

Direct versus indirect tests of memory: Directed forgetting meets the generation effect

COLIN M. MACLEOD and KAREN A. DANIELS
University of Toronto, Scarborough, Ontario, Canada

Subjects read 20 words and generated 20 others from definitions during a 40-item study phase. Production of each word was followed by an instruction to remember or to forget that word. In free recall, a direct test of memory, words that had been generated were recalled much better than words that had been read. The remember–forget instructional manipulation affected read words but not generated words. In speeded word reading, an indirect test of memory, all studied words showed priming, but read words showed more priming than generated words. Here, the effect of remember versus forget instructions appeared only for generated words. These dissociations of a direct and an indirect test indicate that two powerful encoding manipulations affect separable processes to which these tests are differentially sensitive.

The prevalent lay view of remembering rests on the conscious recollection of past episodes. Yet even over 100 years ago, in his ground-breaking empirical studies, Ebbinghaus (1885/1964) distinguished two ways of remembering: “involuntary recollection” and “voluntary recollection.” This distinction can be seen as a precursor to the current concepts of implicit and explicit memory (Graf & Schacter, 1985), as differentiated by two classes of memory tests. *Indirect* tests reveal the influence of memory through facilitation of performance in the absence of conscious recollection; in contrast, *direct* tests require the conscious recollection of previous information. As illustrations, intentionally trying to recollect references that we wish to cite in an article constitutes a direct use of memory, whereas fluently using the many features of our word processors without realizing that memory is being tapped is an indirect use of memory.

A variety of dissociations in performance have been observed between direct and indirect tests, suggesting that these two ways of using memory are quite different (see Richardson-Klavehn & Bjork, 1988; Roediger & McDermott, 1993, for reviews). Studies of amnesia have made especially important contributions. Despite their frequently catastrophic failure on direct tests of memory, amnesics can show normal performance on indirect tests of memory (Graf & Schacter, 1985; Graf, Shimamura, & Squire, 1985; Warrington & Weiskrantz, 1970). These unconscious effects in indirect tests—usually expressed as a benefit conferred on subsequent processing by previous processing—

have come to be called *priming effects*. As reviewed by Richardson-Klavehn and Bjork, the combined evidence from amnesic and normal subjects reveals a number of trends, of which we will highlight two. First, direct tests generally are more influenced than are indirect tests by semantic elaboration during encoding (Graf & Mandler, 1984; Jacoby & Dallas, 1981), although exceptions do exist (Graf & Schacter, 1985). Second, a shift in modality from study phase to test phase (e.g., from auditory to visual) has little effect on direct measures yet sharply attenuates priming on indirect measures (Bassili, Smith, & MacLeod, 1989; Roediger & Blaxton, 1987), although exceptions exist here as well (Jacoby & Dallas, 1981).

It should be noted, however, that despite such striking dissociations between direct and indirect tests, there also are significant parallel effects. As illustrations, repetition of stimuli appears to facilitate performance on both types of test (Jacoby & Dallas, 1981; but see Challis & Sidhu, 1993), and same-context responses are significantly more effective than are different-context responses in aiding retention on both types of test (Graf & Schacter, 1985).

Roediger (1990) has offered a transfer-appropriate processing account of the pattern of results seen on direct versus indirect tests (see also MacLeod & Masson, 1997, in press; Masson & MacLeod, 1992). Here, indirect tests are seen as relying on the fluent replaying of the encoding of stimuli, typically emphasizing perceptual components of their initial analysis. Direct tests typically rely more on any elaborative encoding done during study, emphasizing semantic/conceptual processing that went beyond simply analyzing the stimulus in context. Thus, a modality change from study to test should disrupt recovery of the initial perceptual analysis, affecting primarily indirect tests, whereas semantic elaboration at encoding should enhance conceptual processing, affecting primarily direct tests, as has been shown repeatedly.¹

