at the bottom of Table 5. An ANOVA of the identification data indicated that there were differences among the four encoding conditions, $F(3,69) = 14.28$, $MS_e = 0.006$. Planned comparisons were conducted as in Experiment 5, and showed that there was reliable priming in the Generate condition relative to the New condition, $F(1,23) = 16.22$, $MS_e = 0.005$, and that performance in the Generate and Read conditions was not reliably different, $F < 1$. Power to detect a difference of .09 between these two conditions was estimated to be greater than .99. In addition, the advantage in the Associate condition relative to the Read condition approached significance, $F(1,23) = 4.26$, $MS_e = 0.010$, $p < .06$.

For the direct recall test, the pattern was as in Experiment 5. There were reliable differences among the three encoding conditions, $F(2,46) = 38.32$, $MS_e = 0.008$. Pairwise comparisons showed that items in the Generate and in the Associate conditions were recalled more often than items in the Read condition, $F(1,23) = 54.45$, $MS_e = 0.011$, and $F(1,23) = 52.03$, $MS_e = 0.006$, respectively.

The masked word identification results of Experiment 6 replicated our typical finding: Read and Generate conditions produced very similar amounts of priming. The only difference between Experiments 5 and 6 was in how the three different encoding tasks were presented. When they were presented in blocked format (Experiment 5), our pattern of priming effects replicated that of Weldon (1991, Experiment 1), in which reading led to more priming than generating.

This was the only instance in all of our experiments using masked word identification with definitions as generation cues where we obtained her pattern. Furthermore, shifting to a mixed encoding format in Experiment 6, but making no other procedural changes, brought the levels of priming in the Generate and Read conditions back into line with one another.

This change in amount of priming seen in the three encoding conditions across the last two experiments was examined by an ANOVA with encoding task (Generate, Read, Associate, and New) and Experiment (5 and 6) as factors. There was no main effect of Experiment, $F < 1$, but there was a significant main effect of encoding task, $F(3,174) = 23.57$, $MS_e = 0.008$, and a significant interaction, $F(3,174) = 2.94$. The reliable interaction confirms that the amount of priming changed across experiments. Three additional ANOVAs were computed to determine where the change occurred. These ANOVAs compared each of the Generate, Read, and Associate encoding conditions in turn to the New condition, with Experiment as a second factor. All of these ANOVAs found a reliable priming effect and no effect of Experiment, but the question of interest was whether the priming effect interacted with Experiment. The ANOVAs comparing Generate to New and Associate to New found no interaction, Fs < 1, indicating that the amount of priming produced by these two encoding tasks was not significantly different in the two experiments. In contrast, the interaction between the Read vs. New comparison and Experiment did approach significance, $F(1,58) = 3.04$, $MS_e = 0.008$, $p < .10$, supporting the conclusion that priming was greater in the Read condition when a blocked design was used rather than a mixed design.

General Discussion

In this series of experiments, we have replicated a number of fundamental results involving relative amounts of priming on indirect tests of memory arising from read and generation encoding tasks. First, we replicated our finding (Masson & MacLeod, 1992) of similar amounts of priming for these two encoding tasks on the masked word identification task. Second, we replicated the finding that reading leads to more priming than generating on the word fragment completion task (Blaxton, 1989; Srinivas & Roediger, 1990; Weldon, 1991). Third, under conditions similar to those used by Weldon (1991, Experiment 1), we replicated her finding of more priming on masked word identification after reading than after generating. In obtaining two different patterns of results with the masked word identification task, we demonstrated that differential priming for the Read and Generate encoding conditions occurs only under rather circumscribed conditions. In conjunction with the Weldon results, these findings allow the following conclusion: For the Generate task used by Weldon and by us, only when the Read and Generate encoding tasks are combined with at least a third encoding task—and the resulting collection of tasks is presented in a blocked format—does an advantage of Read over Generate appear in masked word identification.

To conduct an especially powerful test of the possibility that Generate and Read conditions produce different amounts of priming under design parameters that vary from those adopted by Weldon (1991, Experiment 1) and by us in Experiment 5, we combined the data from those two conditions across Experiments 1, 2 (including data only from the within subject condition), 4, and 6 for a total sample size of 94 subjects. The mean identification proportions for the Generate and Read conditions (.60 vs. .62) were not reliably different, $F(1,93) = 1.90$, $MS_e = 0.011$. The power of this analysis to detect a difference of .09 (equal to that found by Weldon in Experiment 1 and by us in Experiment 5), was greater than .99.

