In Search of Mood-Dependent Retrieval

Gordon H. Bower and John D. Mayer
Department of Psychology, Jordan Hall
Stanford University, Stanford, CA 94305

In six experiments, we examined mood-dependent retrieval and mood-congruent learning. In each experiment, subjects learned two or more different word lists in different (happy or sad) moods, and were then tested for recall of the lists while in one or the other mood. In five of the six experiments, the lists were composed of pleasant and unpleasant items. Faster learning of mood-congruent material was observed in experiments that tested it; this effect was modifiable by explicit learning instructions in Experiment 3. Mood-dependent retrieval (MDR), i.e., better recall of the list learned in a matching mood, was not observed in the first three experiments, where mood was a prevailing background incidental to the list learning events. MDR did appear in a fourth experiment which induced subjects to attribute the cause of their emotional reaction to the material being learned. However, that result failed to replicate in Experiment 5. The sixth experiment failed to replicate a finding by Eich (in press) of MDR when subjects generated the items to be remembered. The inconsistent results on MDR are discussed in terms of the complexities of laboratory experiments on induced emotions.

Context-dependent memory exists when memory retrieval improves with increasing similarity between a subject's external or internal context at learning and at recall. For example, if a subject learns a word list while drunk, memory for the list should improve when the subject becomes drunk again, other factors being equal. Context-dependent memory has been observed in many situations, including material learned in different drug states (Eich, 1980), different rooms (Metzger, Boschee, Haugen, & Schnobrich, 1979; Smith, 1979), different times of day (Holoway, 1978),

---

Authors' Notes: We gratefully acknowledge the assistance of Maureen Holt, Dan Morrow, Deedee Perez, and Jamie Uyehara, who assisted us in preparing and conducting the experiments. Jamie Uyehara also assisted in editing the manuscripts.

This research was supported by a grant, MH-139540, to G. Bower from the National Institute of Mental Health. J.D. Mayer was supported by National Institute of Mental Health post-doctoral fellowship, MH-08978. Mail correspondence to the first author or to the second author at Psychology Division of Natural Science, State University of New York, Purchase, NY 10577.

Kuiken, D. (Ed.). (1989) Mood and Memory: Theory, research, and applications. [Special Issue]. *Journal of Social Behavior and Personality*, 1989, Vol. 4, No. 2, 121-156.

and even when learned above and below water (Godden & Baddeley, 1975). Such effects are typically small but, with the exception of some recent reports (Fernandez & Glenberg, 1985), were regularly detected (e.g., Smith, 1982; Smith, Glenberg, & Bjork, 1978). Mood-dependent retrieval (MDR) is a special case of context-dependent memory. The MDR effect exists if memory retrieval increases with the degree of similarity between a subject's *mood* at learning and at recall (Bower, 1981; Weingartner, Miller, & Murphy, 1977).

Previous Findings

Bower, Monteiro, and Gilligan (1978) reported three experimental attempts to produce MDR, all using hypnotic suggestions to induce moods. In the first two experiments, subjects learned to recall a single word list while feeling happy or sad. Later, half of each group of subjects recalled the list while feeling happy, the other half while feeling sad. Neither experiment yielded any evidence for MDR. A number of other experiments using one-list designs have also found either no evidence for MDR (Isen, Shalker, Clark, & Karp, 1978; Mecklenbraucker & Hager, 1984; Schare, Lisman, & Spear, 1984) or asymmetric MDR only (Leight & Ellis, Exp. 2, 1981; Macht, Spear, & Levis, 1977).

In the third experiment[1] of the Bower et al. (1978) report, a memory interference design was used in which each subject learned two word lists, one while happy, the other while sad. Later, subjects recalled both lists, some while feeling happy, others while sad. These subjects appeared to show mood-dependent retrieval. Recall percentages averaged 70% when the learning and testing moods matched, but were considerably lower, 46%, when they mismatched. Thus, it was concluded that MDR could be produced when interference amongst memories was high, under circumstances which in some ways resembled ecologically valid conditions of multiple learning in multiple moods and contexts.