Our goal was to explore the direct–indirect test distinction further by manipulating two powerful encoding vari-

This research was supported by Natural Sciences and Engineering Research Council of Canada Grant A7459, and formed the senior honors thesis of K.A.D., who is now at the Georgia Institute of Technology. We thank Monica Ghosal and Shelley Hodder for their assistance in collecting the data, and Barbara Basden, John Gardiner, Pierre Perruchet, and Henry Roediger III, for their helpful comments during the review process. Correspondence should be addressed to C. M. MacLeod, Division of Life Sciences, University of Toronto, Scarborough, ON, M1C 1A4 Canada (e-mail: macleod@scar.utoronto.ca).

ables: directed forgetting and generation versus reading. These variables are particularly of interest because they permit us to investigate the separate effects of initial perceptual/conceptual encoding and of subsequent rehearsal on direct versus indirect tests. In what follows, we will consider the literature relating to each of these issues in turn before describing in detail the purpose of our study.

Directed Forgetting

Over the past 30 years, it has been well established that stimuli accompanied by an instruction to forget ordinarily show reduced retention relative to those accompanied by an instruction to remember (for recent reviews, see Johnson, 1994; MacLeod, 1998). Known as *directed forgetting*, this phenomenon can be seen as a proper subset of the more general phenomenon of intentional forgetting, which spans social psychology, psychology and the law, animal cognition, and a variety of other domains (see the recent book edited by Golding & MacLeod, 1998). Yet by far the majority of the directed forgetting work in the memory literature has exclusively examined performance on direct tests of memory, primarily recall and recognition.

It is worth noting at the outset that directed forgetting can be realized in either of two ways. Under the *item method*, each individual item presented for study is assigned its own instruction, usually displayed just after the item. Under the *list method*, a single instruction is presented after each half of the list of items, with one half of the list to be forgotten and the other half to be remembered. The item method produces a larger remember-forget (R-F) difference than does the list method in both recall and recognition; indeed, the list method generally does not produce a reliable R-F difference in recognition (see, e.g., MacLeod, 1999).

Does directed forgetting affect performance on indirect tests as well? The picture is not clear: The answer seems to depend both on the particular indirect test employed and the method of delivering the instructions to remember versus forget, but only a very narrow set of conditions has so far resulted in observation of the effect. Using word fragment completion and repetition priming in lexical decision as the indirect tests, MacLeod's (1989) initial findings with the item method indicated differential priming favoring words to be remembered (R words) over words to be forgotten (F words), the same pattern as on direct tests. Despite replicating MacLeod's (1989) finding with word fragment completion, however, Paller (1990) found no reliable R-F difference using word stem completion under the item method. Also using the item method, Basden, Basden, and Gargano (1993, Experiment 4) replicated MacLeod's (1989) finding, but convincingly showed that even on a word fragment completion test, the effect was restricted to a narrow range of conditions.

Using the list method, neither E. L. Bjork and R. A. Bjork (1996, Experiment 2) nor Basden et al. (1993) found differential priming for R words and F words on a word fragment completion test. Basden and Basden (1996, Experiment 2) did find an R-F difference on an indirect test

that required subjects to provide words in response to definitions—presumably a more conceptual indirect test—although only under the item method. Thus, the most general conclusion emerging now from the directed forgetting literature is that there ordinarily is not an R-F difference on indirect tests of memory, although as yet there are only a few relevant studies. It also appears to be the case that indirect tests can be very sensitive to what otherwise might seem to be minor changes over experiments.

Two other points regarding directed forgetting studies are especially relevant. First, as Basden et al. (1993) have argued (see also MacLeod, 1998, 1999), the operative mechanism in directed forgetting appears to differ as a function of the method of delivery of the instructions. In the list method, F words appear to be encoded as well as R words, but inhibited from retrieval. In contrast, in the item method, F words appear to be rehearsed less and hence are not as well encoded. In our study, we used the item procedure because we wished to focus on directed forgetting due to selective rehearsal.

