The Fate of Repetition Effects When Recognition Approaches Chance

Morris Moscovitch and Shlomo Bentin

We examined whether repetition priming effects remain above baseline when explicit recognition is reduced to chance or near chance levels by forgetting. Subjects studied a set of words, and memory was tested explicitly by yes/no (Experiments 1 & 3) or forced-choice recognition (Experiment 3) after a 20-min delay filled with an interfering task. Memory was then tested implicitly by perceptual identification (Experiment 3) or lexical decision (Experiments 1 & 2) for words seen only at study, at recognition, or both. In all experiments, recognition d' was about 0.75, and repetition effects remained above baseline and constant across conditions. At delays of 24 hr (Experiment 4) yes/no recognition fell to near chance (d' < 0.6) levels in a third of the subjects. Repetition effects as measured by lexical decision were not significantly above baseline for words seen only once in those subjects and 4 other subjects from Experiment 1 who had comparable recognition rates. Reducing recognition by forgetting to near chance levels in normal people is very difficult but when it does occur, repetition effects in normal people, in contrast with those in amnesics, are also significantly reduced or eliminated.

Explicit tests of memory, such as recognition and recall, require conscious recollection of the past, whereas implicit or indirect tests of memory do not (Graf & Schacter, 1985; Moscovitch, 1984). Instead, on implicit tests, memory for a particular item or event is inferred from changes in performance on tests involving that item, with the subject often being unaware that memory is being probed. On some implicit tests, memory is reflected in the increased speed or accuracy with which a repeated item is processed compared with when it was first presented or compared with a newly presented item. We label this effect stimulus-specific repetition priming effect, or simply repetition effect.

Evidence that performance on explicit and implicit tests of memory can be dissociated one from the other has accumulated during the last decade. The most dramatic and, perhaps, the most compelling evidence comes from amnesic patients who often score at chance or baseline levels on explicit tests but perform normally on implicit tests of memory (for reviews, see Moscovitch, 1982, 1984; Schacter, 1987; Shimamura, 1986). A comparable, but less dramatic, pattern has been observed in healthy subjects who are tested at long delays from study. The drop in their performance on explicit tests is steep in comparison with the decline on some implicit tests such as repetition effects in lexical decision tasks (Scarborough, Cortese, & Scarborough, 1977), perceptual identification (Jacoby & Dallas, 1981), speeded reading of geometrically transformed script (Cohen & Squire, 1980; Kolers, 1976; Moscovitch, Winocur, & McLachlan, 1986), word-fragment completion (Sloman, Hayman, Ohta, Law, & Tulving, 1988), word-stem completion with unique solutions (Graf, Squire, & Mandler, 1984), and picture naming (Mitchell & Brown, 1988).

The similarity in the pattern of performance on explicit and implicit tests between normal people tested at long delays and amnesic patients has led to the tacit assumption that evidence from both can be used interchangeably to support or refute theoretical positions regarding the organization of memory. It is possible, however, that organic amnesia and extensive, but normal, forgetting have different effects on performance. For example, although amnesic patients with virtually no explicit memory can still perform relatively well on some implicit tests, it is not well established whether healthy subjects matched with severe amnesia for performance on explicit tests of memory, such as recognition, will perform up to the level of amnesics on implicit tests.

By extrapolating from reports of a more rapid decline in performance on explicit than on implicit tests of memory, it is possible to conclude that normal people, like amnesics, will continue to perform well on implicit tests even if their performance on explicit tests has dropped to chance or baseline levels. This conclusion, however, is not warranted for a number of reasons. First, with the exception of studies examining memory for unattended or poorly perceived material, which we discuss shortly, there are, to our knowledge, no studies comparing performance on explicit and implicit tests in normal people when performance on explicit tests has been reduced to chance by forgetting. It is conceivable, despite the rapid initial decay, that performance on recognition remains above chance even after extremely long delays. Second, if performance on explicit tests can be reduced
to chance, then it is still possible that no savings would be apparent on implicit tests.

An example from a widely cited study by Kolers (1976) on the longevity of repetition effects speaks to the first point. In that study, Kolers found that subjects reread pages of geometrically transformed script faster the second time than the first, even though 13 to 15 months elapsed between presentations. To quote Kolers (1976), "Although small [about 5%], the [repetition] effect is reliable and its significance lies with its occurrence, not its size" (p. 559). What has been often overlooked is that performance on an explicit recognition test for the same material also remained significantly above chance: 65% of passages were classified correctly as old or new. Kolers's statement, therefore, can be applied with equal force to performance on the explicit recognition tests in normal people. In marked contrast, the performance of amnesic patients on a comparable explicit test drops to chance by 2 hr, whereas that on the comparable implicit test is normal (Moscovitch et al., 1986).

Two other findings are worth noting in Kolers's study. When compared with the scores achieved on immediate testing (Kolers, 1973), performance on both implicit and explicit tests declined noticeably, although that on the explicit test was steeper. Second, overall performances on the two tests correlated poorly with each other. It is important to emphasize, however, that not all repetition effects are as long-lasting as those reported by Kolers. Some, such as stem completion with nonunique solutions, disappear within minutes or hours (Diamond & Rozin, 1984; Squire, Shimamura, & Graf, 1987), which is much quicker than performance on explicit tests of memory. However, if one considers only long-lasting repetition effects, the data reported by Kolers are similar to those reported by others in the literature (e.g. Komatsu & Ohta, 1984; Mitchell & Brown, 1988; Sloman et al., 1988). In all instances, repetition effects decline slowly in comparison with recognition, but performance on tests of recognition remain substantially above chance when tested at the same intervals.