Taken together with our earlier work comparing Read and Generate tasks (Masson & MacLeod, 1992), the experiments reported here show that for the masked word identification task the more general finding is similar amounts of priming with the two encoding tasks. The results reported by Weldon

(1991, Experiment 1) represent an exception that appears to depend on a particular confluence of design features. Our earlier experiments, however, also showed that the Read task can yield greater priming than certain Generate tasks, particularly those that induce subjects to integrate target words with their generation cues (Masson & MacLeod, 1992, Experiments 2, 6, 8, and 9; see also Jacoby, 1983b). In the Masson and MacLeod study, we argued that integrated encoding of a target and its generation cue made it less likely that the target word would recruit the earlier encoding episode when that target was later tested in isolation.

Our finding that word fragment completion and masked word identification yield different patterns of read/generate effects poses a theoretical puzzle, but it is certainly not an anomalous outcome. Earlier research obtaining dissociations between these two tasks led investigators to suggest that these two tasks may depend on different processing operations (Schwartz, 1989; Witherspoon & Moscovitch, 1989). To understand why word fragment completion and masked word identification should produce different patterns of results with Read and Generate encoding tasks, we consider next how presentation of a word fragment as opposed to a masked but complete target word might recruit memory for earlier encoding episodes differently. We then turn to the question of why finding different amounts of identification priming in Generate and Read conditions depends on using blocked presentation of multiple encoding tasks.

Recruitment of Processing Episodes

Earlier accounts of priming in masked word identification following prior encoding episodes have emphasized the overlap in perceptual processes required by the encoding and test tasks (e.g., Jacoby, 1983a, 1983b; Jacoby & Dallas, 1981; Reinitz & Alexander, 1996; Weldon, 1991). We have proposed that, in addition to extracting perceptual information, the brief availability of the entire target word in the masked word identification task is sufficient to make contact with relevant conceptual episodes in memory (Masson & MacLeod, 1992, 1996). This is a reasonable claim given demonstrations of semantic or associative priming effects obtained under binocular masking conditions such as those used in masked word identification (Carr & Dagenbach, 1990; Dagenbach, Carr, & Wilhelmsen, 1989; de Groot, 1983; Lukatela & Turvey, 1994) and the tendency to report semantically related intrusions (e.g., reporting *jazz* instead of the target word *blues*) when attempting to identify briefly presented target words (Allport, 1977). Building on these findings, we suggest that a briefly presented target can make contact with an earlier episode in which the word was generated from a conceptual cue, but never physically seen.

In contrast to the masked word identification task, in which the entire target word is visible, a word fragment or even a three-letter word stem, as used in a substantial number of studies of indirect tests of memory (e.g., Toth, Reingold, & Jacoby, 1994), may not provide enough constraints for conceptual knowledge about the target word to be retrieved. If word fragments or stems are indeed less likely to recruit conceptually-based encoding episodes, then the typical advantage of Read over Generate encoding conditions seen on these tasks, and replicated here with word fragments, is easily explained.

Our claim regarding the differential ability of word fragments or stems on the one hand, and briefly presented whole words on the other hand, to retrieve conceptual knowledge relevant to the target word runs counter to a proposal by Keane, Gabrieli, Fennema, Growdon, and Corkin (1991). In a study comparing patients with Alzheimer's disease to normal controls, they found that the two groups produced equivalent priming on a variant of the masked word identification task for target words that they had read earlier, whereas the patients were impaired in the amount of priming obtained on a word stem completion task. Keane et al. attributed this pattern to (1) preserved perceptual processing, which they assumed was responsible for priming in the identification task, and (2) compromised conceptual processes that ordinarily contributed to priming on the stem completion task.

We have two concerns regarding the Keane et al. (1991) position. First, the proposition that priming in word stem completion is driven by a substantial, automatic conceptual component is contradicted by the Toth et al. (1994) results based on that very task. Their analysis indicates no unconscious influence of conceptually driven encoding processes to a subsequent word stem completion task. Second, even if the impaired performance of Alzheimer's patients on the stem completion task were due to conceptual processing deficits, it is not clear that the compromised conceptual processes are the same ones that contribute to enhanced performance on the masked word identification task. There is substantial evidence (noted earlier) supporting the view that a briefly presented target word is capable of automatically retrieving relevant conceptual knowledge. Furthermore, this function appears to be spared in Alzheimer's patients, inasmuch as normal semantic priming effects have been obtained with these patients (e.g., Nebes, Martin, & Horn, 1984). Thus, preserved priming in masked word identification with Alzheimer's patients may be produced, in part, by the same conceptual processing operations that we propose contribute to priming effects in normal subjects.