Second, in a study by Golding, Long, and MacLeod (1994) that examined recall only, we found that the ability to forget information on cue was a function of the relatedness of studied items. Subjects studied a list containing some semantically related successive words (e.g., *crab-leg*) and some unrelated successive words (e.g., *cheese-fire*, constructed from the related pairs *cheese-cake* and *camp-fire*). Using the item method, successive words sometimes received the same instruction, sometimes different instructions. Although recall was best when both words in a pair were given a remember cue (i.e., the R-R condition), we also found good recall of semantically related word pairs in the R-F condition, much better than when the first word of the pair was an F word. We interpreted this finding as suggesting that compliance with an instruction to forget a word is difficult when that word immediately follows an R word to which it is semantically related. The elaborative, conceptual processing already completed for the R word by the time the F word appears undermines the ability to differentially encode the F word and the R word. Perhaps, then, on a direct test, any study task that encouraged elaborative encoding could defeat an instruction to forget by making postencoding rehearsal differences inoperative.

Generation Versus Reading

On a direct recognition test, Jacoby and Dallas (1981) found that words generated from anagrams during study were better recognized than words that were simply read. In contrast, on an indirect masked word identification test where subjects were to identify words appearing on the screen for 30 msec prior to a mask, the opposite pattern appeared: There was more priming for words that were read than for words that were generated. Also using a masked word identification test, Weldon (1991, Experiment 1) supported this finding, aligning it with the well-established pattern in word fragment completion (see Roediger & McDermott, 1993, for a review).

Since then, Masson and MacLeod (1992) and MacLeod and Masson (1997) have confirmed the reliable priming for generated words on the masked word identification test. More recently, MacLeod and Masson (in press) have extended this to another indirect measure, speeded word reading. Here, subjects simply read aloud a fully exposed test word as quickly as they can into a microphone. Sometimes called "naming" (see, e.g., Balota & Chumbley, 1984), speeded word reading is an especially good candidate as an indirect measure of memory because the task is performed very rapidly—indeed virtually automatically. This argument is buttressed by the well-known Stroop (1935) effect, where color words cannot be ignored when the task is to name the ink colors in which they are printed. Word reading should not suffer the intrusion of conscious recollection—the bane of indirect tests—because it is extremely fluent and easy and does not require any of the problem-solving types of activities required of indirect tests such as word fragment completion, where the test stimulus is quite degraded.

Rationale for the Present Experiment

This experiment brings together all of the factors that we have just set out, and addresses two primary questions. The first relates to direct tests of memory. As already described, the generation effect and directed forgetting strongly affect recall and recognition. Yet there is evidence that, for words, directed forgetting cannot overcome an already strong elaborative encoding, at least in the case of preexisting semantic associations (Golding et al., 1994; but see Basden & Basden, 1996, Experiment 1, for a different result with pictures). Would this also be true for new episodic elaborative encodings, as in the case of generated words (but not read words)? We expected recall performance on R words to be better than that on F words when those words were read at study, consistent with the considerable evidence for selective rehearsal of the R words (see MacLeod, 1998, for a review). However, if our logic is correct, words generated at study should not show an R–F difference on a recall test because the elaborative encoding already carried out on these words should be able to withstand an instruction to forget; that is, subsequent differential rehearsal should not be able to override such strong conceptual encoding. Of course, recall of generated words should be much better than that of read words (Slamecka & Graf, 1978).

Turning to indirect test performance, we chose to use the word reading task because of the virtues of this procedure as an indirect measure that we described earlier. On the basis of prior findings regarding directed forgetting (Basden et al., 1993; MacLeod, 1989; Paller, 1990) and regarding the generation effect (MacLeod & Masson, 1997, in press; Masson & MacLeod, 1992), we predicted that all studied words should show priming. We also anticipated more priming for read words than for generated words, given that the perceptual processing of read words during initial encoding is probably optimal for the indirect speeded reading test. However, when the weight of perceptual encoding is reduced, as for the generated words, it

might be possible to observe effects of differential rehearsal, so we anticipated that there could be an R–F difference in priming on the indirect speeded reading test for the generated words only.