In contrast, recognition can be reduced to chance by perceptually degrading the stimulus at study (Bonnano & Stilling, 1986; Kunst-Wilson & Zajonc, 1980; Mandler, Nakamura, & Van Zandt, 1987; Seaman, Brody, & Kauff, 1983; Seaman, Marsh, & Brody, 1984) or by diverting attention from it (Bentin, Kutus, & Hillyard, 1990; Eich, 1984; Merikle & Reingold, 1991). Despite chance recognition, performance on implicit tests remained significantly above chance and, in many cases (Merikle & Reingold, 1991; Seaman et al., 1983), remained stable in the face of wide variations in recognition. Comparison of the effects of encoding manipulations and forgetting on implicit and explicit tests of memory only serves to bring our question into relief: Would performance on implicit tests in normal people remain stable and above baseline as recognition is reduced to near chance levels by forgetting? If not, what accounts for the differences between the two types of studies?

The purpose of the experiments reported in this article, therefore, was to try to reduce, by forgetting, performance on explicit tests of memory to chance or near chance levels in normal people to determine whether they would still show robust savings on implicit tests of memory. If they would, it would strengthen the claim that memory as inferred from performance on some implicit tests is more resilient to forgetting than memory when it is tested explicitly. On the other hand, this claim would be placed in doubt by a failure to find preserved memory on implicit tests in normal people when recognition induced by forgetting is near chance. Such an outcome would favor the view that performance on both implicit and explicit tests depends, in part, on gaining access to a common trace or engram. To anticipate the results, we note at the outset that we had great difficulty in reducing recognition to near chance by forgetting. This itself is an interesting finding when compared with the ease with which encoding manipulations can achieve these results. In those few subjects in whom we succeeded in reducing recognition substantially, performance on implicit tests also dropped significantly. The implications of these results for theories of normal and pathological memory and forgetting are discussed when we present our findings.

The design and procedure was similar for all of the experiments we conducted. Each experiment consisted of four successive phases: a study phase, an interpolated interfering task, an explicit recognition test, and an implicit memory test in which a repetition effect was sought.

In the study phase, subjects made lexical decisions to 240 strings of letters that were evenly divided into words and nonwords. Depending on the experiment, the recognition procedure was either yes/no recognition or modified forced-choice. Memory was next tested implicitly in a lexical decision or perceptual identification task. The words in the implicit test were encountered in the study phase, only in the recognition phase, in both the study and recognition phase, or were new words that had never been presented in the experiment up to that point. Repetition effects, a measure of savings on the implicit test, were calculated by computing the difference in latency or identification accuracy between old, repeated items and new items encountered for the first time.

Although items were not counterbalanced within experiments, different words from the same pool were chosen at random across Experiments 1, 3, and 4 to form the different conditions. For example, the set comprising the repeated words in one experiment was different from the set in the other experiments. The same is true of the other conditions. The only exception is that, in Experiments 1 and 2, the same items were used in corresponding conditions in both experiments. The words in the different conditions were equated for frequency, imageability, and length.

The first three experiments were devoted to assessing the relative sensitivity of different implicit tests of memory, yes/no recognition and modified forced choice, and two different implicit tests, improvement in lexical decisions and in identification accuracy with repetition. On the basis of this information, a fourth experiment was conducted in an attempt to lower performance on the explicit test even further and determine what effect that manipulation had on performance on the implicit test.
Experiment 1

In previous studies, we had established that repetition effects in a lexical decision task decline sharply, from a lag of 0 to a lag of 4 between repeated words, and then are virtually unchanged to a lag of 15 (Bentin & Moscovitch, 1988) and even longer (Bentin & Moscovitch, 1990). Recognition, on the other hand, decayed steadily over the same intervals but still remained well above chance, yielding a d' of about 1.8. Similar results were reported by Ratcliff, Hockley, and McKoon (1985; but see Dannenbring & Briaud, 1982). Interpolating an interfering, mathematical task between the presentation of each word and extending the lag to 29 items had no noticeable influence on the repetition effect (Moscovitch, 1985). Recognition, on the other hand, did deteriorate but remained well above chance at about 80%. Many amnesic patients, however, were scoring at chance on recognition, even without the interpolated task, but showed repetition effects that were no different from normal (Moscovitch, 1985). The purpose of Experiment 1, therefore, was to determine whether repetition effects in normal people could survive even further reductions in recognition. In addition, we wished to compare the repetition effect for words that were explicitly recognized with those that were not, something the design of the previous experiment did not allow us to do.

Although stochastic independence between performance on implicit and explicit tests of memory had been reported in previous studies (Jacoby & Witherspoon, 1982; Tulving, Schacter, & Stark, 1982; for review, see Richardson-Klavehn & Bjork, 1988; Schacter, 1987, and Tulving, 1985), we were interested in the effect that low overall recognition scores had on the phenomenon.