When Recruitment of Prior Episodes Fails

We have emphasized the differences between word fragment completion and masked word identification with respect to priming induced by prior processing episodes. It is important to realize, however, that the two tasks did yield similar outcomes under a circumscribed set of conditions for the identification task. Presentation of three different encoding tasks in blocked format led to more priming in the Read than in the Generate encoding condition on the identification task (Experiment 5). This result replicates the earlier finding by Weldon (1991, Experiment 1) in which four different encoding tasks were presented in separate blocks. Taken in isolation, the results of Experiment 5 suggested that priming in the Generate condition had been reduced relative to the Read condition. By considering the comparison between the results of Experiments 5 and 6, however, it appears that the impact of mixed versus blocked presentation of encoding tasks is actually on performance in the Read condition. In particular, mixed presentation of encoding tasks appears to reduce identification of Read words.

In an earlier article, we considered the possibility that the degree of conceptual processing applied to Read items might be affected by mixed versus blocked presentation of encoding tasks (Masson & MacLeod, 1992, p. 164). Following the proposal by Begg and Snider (1987), we suggested that mixing read and generated items in the same list might cause subjects to engage in less discriminative encoding of read words (i.e., to become ''lazy readers''). In the interpretive/elaborative encoding framework we proposed, this reduced discriminative encoding would potentially affect both types of encoding. Thus, even during the initial interpretive encoding of an item, subjects might fail to distinguish adequately between the target item and other items that might be recruited during the read aloud task by virtue of their conceptual or orthographic relation to the target.

If interpretive encoding in the Read task were compromised in such a way, we would expect reduced priming on a subsequent masked word identification task. This outcome would be expected because it is memory for the interpretive encoding that we propose underlies priming in that task (Masson & MacLeod, 1992). On this account, then, finding similar amounts of priming on read and generated items would be due to depressed performance on the Read items. That is, the lack of discriminative encoding of a Read item would reduce the likelihood that its encoding episode would be recruited when the target word was presented on the masked word identification test.

Although this account seems promising, it fails to explain why Read did not produce greater identification accuracy than Generate when encoding task was manipulated between subjects (Experiment 2) or when only two encoding tasks were blocked (Experiment 4). Further work will be needed to assess fully this account of the changing pattern of priming effects. We have, however, established two important points in the debate regarding indirect memory for episodes involving Read and Generate tasks, as revealed on the masked word identification task. First, we have established what procedural difference underlies the discrepancy in priming patterns found by Weldon (1991, Experiment 1) and ourselves (Masson & MacLeod, 1992). Second, we have preliminary evidence from the comparison between Experiments 5 and 6 that the effect of mixed versus blocked presentation of encoding tasks is to alter the amount of priming observed among Read items.

Intentional Retrieval

The experiments reported here were not designed to address the question of whether, or to what extent, intentional retrieval strategies influence the amount of priming found in the masked word identification task. We have worked from the assumption that influences from that quarter are minimal. Support for this assumption comes from studies that have produced dissociations between the masked word identification task and direct tests of memory (e.g., Allen & Jacoby, 1991; Jacoby & Dallas, 1981; Jacoby 1983a, 1983b). In our earlier work, we have also obtained dissociations between masked word identification and recognition memory (Masson & MacLeod, 1992, Experiments 5, 6, and 8). Moreover, Richardson-Klavehn, Lee, Joubran, and Bjork (1994) showed that dissociations between indirect and direct tests of memory are attributable to differences in intentional recollection, even though subjects may be aware of prior occurrence of target words on the indirect test. Thus, awareness of past occurrence does not necessarily imply that indirect test performance is driven by deliberate attempts to remember previously encoded items. Finally, the change in amount of priming found in the Read encoding condition across Experiments 5 and 6 cannot plausibly be ascribed to intentional retrieval strategies; were such strategies involved in that change, performance in the Generate and Associate conditions should have been affected as well, perhaps even more. Therefore, differential application of intentional retrieval strategies does not constitute a viable account of why Read and Generate encoding conditions sometimes produce similar amounts of priming and why Read encoding sometimes produces more priming.

Conclusion

We have argued that the difference between masked word identification and word fragment completion with respect to the effects of Read versus Generate tasks is due to differences in automatic recruitment of conceptual knowledge. Our view is that the availability of even a brief display of an entire word enables more effective conceptually based recruitment of prior processing episodes than does presentation of only part of a word, as is the case for word fragments. Furthermore, we maintain that the presence of conceptual priming on an indirect test does not imply contamination of that test by conscious recollection. Rather, we hold that the recruiting of conceptual aspects of encoding is a normal feature of masked word identification functioning as an indirect test.