METHOD

Subjects

The subjects were 89 undergraduates from introductory psychology at the University of Toronto at Scarborough. They were recruited on a voluntary basis and received bonus points for their participation. Fifty-four subjects took part in the speeded reading task. The data from 4 of these subjects were discarded due to microphone malfunction, leaving 50 in this condition.² The remaining 35 subjects took part in the free-recall task only.

Materials

The 60 stimuli were the definitions and target words used by Weldon (1991) and presented in the appendix to MacLeod and Masson (1997). An example would be "a large, triangular structure in Egypt–p?" and the corresponding target "pyramid." The program randomly assigned the words to the generate, read, or new condition for each subject. An additional 12 words were chosen to serve as buffers in the study phase and the reading task. The buffer words matched those in Weldon's list according to frequency of usage and word length, and were selected from the Paivio, Yuille, and Madigan (1968) norms.

Apparatus

Subjects were tested using IBM-compatible 286 computers equipped with Tatung CM-1496 14-in. VGA color monitors. When the subject spoke into a Realistic Highball-7 microphone, the signal was amplified by a Realistic SA-150 stereo amplifier and input to the computer via a modified keyboard that permitted millisecond accuracy response timing using the routines taken from Graves and Bradley (1987, 1988). All programs were written in QuickBASIC 4.5. The screen background color was white (palate no. 15); the stimuli were presented in black (palate no. 0).

Study Procedure

Each subject studied 40 words, with the remaining 20 words in the set reserved to provide new, unstudied baseline items in the reading test. Half of the studied words were presented in isolation (e.g., *cactus*); the subject read the word aloud and then the instruction for that word appeared. For the other half of the studied words, a definition was presented (e.g., "this protects you from the rain–u?") and remained on the screen until the subject responded aloud, at which point its instruction appeared. Each instruction remained on the screen for a period of 3 sec: "RRRRR" indicated that the preceding word was to be remembered, and "FFFFF" indicated that it could be forgotten. There were 10 remember and 10 forget instructions for each of the 20 read and 20 generate words, presented in random order to the subject.

Because only a subject's first response was accepted during study, accuracy was emphasized. The experimenter pressed one of two keys to score accuracy. If the subject did not produce the intended answer, that word was eliminated from further analysis on the retention tests. (Note that leaving these incorrectly generated words in the data set for the analyses did not alter the pattern of results, just as Masson and MacLeod, 1992, showed.) Stimuli were centered top to bottom and left justified on the computer screen, as were the instructions to remember or forget.

Test Procedure

On the speeded word reading task, the 50 subjects were required to read aloud each test word that appeared on the screen as quickly as possible. Again, stimuli were centered top to bottom and left jus-

tified and stayed on the screen until the subject responded. In addition to the 40 studied words, 20 unstudied words were incorporated to provide a baseline measure, with the order of all 60 words randomized. Subjects spoke each test word into a microphone, allowing response latency to be recorded by the computer. All subjects were given a practice session using the microphone prior to the experiment. The experimenter again recorded accuracy by pressing one of two keys.

The 35 subjects who did the free recall test were instructed to take as much time as they needed to write down, in any order, as many of the words as they could remember. They were explicitly told to disregard the study phase instructions to remember or forget.

RESULTS

Study Phase

During the study phase, the mean proportions of errors were .000 for the words that were read and .055 for the words that were generated.

All principal analyses were 2×2 analyses of variance (ANOVAs), with the two within-subjects variables being encoding condition (read vs. generate) and instruction (remember vs. forget). Additional t tests were conducted on the speeded word reading data for comparisons of studied to new (unstudied) items; this was accomplished by averaging over the four studied conditions and comparing this mean to the single new condition.