Method

Subjects. Twenty-four undergraduate students, all native speakers of Hebrew, participated in Experiment 1 either for course credit or payment. All had normal or corrected-to-normal vision.

Stimuli. Stimuli consisted of 232 Hebrew words (all nouns) and 160 pronounceable nonwords. In the absence of a reliable word-frequency count in Hebrew, word frequency was determined on the basis of subjects' ratings. In several previous studies, subjects rated the frequency of words on a 5-point scale (1 = low frequency, 5 = high frequency). The words included in Experiment 1 were sampled from that pool. Their average rated frequency ranged between 2 and 3 points (M = 2.78). Both words and nonwords were between 3 and 5 letters long, which, because the regular unwoveled print was used, were equivalent to 4- to 7-letter words in English. The words used in each phase of the experiment were randomly selected, and there were no frequency or word-length differences among these groups. All stimuli were computer-generated on a cathode-ray tube (CRT), and a standard set of Hebrew characters without the vowel dots was used. They were presented in double-size and double-width format between two lines of asterisks centrally located and subtending a visual angle of 7° × 4°.

Procedure. In the study phase, subjects made lexical decisions to 240 strings of letters evenly divided between words and nonwords and presented in a random order. Stimuli were exposed for 1000 ms each and separated by 2,500 ms interstimulus intervals. Responses were made throughout the study by pressing a button in a response box. The right hand was used for “word” responses and the left for “nonword” responses.

An interference task was interpolated between study and recognition. The task lasted about 20 min, during which subjects distinguished between words and numbers. The words in the interference task were 80 regular nouns sampled from the same original pool as the experimental words.

In the explicit memory test that followed, 40 of the words encountered in the study phase were presented along with 80 lures. Each item appeared alone on the CRT screen and the subject had to judge whether it was an old item (from the study phase) or a new one. The right hand was used to respond “old” and the left hand, “new.”

The implicit test was administered immediately afterwards. Subjects viewed 240 letter strings evenly divided between words and nonwords. Forty of the words were previously encountered only in the study phase, 24 words were encountered only in the explicit phase (i.e., they were randomly sampled from the new words in the recognition test), 24 were presented both in Phase 1 and in Phase 2 (i.e., they were old items in the recognition task), and 32 were new words. Subjects decided as quickly and as accurately as possible whether each item was a word or a nonword. Subjects were told that some words might appear that they had seen before but that this fact was not relevant to the present task. They were to treat those words as they would any others.

Results and Discussion

Explicit test (yes/no recognition). As Table 1 shows, the proportion of words correctly judged as old or new was .60 and .71, respectively (d' = 0.86). Hits minus false alarms (FAs), another measure of recognition accuracy (corrected for guessing), was 0.31. Reaction times (RTs) were averaged separately for each of the four possible outcomes: hit, correct rejection (CR), FA, and miss (see Table 1). A two-factor analysis of variance (ANOVA) showed that the correct responses (hit and CR) were faster than incorrect responses (miss and FA), F(1, 23) = 14.30, MS_e = 7.085, p < .001, and that the responses to new items (CR and FA) were as fast as to old items (hit and miss; F < 1). The interaction between these two effects was not reliable, F(1, 23) = 2.63, MS_e = 8.150, p < .12.

Implicit test (lexical decision). The results are presented in Table 2. At study, RTs to words (592 ms) were longer than to nonwords (730 ms). RTs in the implicit test were broken down according to whether the words happened only at study, only at recognition, on both occa-

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unrepeated words</th>
<th>Repeated words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
<td>FA</td>
</tr>
<tr>
<td>RT, %</td>
<td>832</td>
<td>867</td>
</tr>
<tr>
<td>Note. Mean RTs are expressed in milliseconds; d' = 0.86. CR = correct rejection; FA = false alarm.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2
Mean Reaction Times (RTs) in the Implicit Test
(Experiment 1)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study list</th>
<th>Unrepeated words</th>
<th>In study</th>
<th>CR</th>
<th>FA</th>
<th>Seen twice before</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>592</td>
<td>615</td>
<td>568</td>
<td>577</td>
<td>540</td>
<td>573</td>
</tr>
<tr>
<td>SEM</td>
<td>24</td>
<td>29</td>
<td>22</td>
<td>27</td>
<td>23</td>
<td>33</td>
</tr>
</tbody>
</table>

Note. Mean RTs are expressed in milliseconds. CR = correct rejection; FA = false alarm.

In Experiment 1, memory was tested explicitly by a yes/no recognition procedure in which the ratio of new to old items was 2:1. The purpose of Experiment 2 was to determine whether a modified forced-choice recognition procedure would prove to be a more sensitive, explicit test of memory, perhaps a more worthy rival to the implicit test. The modification consisted of presenting the subjects with a printed list of 120 items of which they had to choose 40 that they believed they had seen before.

Method

Subjects. Twenty-four undergraduates, all native speakers of Hebrew, participated in the experiment either for course credit or for payment. All had normal or corrected-to-normal vision. None of the subjects had taken part in Experiment 1.