We also have established that priming on the masked word identification task brought about by reading target words can rise above the level of priming achieved by generating words. This occurs under a highly specific set of conditions (blocked presentation of at least three encoding tasks). Although we have been unable to develop a satisfying explanation as to why these particular conditions render the read encoding task more effective than the generation task, we have identified a candidate account for the shifting fortunes of Read and Generate tasks that offers guidance for further exploration of how processing fluency is enhanced by memory for prior events. We have also made it clear that the normal pattern in masked word identification is for priming to be equivalent for read and generated words.

APPENDIX

THE 60 TARGET WORDS, FRAGMENTS, AND DEFINITIONS USED IN EXPERIMENTS 1–4 (FROM WELDON, 1991)

Word	Fragment	Definition
ambulance	a--ula-c-	paramedic's vehicle - a?
arrow	-r--w	an archer shoots a bow and - a?
ashtray	--ht-ay	place to put a cigarette butt - a?
balloon	--l-oo-	helium - b?
bicycle	--c--le	10-speed - b?
cactus	-ac-u-	desert plant - c?
camera	-ame-a	a photographer takes a picture with a - c?
canoe	---oe	Indians paddle a - c?
carrot	c---ot	Bugs Bunny ate this orange vegetable - c?
diamond	-ia--nd	gem in an engagement ring - d?
elephant	-lep--n-	a large grey animal with tusks - e?
envelope	-nve---e	you put a stamp on an - e?
escalator	-sca--t--	moving stairs - e?
fireplace	-ir-p-a-e	you burn logs in this in the winter to get warm - f?
flower	f--we-	daisy - f?
football	-oo--al-	autumn sport - f?
giraffe	--r-f-e	African animal with a long neck - g?
grasshopper	-ra-s-o-p--	green insect, related to locusts and katydids - g?
helicopter	-e-ico---r	an aircraft that can fly straight up and down - h?
igloo	i-lo-	Eskimo's ice house - i?
intestines	in--stin-s	where food goes after digestion in the stomach - i?
kangaroo	-a-g-r-o	Australian animal that has a pouch and hops - k?
lobster	l-bs---	this sea animal has claws (similar to a crab) - l?
microscope	-i-ros-op-	instrument used to magnify blood cells - m?
motorcycle	-oto--y---	a Harley-Davidson is a type of - m?
mountain	m-un--in	larger than a hill - m?
needle	n--d-e	this is difficult to find in a haystack - n?
newspaper	-ew--ape-	the New York Times is one - n?
octopus	--topu-	a sea animal with eight tentacles - o?
ostrich	-s-ri-h	a bird with valuable feathers that buries its head in the sand - o?
parachute	-ar-ch-t-	sky-divers need a - p?
peanut	-e-nu-	snack food that can be dry-roasted and/or salted - p?
pencil	-e-c-l	an eraser is on the end of a - p?
penguin	pe--ui-	an Antarctic bird that wears a tuxedo - p?
piano	-i--o	baby grand - p?
pyramid	-yr-mi-	a large, triangular structure in Egypt - p?
refrigerator	-efri---ato-	an appliance that keeps food cool - r?
rhinoceros	--in-ce-os	an animal that resembles a hippopotamus, but has a horn - r?
sandwich	s--d-i-h	ham and cheese on rye - s?
saxophone	sa-op-on-	a jazz instrument, related to the trumpet and trombone - s?
scissors	--isso--	you cut paper with these - s?

APPENDIX—*Continued*

Word	Fragment	Definition
screwdriver	s--ew-ri-e-	tool for twisting bolts into the wall - s?
skunk	-ku-k	animal that sprays an awful odor - s?
snowman	-no--an	Frosty the - s?
squirrel	s-ui-re-	animal that stores acorns - s?
submarine	--bm-ri-e	a vessel that travels underwater and has a periscope - s?
suitcase	--itc-se	luggage - s?
sweater	-w-ate-	knit a wool - s?
telescope	--le-c-pe	instrument for star-gazing - t?
television	-e-ev--i-n	you watch prime time shows on - t?
thermometer	-her--m-te-	instrument containing mercury for measuring temperature - t?
toaster	to-s--r	applicance for making sliced bread warm and crisp in the morning - t?
toothbrush	to-t--rus-	object for oral hygiene - t?
tornado	-orn-d-	another name for a cyclone or twister - t?
turkey	t-r--y	main course on Thanksgiving - t?
typewriter	t--e-rit--	secretaries have this piece of equipment - t?
umbrella	-m-re-l-	this protects you from the rain - u?
unicorn	-n-cor-	a mythical animal with a horn on its forehead - u?
violin	-io--n	string instrument played by Itzhak Perlman - v?
watermelon	-ate--elo-	a summer fruit related to a cantaloupe - w?