Speeded Word Reading Test

Both accuracy and latency data were collected for the speeded reading task. The subjects were highly accurate when reading the words from the computer screen. The mean accuracy rates for the five conditions were as follows: .998 for read-remember, .998 for read-forget, .998 for generate-remember, 1.000 for generate-forget, and .993 for new.³ Accuracy was greater for studied items than for new items [$t(49) = 2.01, p = .05$], demonstrating priming in the accuracy data. Not surprisingly, though, a 2×2 ANOVA confirmed that there were no reliable differences in accuracy among the four studied conditions (all three $F_s < 1$).

The latency data were of primary interest. Table 1 displays the mean response latencies and the corresponding standard errors for each of the four studied conditions, as

Table 1
Mean Response Time (RT, in Milliseconds) and Standard Error (SE) on the Indirect Speeded Word Reading Test and Mean Proportion Correct (PC) and Standard Error (SE) on the Direct Free Recall Test, as a Function of Encoding Condition

Condition	Test			
	Speeded Reading		Recall	
	RT	SE	PC	SE
Read				
Remember	547	8.28	.269	.036
Forget	548	9.51	.146	.021
Generate				
Remember	553	8.88	.424	.030
Forget	563	9.62	.416	.031
New	565	10.00		

well as for the new condition. Note that trimming the individual subject latencies to censor those shorter than 300 or longer than 1,200 msec removed almost no data (only five latencies in total). Response latency was faster for studied items than for new items [$t(49) = 4.36, p < .001$], demonstrating priming in the latency data. The principal analysis, a 2×2 ANOVA, revealed reliable main effects of encoding condition [$F(1,49) = 16.72, MS_e = 348.76, p < .001$] and of instruction [$F(1,49) = 4.16, MS_e = 413.27, p < .05$]. Priming was greater for read items (548 msec) than for generated items (558 msec), consistent with typical findings on indirect measures (e.g., Jacoby & Dallas, 1981). Also, R words (550 msec) showed more priming than F words (556 msec), again consistent with prior research (MacLeod, 1989). The interaction of these two variables, although not reliable, did point to a clear trend, evident in Table 1 [$F(1,49) = 2.66, MS_e = 439.73, p = .11$].

Our primary concern was with the R-F difference as a function of encoding condition. Planned comparisons revealed that this difference was not significant for read words ($F < 1$), but that it was significant for generated words [$F(1,49) = 4.94, MS_e = 1,159.44, p < .05$]. Put simply, the R-F difference in speeded word reading was restricted to the words generated during study.

Free Recall Test

For free recall, the mean proportions of words accurately recalled following the study phase are displayed in Table 1. Subjects recalled an average proportion of .314 of the 40 study words during the recall task. The 2×2 ANOVA showed reliable main effects of both encoding condition [$F(1,34) = 73.44, MS_e = .022, p < .001$] and instruction [$F(1,34) = 5.16, MS_e = .029, p < .05$]. That more generated words (.420) were recalled than read words (.208) concurs with prior findings for direct tests of memory (MacLeod & Masson, 1997, in press; Masson & MacLeod, 1992; Slamecka & Graf, 1978; Weldon, 1991). The overall better recall of R words (.346) than of F words (.281) is also in keeping with the directed forgetting literature (see MacLeod, 1998). The interaction was also significant [$F(1,34) = 4.82, MS_e = .024, p < .05$].

With regard to the planned comparisons, whereas recall of the generated words showed no significant effect of instruction ($F < 1$), recall of the read words was significantly better following remember instructions than following forget instructions [$F(1,34) = 9.97, MS_e = .053, p < .01$]. Put simply, the R-F difference in recall was restricted to the words that had been read during study.