Procedure. Except for the explicit recognition test, the stimuli and procedure were identical to those in Experiment 1. As in Experiment 1, assignment of words to each condition was random. Modified forced-choice recognition was tested by requesting each subject to circle 40 words seen in the initial lexical decision task from a printed list that included an additional 80 words as lures.

Results and Discussion

Performance on the modified forced-choice recognition tests yielded a $d'$ of 0.84, a similar value to that of 0.86 in Experiment 1. By using the proportion of hits minus FA as a corrected measure of recognition accuracy, we obtained a value of .29, which was also similar to the value of .31 found in Experiment 1.

As Table 3 shows, performance on the implicit lexical decision task was also similar to that of Experiment 1. A one-way ANOVA with repeated measures showed a significant effect of stimuli category, $F(5, 115) = 6.73$, $MS_e = 1.160$, $p < .0001$. Post hoc Tukey comparisons showed that new words yielded longer RTs than did any of the categories of repeated (old) words. RTs to old words did not differ significantly among themselves. A planned $t$ test revealed that the RTs to new words in the implicit test were not reliably different from the RTs to words in the study list.

Table 3
Mean Reaction Times (RTs) in the Implicit Test
(Experiment 2)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Study list</th>
<th>Unrepeated words</th>
<th>In study</th>
<th>CR</th>
<th>FA</th>
<th>Seen twice before</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>591</td>
<td>611</td>
<td>569</td>
<td>579</td>
<td>577</td>
<td>565</td>
</tr>
<tr>
<td>SEM</td>
<td>15</td>
<td>20</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>17</td>
</tr>
</tbody>
</table>

Note. Mean RTs are expressed in milliseconds. CR = correct rejection; FA = false alarm.
The results of Experiment 2 indicate that, insofar as the materials and procedures that we used are concerned, modified forced choice is no more sensitive than yes/no recognition as an explicit test of memory. Because we did not use forced choice in an easier recognition condition, as we did for old/new recognition (Bentin & Moscovitch, 1990), we cannot tell whether performance on the two types of recognition tests decayed at the same rate. We can conclude, however, that under the comparable conditions of Experiments 1 and 2, recognition as measured by the two types of tests is roughly equivalent.

It is also worth noting that a similar repetition effect of 40–50 ms was found in both Experiments 1 and 2 and was the same whether the words were seen only at study, only at recognition, or at both instances.

**Experiment 3**

Just as it is possible for one recognition test to be more sensitive than another, so it is possible for one type of implicit test to be more sensitive than another. In Experiments 1 and 2 we found that repetition effects in lexical decision, an implicit measure of memory, did not decay noticeably over the intervals and interpolated task that we used. Would the same be true of repetition effects in perceptual identification, which is another widely used measure of performance on implicit memory tests? Although there is already substantial evidence in the literature that repetition effects in perceptual identification are long-lasting (Jacoby, 1983; Jacoby & Dallas, 1981), there is no study, to our knowledge, that compared decay of repetition effects in one type of test with decay of repetition effects in another. In this case, the comparison cannot be direct because latency measures repetition effects in lexical decision, whereas it is accuracy that measures perceptual identification. Even if the measures were ostensibly the same in the two cases, it is unlikely that a direct comparison would be legitimate because both the baselines and the metric would have to equated. In addition, the underlying process that was being reflected in the two measures would have to be equivalent. To avoid these problems, the comparison we propose is one between patterns of performance on the two types of implicit memory tests. As we noted earlier, the repetition effect was equivalent for all repeated items regardless of whether they had been seen only once (either at the start of the experiment or at recognition) or twice (at both study and at the recognition test immediately preceding the implicit test). In short, there was no indication of a significant decay with time and interpolated activity over the intervals we sampled. Would the same be true of repetition effects in the perceptual identification test?

**Method**

**Subjects.** The participants in the study were 24 undergraduates who met the same inclusion criteria as the subjects in Experiments 1 and 2. None of them had participated in the previous experiments.

**Procedure.** The stimuli were the same as in Experiments 1 and 2, as was their random assignment to each condition. As before, Experiment 3 consisted of a study phase, an interpolated task, an explicit recognition test, and an implicit test. The interpolated task consisted of trials during which subjects tried to identify perceptually degraded words to establish an exposure duration threshold of 50% correct identification. In each trial, a word was exposed, immediately followed by a mask consisting of @@@@ that lasted for 500 ms. Following each trial, the subject identified the word by naming. To establish the threshold, lists of 20 trials each were prepared. The exposure time of the words in the first list was 45 ms. In the second list, the exposure time was 15 ms. On the basis of the subject's performance on these two lists, the exposure time for words in the next four lists was increased and decreased until identification was about 50% correct. The last list was used to test the reliability of the threshold exposure time. This threshold exposure duration would then be used in the implicit perceptual identification test. None of the words used in this interpolated task appeared in any other part of the experiment.

Because we found no difference between old/new free recognition and forced-choice recognition in Experiments 1 and 2, the type of explicit recognition test administered in Experiment 3 was distributed evenly among the subjects. In the implicit memory test, 120 words were presented and followed by a mask, as in the interpolated task. On each trial, the subject attempted to identify the presented word by naming it. The exposure duration for the words was determined for each subject individually and set at the measured threshold value (that which yielded 50% correct identification in the interpolated task). The range of exposure durations used was between 17 and 26 ms.