REFERENCES

ALLPORT, D. A. (1977). On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.), *Attention and performance VI* (pp. 505–533). Hillsdale, NJ: Erlbaum.

BEGG, I., & ROE, H. (1988). On the inhibition of reading by generating. *Canadian Journal of Psychology,* **42,** 325–336.

BEGG, I., & SNIDER, A. (1987). The generation effect: Evidence for generalized inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **13,** 553–563.

BEGG, I., SNIDER, A., FOLEY, F., & GODDARD, R. (1989). The generation effect is no artifact: Generating makes words distinctive. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15,** 977–989.

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.

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.

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.

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.

GRAVES, R., & BRADLEY, R. (1987). Millisecond interval timer and auditory reaction time programs for the IBM PC. *Behavior Research Methods, Instruments, and Computers,* **19,** 30–35.

GRAVES, R., & BRADLEY, R. (1988). More on millisecond timing and tachistoscope applications for the IBM PC. *Behavior Research Methods, Instruments, and Computers,* **20,** 408–412.

JACOBY, L. L. (1983a). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **9,** 21–38.

JACOBY, L. L. (1983b). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior,* **22,** 485–508.

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., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General,* **118,** 126–135.

KEANE, M. M., GABRIELI, J. D. E., FENNEMA, A. C., GROWDON, J. H., & CORKIN, S. (1991). Evidence for a dissociation between perceptual and conceptual priming in Alzheimer's disease. *Behavioral Neuroscience,* **105,** 326–243.

KELLEY, C. M., & JACOBY, L. L. (1990). The construction of subjective experience: Memory attributions. *Mind & Language,* **5,** 49–68.

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.

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.

MASSON, M. E. J., & MACLEOD, C. M. (1992). Reenacting 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.

MORRIS, C. D., BRANSFORD, J. D., & FRANKS, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior,* **16,** 519–533.

NEBES, R. D., MARTIN, D. C., & HORN, L. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of Abnormal Psychology,* **93,** 321–330.

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.

NELSON, D. L., SCHREIBER, T. A., & HOLLEY, P. E. (1992). The retrieval of controlled and automatic aspects of meaning on direct and indirect tests. *Memory & Cognition,* **20,** 671–684.

REINITZ, M. T., & ALEXANDER, A. (1996). Mechanisms of facilitation in primed perceptual identification. *Memory & Cognition,* **24,** 129–135.

RICHARDSON-KLAVEHN, A., LEE, M. G., JOUBRAN, R., & BJORK, R. A. (1994). Intention and awareness in perceptual identification priming. *Memory & Cognition,* **22,** 293–312.

ROEDIGER, H. L., III (1990). Implicit memory: Retention without remembering. *American Psychologist,* **45,** 1043–1056.

ROEDIGER, H. L., & BLAXTON, T. A. (1987). Retrieval modes produce dissociations in memory for surface information. In D. Gorfein & R. R. HOFFMAN (Eds.), *Memory and cognitive processes: The Ebbinghaus centennial conference* (pp. 349–379). Hillsdale, NJ: Erlbaum.

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., WELDON, M. S., & CHALLIS, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing 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.

SCHWARTZ, B. L. (1989). Effects of generation on indirect measures of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15,** 1119–1128.

SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory,* **4,** 592–604.

SRINIVAS, K., & ROEDIGER, H. L., III (1990). Classifying implicit memory tests: Category association and anagram solution. *Journal of Memory and Language,* **29,** 389–412.

STRAIN, E., PATTERSON, K., & SEIDENBERG, M. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **21,** 1140–1154.

THAPAR, A., & GREENE, R. L. (1994). Effects of level of processing on implicit and explicit tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 671–679.

TOTH, J. P., REINGOLD, E. M., & JACOBY, L. L. (1994). Toward a redefinition of implicit memory: Process dissociations following elaborative processing and self-generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **20,** 290–303.

TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review,* **80,** 352–373.

WELDON, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **17,** 526–541.

WELDON, M. S., & ROEDIGER, H. L., III. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition,* **15,** 269–280.

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., & MOSCOVITCH, M. (1989). Stochastic independence between two implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **15,** 22–30.

(Received June 4, 1996)
(Revision received October 17, 1996)