DISCUSSION

We have observed three principal dissociations within a single experiment. First, we replicated the familiar read-generate dissociation across tests: Our indirect speeded reading test showed better memory (more priming) for read words, whereas our direct recall test showed better memory (greater recall) for generated words. This, of course, is the pattern reported by Jacoby (1983; Jacoby &

Dallas, 1981) and fits nicely with transfer appropriate processing accounts (e.g., Masson & MacLeod, 1992; Roediger, 1990). Under such views, encoding at a more perceptual level is most beneficial to performance on the indirect test, which is typically more perceptually driven, whereas encoding at a more conceptual level is most beneficial to performance on the direct test, which is typically more conceptually driven.

Second, there was a reliable effect of instruction for the read words but not for the generated words on the direct recall task. The presence of the effect for read words is entirely consistent with the extensive directed forgetting literature (see Johnson, 1994; MacLeod, 1998, for reviews of the many relevant studies). Recall that under the item method, a remember instruction appears to motivate the subject to rehearse that item; by contrast, a forget instruction functions to discourage rehearsal of that item (see Basden et al., 1993; R. A. Bjork, 1989). Given the largely perceptual encoding for the words read at study, R words are better recalled because they have benefited from more rehearsal following initial encoding. Now consider generated words. Here, our claim is that the semantic encoding of both the R words and the F words was already so rich that differential rehearsal was virtually irrelevant, and was essentially overridden. Consistent with the Golding et al. (1994) finding that it is difficult to forget an F word that is semantically related to a prior R word, it appears that if every word is processed elaboratively prior to presentation of its instruction, then the normal potency of the forget instruction (which hinges on differential rehearsal) is completely eliminated on a direct test such as free recall.

The third dissociation occurred within the indirect speeded word reading task. Here, there was also a selective effect of the directed forgetting manipulation, but now only the generated words and not the read words showed a reliable R-F difference. A critic might wish to suggest that this was the result of contamination from explicit retrieval, given the better memory for generated words on direct tests. However, two arguments militate against such a criticism. The first is empirical: There was no corresponding R-F difference in recall for generated items to support such contamination. The second is more methodological (see MacLeod, 1996; MacLeod & Masson, in press, for relevant arguments): As an indirect measure of memory, speeded word reading has the virtue that the test item is presented without degradation and hence without the need for problem solving, thereby obviating any need for conscious recollection.

So how can we explain this dissociation within the indirect test? We rely again on the idea of differential rehearsal. The logic is really the same as that used to explain the dissociation within the direct test: When initial encoding optimizes performance on any test, directed forgetting cannot exert its effect. In recall, performance was optimized by elaborative encoding of the generated words, making differential rehearsal redundant; thus, an R-F difference could appear only for read words. In contrast,

priming on speeded reading was optimal when words were read during encoding, so the instructional manipulation could not have any effect in this condition. For generated words, however, where the benefit of additional rehearsal for R words could be seen because priming was not optimal, the R-F difference did indeed appear.

In summary, these three dissociations help to differentiate the processing that underlies performance on an indirect test from that underlying performance on a direct test. Under the transfer-appropriate processing framework, it is the match between processing at study and processing at test that is crucial in determining the success of remembering, and that explains the oft-observed between-test dissociation of the read and generate conditions. But what of the two within-test dissociations due to directed forgetting that we have reported here? Our argument is that the benefits of the additional rehearsal that accrues to R items are visible only when performance on a test is not optimized by the type of encoding. This occurs in direct tests where encoding was not elaborative, and in indirect tests where encoding was not perceptual.