As in the previous experiments, the stimulus list comprised 40 words initially presented during study, 24 words initially presented in the recognition test, 24 words initially presented at study and repeated at recognition, and 32 new words.

**Results and Discussion**

Perhaps because the intervening task was a little more demanding, the overall d' in Experiment 3 was 0.75 and the proportion of hits minus FAs was .26, both slightly lower than in Experiments 1 and 2. The values of d' and proportion of Hits minus FA for the yes/no recognition condition were .86 and .28, respectively; for the forced-choice recognition, they were .63 and .23, respectively.

As Table 4 shows, the pattern of repetition effects obtained in perceptual identification resembled those in lexical decision. The repetition effect was as large for items seen only at study as for items seen at study and at recognition. An ANOVA on the different categories of words showed an overall main effect, F(3, 69) = 8.51, MS_e = 109, p < .0001. Post hoc Tukey-A tests showed that new items were iden-

**Table 4**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unrepeated words</th>
<th>Seen once before</th>
<th>In study</th>
<th>In recognition</th>
<th>Seen twice before</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>50.0</td>
<td>63.5</td>
<td>60.0</td>
<td>62.6</td>
<td></td>
</tr>
<tr>
<td>SE_M</td>
<td>2.3</td>
<td>2.6</td>
<td>2.9</td>
<td>3.1</td>
<td></td>
</tr>
</tbody>
</table>
tified more poorly than any of the repeated old items \(p < .01\), with no significant differences among the old items.

The results of Experiment 3 show that the pattern of repetition effects on perceptual identification resembles that on lexical decision, despite the fact that the tests themselves are not similar and that performance is measured by accuracy in one test and by latency in the other. In both tests, the repetition effect for old words was not affected by recency or by the number of repetitions. The resemblance between the two tests was remarkable in that both showed the smallest repetition effect for words that were correctly rejected at recognition. Although subsequent studies might find differences in sensitivity between the two tests, at the moment there is no reason to regard one as being a more sensitive test than the other.

The similarity in the pattern of repetition effects should not be interpreted as evidence that performances on the two tests are correlated with one another or that performance on one test is statistically dependent on the other. Should a direct test of this prediction be conducted, it may well prove that stochastic independence holds for these two tests as it does for other implicit memory tests (Witherspoon & Moscovitch, 1989). Instead, the results suggest that a common component or components mediate performance on the two tests. At the moment, one can only speculate as to the identity of those components.

As noted in the introduction, not all implicit tests of memory produce comparable decay rates. Some decay within minutes, others take hours, still others take days, whereas some study only a slow decline whose limits have yet to be reached after more than a year. Experiments 1, 2, and 3, therefore, have been successful in demonstrating (a) comparable levels of performance on two types of recognition tests and (b) comparable patterns of performance on two types of implicit memory tests. In all cases, performance was well above chance or baseline levels. The question still remains of whether performance on implicit tests would remain above baseline if performance on either of the recognition tests were reduced to chance or near chance levels. Experiment 4 brings us closer to an answer.

**Experiment 4**

To reduce recognition even more, the delay between study and test in Experiment 4 was increased by 24 hr. Because there was no noticeable difference between the two types of recognition tests or between the two types of implicit tests that were used in Experiments 1, 2, and 3, we chose the yes/no procedure for recognition and the lexical decision task for the implicit test. We chose the lexical decision task for a number of reasons, not the least of which was that we have had extensive experience in using the task to study memory, reading, and other forms of priming. In addition, it has been shown that even after 48 hr, repetition effects in lexical decision decline only moderately (Scarborough et al., 1977) and recognition remains above chance. Amnesic patients, on the other hand, demonstrate normal repetition effects on lexical decision even when their recognition is at chance (Moscovitch, 1985). These data on people with normal and severely impaired memory can serve both as a yardstick against which the performance of our subjects can be compared and as a guide for interpreting the results of this experiment.

**Method**

**Subjects.** Participants in Experiment 4 were 24 undergraduates at the Hebrew University who met the same inclusion criteria as subjects in Experiments 1, 2, and 3. None of these subjects participated in the previous experiments.

**Stimuli and procedure.** The stimuli were identical to those used in Experiment 1, and assignment to each condition was random. The procedure was the same as that of Experiment 1, except that the delay between the interpolated task and the recognition test was 24 hr.

**Results**

**Explicit test (yes/no recognition).** Despite the 24-hr study-test interval the explicit recognition performance of most subjects was as good as in Experiment 3. The average value of \(d'\) was 0.77, which was not significantly different, \(t(44) = 0.88\), from the \(d'\) in Experiment 1 (which differed from the Experiment 4 only in the time elapsed from study to the recognition test). The proportion of hit minus the proportion of FA was 0.32 (see Table 5). An ANOVA revealed that, as in Experiment 1, correct responses (hit and CR) were faster than incorrect responses (miss and FA), \(F(1, 23) = 18.51, M_{S_e} = 2.283, p < .0003\), and responses to new words (hit and miss) were as rapid as to old words \((F < 1)\). Accuracy significantly interacted with stimulus type, \(F(1, 23) = 10.51, M_{S_e} = 5.634, p < .004\). Post hoc Tukey-A comparisons showed that hit responses were significantly faster than miss responses, whereas CR and FA responses were equivalent.

**Implicit test (lexical decision).** The pattern of repetition effects observed in the implicit memory test was similar to that observed in Experiments 1, 2, and 3 (see Table 6).

An ANOVA showed a significant main effect of repetition, \(F(3, 69) = 14.52, M_{S_e} = 620, p < .0001\). Post hoc Tukey-A comparisons showed that RTs to new words were slower than RTs to words in any of the other categories and that the number of repetitions or the recency of the initial presentation relative to test made no difference. The RTs to repeated nonwords in the implicit test (574 ms) were similar to RTs to new nonwords (577 ms).

| Table 5 |

**Mean Reaction Times (RTs) and Percentages in the Explicit Recognition Test (Experiment 4)**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unrepeated words</th>
<th>Repeated words</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CR</td>
<td>FA</td>
</tr>
<tr>
<td>RT</td>
<td>810</td>
<td>803</td>
</tr>
<tr>
<td>(SE_{M})</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>%</td>
<td>63.5</td>
<td>36.5</td>
</tr>
</tbody>
</table>

*Note.* Mean RTs are expressed in milliseconds; \(d' = 0.77\). CR = correct rejection; FA = false alarm.
Table 6  
Mean Reaction Times (RTs) in the Implicit Test  
(Experiment 4)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unrepeated words</th>
<th>Repeated words</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In study</td>
<td>In recognition</td>
<td>Both</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>556</td>
<td>527</td>
<td>518</td>
<td>512</td>
<td></td>
</tr>
<tr>
<td>SEM</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Note. Mean RTs are expressed in milliseconds.

RT data from the two groups yielded a significant main effect of repetition, $F(3, 66) = 12.34$, $MS_e = 845$, $p < .0001$. The average repetition effect in the high and low $d'$ groups was 56 ms and 26 ms, respectively. A t test for independent groups revealed that this difference was statistically reliable, $t(22) = 2.46$, $p < .025$. Subsequent comparisons between the two groups, in each of the conditions separately, yielded significant group differences for items that were previously presented only at study, $t(22) = 4.08$, $p < .05$ and only at recognition, $t(22) = 7.03$, $p < .01$, but not for the doubly repeated items ($t < 1$).

Ideally, we would have preferred that recognition in the entire group be at least as low as in the sample of 12 subjects. However, lowering recognition to chance is very difficult. Before interpreting these results, it is important, therefore, to establish that the performance of the 12 subjects with poor recognition was not artifactual. First, these subjects were selected at random from a pool of undergraduate volunteers. Their performance was in other ways comparable with that of the other subjects. Their mean RT for lexical decisions at study (551 ms) was not reliably different from that of the other group (609 ms; $p > .1$). There is no reason to believe that something other than poor memory accounts for their performance, although the shorter RTs among the poor recognizers may suggest a relationship between rapid encoding and forgetting. Second, it is noteworthy that the largest proportion of subjects with poor memory, at least twice that in any other condition, was found in Experiment 4, in which delay and interference were highest. This suggests that factors affecting forgetting rather than only encoding are responsible for subjects’ poor memory.

Moreover, there is already ample evidence in the literature that poor encoding, to the point that the subjects may not be aware that a stimulus was presented, results in normal, often long-lasting, repetition effects on implicit tests of memory. Our stimuli, on the other hand, were clearly visible and certainly of sufficient duration for the subjects to respond with nearly perfect accuracy in lexical decision on the initial presentation. Poor encoding is not likely to be the cause of the poor repetition effects seen in our subjects. The most parsimonious explanation for their poor performance is that they were more prone than others to forget the target items.

The results of Experiment 4 suggest, therefore, that when recognition approaches chance levels, long-lasting repetition effects are substantially reduced or even eliminated. For the group of subjects with a low recognition performance ($d' < 0.61$), repetition effects were absent for words seen only once, either at study or at recognition. In contrast, a single previous presentation was sufficient to produce repetition effects that were not significantly reduced over a period of 24 hr in subjects with a higher recognition level ($d' > 1$). In a similar manner, presenting items twice, at study and at recognition, led to substantial repetition priming effects, even in subjects with poor memory, presumably because those items were remembered better.

Table 7  
Mean Reaction Times (RTs) in the Implicit Test for Subjects With Low $d'$ (Experiment 4)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Unrepeated words</th>
<th>Seen once before</th>
<th>Seen twice before</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In study</td>
<td>In recognition</td>
<td>before</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>540</td>
<td>521</td>
<td>515</td>
<td>497</td>
<td></td>
</tr>
<tr>
<td>SEM</td>
<td>22</td>
<td>24</td>
<td>30</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

Note. $N = 8$. Mean RTs are expressed in milliseconds.
Figure 1. Mean differences in reaction time on the implicit test between new items and old items seen at study, at recognition, or on both occasions for subjects with high $d'$ and low $d'$ scores on explicit recognition tests in Experiments 1 and 4.