REFERENCES

- 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 & Performance*, *10*, 340-357.
- BASDEN, B. H., & BASDEN, D. R. (1996). Directed forgetting: A further comparison of the list and item methods. *Memory*, *4*, 633-653.
- BASDEN, B. H., BASDEN, D. R., & GARGANO, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 603-616.
- 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.
- BJORK, E. L., & BJORK, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness & Cognition*, *5*, 176-196.
- BJORK, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.
- CHALLIS, B. H., & SIDHU, R. (1993). Dissociative effect of massed repetition on implicit and explicit measures of memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 115-127.
- EBBINGHAUS, H. (1885/1964). *Memory: A contribution to experimental psychology*. New York: Dover. (Original work published 1885)
- GOLDING, J. M., LONG, D. L., & MACLEOD, C. M. (1994). You can't always forget what you want: Directed forgetting of related words. *Journal of Memory & Language*, *33*, 493-510.
- GOLDING, J. M., & MACLEOD, C. M. (Eds.) (1998). *Intentional forgetting: Interdisciplinary approaches*. Mahwah, NJ: Erlbaum.
- GRAF, P., & MANDLER, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning & Verbal Behavior*, *23*, 553-568.
- GRAF, P., & SCHAFFER, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 501-518.
- GRAF, P., SHIMAMURA, A. P., & SQUIRE, L. R. (1985). Priming across modalities and priming across category levels: Extending the domain of preserved function in amnesia. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *11*, 385-395.
- GRAVES, R., & BRADLEY, R. (1987). Millisecond interval timer and au-

- ditory reaction time programs for the IBM PC. *Behavior Research Methods, Instruments, & Computers*, **19**, 30-35.
- GRAVES, R., & BRADLEY, R. (1988). More on millisecond timing and tachistoscopic applications for the IBM PC. *Behavior Research Methods, Instruments, & Computers*, **20**, 408-412.
- JACOBY, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.
- JOHNSON, H. M. (1994). Processes of successful intentional forgetting. *Psychological Bulletin*, **116**, 274-292.
- MACLEOD, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 13-21.
- 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. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Erlbaum.
- MACLEOD, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*, **6**, 123-129.
- MACLEOD, C. M., & MASSON, M. E. J. (1997). Priming patterns are different in masked word identification and word fragment completion. *Journal of Memory & Language*, **36**, 461-483.
- MACLEOD, C. M., & MASSON, M. E. J. (in press). Repetition priming in speeded word reading: Contributions of perceptual and conceptual processing episodes. *Journal of Memory & Language*.
- 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.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs*, **76** (1, Pt. 2), 1-25.
- PALLER, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differently by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1021-1032.
- RICHARDSON-KLAVEHN, A., & BJORK, R. A. (1988). Measures of memory. *Annual Review of Psychology*, **39**, 673-690.
- 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.
- SLAMECKA, N. J., & GRAF, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 592-604.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- WARRINGTON, E. K., & WEISKRANTZ, L. (1970). The amnesiac syndrome: Consolidation or retrieval? *Nature*, **228**, 628-630.
- WELDON, M. S. (1991). Mechanisms underlying priming on perceptual tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 526-541.

NOTES

1. It is important to realize, though, that the mapping of direct test to conceptual processing and indirect test to perceptual processing is clearly an oversimplification. Although tests tend to sort in this way, there certainly exist conceptual indirect tests and perceptual direct tests (see Roediger & McDermott, 1993). Because the speeded word reading test would seem to be more perceptual and the recall test certainly is more conceptual, we have chosen for simplicity to emphasize the more common mapping here. Consequently, our claims regarding indirect tests can be interpreted as referring only to perceptual indirect tests.

2. Actually, two sets of data—each made up of 25 subjects—were collected about a year apart in the speeded word reading task. A preliminary analysis including these two subgroups as a factor revealed no reliable main effect of subgroup and no reliable interactions involving subgroup, so this variable was deleted and the data for all 50 subjects were merged. Note, too, that a recognition test was administered to all subjects after the speeded word reading test. Because performance in the generate condition was at ceiling on the recognition test, it will not be discussed further.

3. Because the absence of variance in the generate-forget accuracy data would have prevented analysis of variance, 1 subject's score in that condition was changed from 1.000 to 0.999.

(Manuscript received November 24, 1997;
revision accepted for publication June 30, 1999.)