On the basis of evidence from our own study as well as other studies, we can chart the time course of repetition effects for words relative to that for recognition. First, there is a steep initial decline in repetition effects that occurs within seconds to minutes after stimulus presentation (Bentin & Moscovitch, 1988; Ratcliff et al., 1985; Sloman et al., 1988; but see Dannenbring & Briand, 1982); this is followed by a relatively long-lasting, stable period for repetition effects in comparison with explicit recognition performance, which gets steadily lower; finally, when recognition approaches chance levels, repetition effects are also noticeably reduced or fall to zero.

General Discussion

The primary question addressed by our study was whether normal people would perform above baseline levels on implicit tests of memory when explicit recognition was reduced to chance or near chance levels by forgetting. The first point to note is that lowering recognition to that level in normal people proved to be difficult. This fact is not appreciated very much, as investigators are more impressed by the initial rapid forgetting than by the longevity of memories as revealed by explicit recognition tests. This is a worthy topic for future research but not one that we can pursue in this article.

Once recognition approaches chance levels, however, our evidence suggests that there are differences in implicit test performance between amnesic patients and normal people whose recognition was reduced by forgetting. Unlike amnesics who often have normal repetition effects in the absence of recognition (see, e.g., Moscovitch, 1985), normal subjects, who score at or near chance on recognition, have repetition effects that are substantially reduced or even eliminated.

The reduction in performance on implicit tests with poor explicit recognition due to forgetting also contrasted with findings from studies that manipulated encoding to lower performance on explicit tests. As we noted, when performance on explicit tests was reduced to chance by degrading the stimulus at study or by diverting attention from it, repetition effects were not significantly reduced but, as with amnesics, remained reliably above baseline (Bentin in press; Bonnano & Stilling, 1986; Eich, 1984; Mandler et al., 1987; Merikle & Reingold, 1991; Seaman et al., 1983; Seaman et al., 1984).

Together, studies of normal and amnesic subjects suggest that the following conditions must be met for explicit remembering: (a) The stimulus event must first be consciously and fully apprehended. The cited encoding studies suggest that without full conscious apprehension of the stimulus at encoding, explicit remembering is very ineffective. Conscious apprehension is not, however, a sufficient condition for explicit remembering, as the evidence from amnesic subjects indicates. Amnesic subjects can attend, perceive, and apprehend the stimuli, yet their conscious recollection is severely impaired. (b) Clearly, an additional process is necessary to make this information available for long-term explicit remembering. This process, often termed consolidation, is identified with the operation of the hippocampus and related limbic structures (Scoville & Milner, 1957; Squire, 1987). In amnesic subjects, the deficit in conscious recollection arises either because the apprehended information is not delivered to the hippocampal system or because that system is absent or malfunctioning. The information, however, is not stored in the hippocampus. (c) Retrieval cues are necessary to recover the consolidated information and deliver it to consciousness. Forgetting occurs either because the consolidated information has decayed or because retrieval cues are ineffective or both. These points and
the ones made in the discussion that follows are elaborated in Moscovitch (1989, 1992) and Moscovitch and Umiltà (1991) where a neuropsychological model of memory is presented.

None of these processes is necessary for storage and retention of information on implicit tests. For normal performance on implicit tests, it is only necessary that the stimulus be picked up by neural structures necessary for decoding it. The mere pickup of information by these structures, even without full conscious apprehension (as the encoding studies show), is sufficient to form long-term records or engrams in memory. These records are not consciously accessible. To become accessible, the input must be both consciously apprehended and delivered to the hippocampal system.

Consolidation, we suggest, involves the binding of engrams into memory traces by the hippocampus. The memory trace consists of a collection of bound perceptual (and possibly conceptual) engrams related directly to the stimulus, along with other engrams that encode semantic and episodic information about the event in which the stimulus was embedded. Performance on implicit tests requires only that the repeated stimulus reactivate the long-term records by engaging directly the same processes involved in their formation. In explicit tests it is necessary that the bound engrams that formed the memory trace be reactivated indirectly through the hippocampal system.

According to our framework, the records or engrams that mediate performance on implicit tests are a subset of the engrams that form the memory trace that mediates performance on explicit tests. Access to the record is achieved directly by the stimulus or indirectly, by the hippocampus, through reactivation of the entire memory trace. That is, the effective cue for performance on implicit tests is the stimulus itself or something closely related to it, whereas the effective cue to the memory trace is richer, containing information that was encoded along with the stimulus as well as any other relevant information that can be brought to bear at the time of retrieval. Forgetting occurs either because the cue is no longer effective or because engrams or traces have decayed. Because the cues are different, memory for explicitly remembered items is independent of memory for implicitly elicited information. Item amnesia can occur without concomitant effects on implicit tests for the same target.

The time course of repetition effects may also differ from that of recognition, not because the memory trace and engram are different but because they are accessed differently. The effective stimulus for reactivating the perceptual process presumably does not vary, thereby leading to repetition effects that remain stable over prolonged periods. Whether semantic records are equally long-lasting is not known. The effective cue on explicit tests of memory, such as recognition, however, is not simply the stimulus itself but the complex event associated with that stimulus. With time, the context becomes less readily available to act as a cue for the memory trace. Thus, performance on explicit tests comes to rely more and more on an impoverished cue that consists primarily of the stimulus item itself and the sense of familiarity to which it gives rise.

By this account, with increased forgetting, performance on explicit and implicit tests becomes driven primarily by the same effective cue. When recognition approaches chance or near chance levels, it is either because the stimulus item alone is no longer an effective cue for evoking the memory trace or because the trace itself may have dissipated or been altered with time. Our finding of a drop in repetition effects as recognition approaches chance is more consistent with the second alternative. Because there is no reason to believe that the effective stimulus for reactivating the record (i.e., driving performance on implicit tests) is altered, the drop in repetition priming effects as recognition approaches chance probably occurs because of a change in the engrams that form the trace. This situation is different from that in which chance performance on explicit tests results from poor encoding. In that case, repetition priming effects are obtained because testing occurs shortly after the study phase when the engram is still "fresh" and can support performance on implicit tests. According to our proposal, substantially increasing the study-test interval should reduce or eliminate the repetition priming effect, even when encoding is poor.

The correlations between performance on some explicit and implicit tests of memory, but not on others, has been reported by Perruchet and Baveux (1989). Their point, however, was to show that the pattern of dependence results from the extent to which component processes are common to different tests. Our proposal goes further and suggests that even on tests that seem independent at high to moderate levels of performance, at near chance levels performance drops in both types of tests.

Our account clearly borrows some of Tulving's (1983, 1985) ideas about ephoric processes in recognition. There are also similarities with Mandler's (1980) proposal that two component processes contribute to recognition: interim elaboration, which is context dependent, and intratext integration, which leads to a sense of familiarity devoid of context. Mandler and others (Gardiner, 1988; Gardiner & Java, 1990) have associated the latter process with performance on implicit tests of memory. According to this view, long-lasting recognition is mediated by processes associated

---

1 We use the term record (Kirsner & Dunn, 1985) to refer to a modification of the neural circuitry in response to stimulation. The modification contains information about the stimulating event and has processing consequence so that identical, and perhaps related, stimuli can be processed more efficiently and leads to above baseline performance on implicit tests. We reserve the term engram to refer to the informational content of the record that can be accessed or reactivated to contribute to recollection. Whether the engrams are purely perceptual or whether they also have semantic content and show some contextual dependencies is currently disputed (Masson & Freedman, 1990; Roediger, 1990; Tulving & Schacter, 1990).

2 The rapid initial decline on implicit tests may have more to do with recruitment of response factors and episodic memory factors than with reactivation of perceptual records. That is, for a short while after stimulus presentation, the subject may remember the response associated with a particular stimulus, which would lead to faster or more accurate responses on repetition.
with implicit memory. What distinguishes our account from Mandler’s is that we believe that both types of recognition are mediated by the hippocampus and its related structures. Put in purely psychological terms, we believe (contrary to Mandler) that recognition, even when it is based only on a sense of familiarity, is an explicit test of memory. The process that gives rise to a sense of familiarity (insofar as that sense is inextricably a product of the process rather than an inferred or attributed byproduct of it; Jacoby, 1983; Johnston, Dark, & Jacoby, 1985) is different from the process that is involved in producing repetition effects.

The distinction we have drawn between our proposal and those of Mandler, Gardner, and Jacoby is not trivial. It implies that recognition, even when it is based on familiarity, is mediated by processes that are distinct from those mediating performance on implicit tests of memory. At the moment, we must admit that there is little evidence in the literature to distinguish among the alternative proposals. Some findings, however, are more consistent with our proposal and we review these briefly.

It is significant to note that performance based only on familiarity was well above chance (Gardiner, 1988; Gardiner & Java, 1990). This was true also, and especially, when nonwords were used. By themselves, these results are consistent with all of the proposals. When evidence from amnesic subjects is taken into account, however, it becomes possible to distinguish between the predictions made by the various proposals. Performance on recognition tests can be at chance in amnesics even when repetition effects are normal for the same material. In addition, repetition effects for nonwords are very difficult to obtain in amnesic subjects. If processes leading to recognition judgments based on familiarity were the same as those mediating repetition effects on implicit tests of memory, we would expect the performance of amnesics to resemble that of normal people more than it does. Such similarity as does exist between familiarity effects in recognition and repetition effects on implicit tests of memory probably occurs because both are based on reactivation of the same engram or record.

Bentin and Moscovitch (1990; Bentin, Moscovitch, and Heth, 1992) found electrophysiological evidence of repetition effects in both implicit and explicit tests of memory. This result is consistent with the idea that a common record or engram is reactivated in both cases. The additional observation that repetition effects on explicit and implicit tests are associated with different components of the evoked response suggests that different neural processes mediate their occurrence on the two types of tests. It would be important to determine whether one is associated with recognition due to familiarity and the other with repetition effects on implicit tests.

To summarize, our finding, that performances on implicit and explicit tests are interdependent when explicit recognition is very low, suggests that they are both mediated by common engrams. In implicit tests, only the record is accessed by using the repeated stimulus as a cue. In the explicit test, reactivation of the entire memory trace is necessary, a process which is performed by the hippocampal system and which requires conscious processing.

References


Psychology: Human Learning and Memory, 2, 554–565.

Received April 22, 1991
Revision received April 6, 1992
Accepted April 8, 1992