Initially, it appeared that the interference design was sufficient to regularly obtain MDR. The effect was found in an undergraduate Honors thesis by Brett Thompson (directed by, and reviewed in, Bower, 1981) and a study by Bartlett and Santrock (1979). These and related findings provided some of the basis for review papers by Bower (1981) and Bower and Cohen (1982). Subsequently, several researchers using the two-list interference design reported MDR (Bartlett, Burleson, & Santrock, 1982; Goerss & Miller, 1982; Schare, Lisman & Spear, 1984).

[1] *This third experiment was Stephen Gilligan's first-year graduate project supervised by Bower. Such projects have the approximate scope of a Masters' thesis.*

Introduction to the Present Experiments

While it appeared that the MDR effect had been well established shortly after its initial demonstrations, some disturbing findings began to appear as well. Ellis (1983) found no MDR in the two-list design when the happy and sad moods were induced by having subjects view a funny or a sad movie segment. Bayer (1982, personal communication), working in Professor Spanos' lab at Carlton University, was unable to obtain the MDR effect using hypnosis in a partial replication of the two-list interference experiment. Later, Wetzler (1985) reported a failure to obtain MDR in a replication of the two-list design.

In light of the doubts raised by the Ellis and Bayer failures, it appeared useful to attempt a full replication of Experiment 3 of the 1978 Bower et al. report. Such a replication was conducted in Experiment 1 below, with minor variations in detail.

While our primary focus will be on MDR, the experiments provide opportunities to observe two other mood effects that have been reported. The first effect is mood-congruent learning, that people better learn material that agrees in content with their mood (Bower, Gilligan, & Monteiro, 1981). Thus, from studying a mixed list of pleasant and unpleasant words, happy subjects should learn more pleasant words, whereas sad subjects should learn more unpleasant words. The second effect to be examined is whether sad subjects learn more slowly overall than do happy subjects. This effect, reported for psychiatrically depressed subjects (Weingartner & Silberman, 1982), has also been observed for normal subjects induced to feel sad in the laboratory (Ellis, Thomas, & Rodriguez, 1984).

EXPERIMENT 1[2]

Method

Subjects

Student volunteers, who responded to posted announcements on campus or who were enrolled in the Introductory Psychology course, were prescreened for hypnotic ability on the Harvard Group Scale of Hypnotic Susceptibility (HSHS; Shor & Orne, 1962) or its equivalent. All subjects in the experiment (N = 48) scored 6 or higher on the Harvard

[2] *Study 1 was reported in abbreviated form in Bower and Mayer (1985) so as to inform the research community as quickly as possible of difficulties in replicating MDR. The study is presented here in its original context, as part of a series of studies searching for the MDR effect.*

scale (M = 9.13). Four were dropped from the experiment for non-compliance or errors. Subjects were either paid five dollars or given course credit for their participation in the main experiment.

Word Lists

The words learned were 16 pleasant and 16 unpleasant abstract nouns, with scaled concreteness less than 3.0 on a seven-point scale (Paivio, Yuille, & Madigan, 1968). Seventeen of the thirty-two words were on the list used in the original study by Bower et al. (1978). The replacements were made in order to eliminate unfamiliar words and clarify the affective quality of the words. Eight pleasant and eight unpleasant words were randomly assigned to each of the two lists.

Design

The design closely followed that of the original study. During a first mood, subjects learned the first of the 16-word lists by free recall; during the second mood, they learned the second list. After a filled interval, a third mood was induced, matching the first or second. The subjects then tried to recall aloud all words from the first list, and after that, all words from the second list. In addition, during this same final mood, a recognition memory test for both lists was administered. List order during learning was counterbalanced across subjects and conditions.

Each subject received three successive mood inductions which differed according to his or her experimental condition. In the control condition, subjects learned their first list in one mood (e.g., happy), their second list in the same mood, and then performed the recall and recognition tasks in yet a third instance of the same mood. This condition will be symbolized as AAA, to represent the fact that the mood was of type A (either happy or sad) during all three mood periods. Our second condition will be symbolized as ABA, to represent the fact that the person's mood was reversed after learning the first list and before learning the second list. These subjects were returned to their first-list mood at the time recall of both list was requested. The third condition will be symbolized as ABB, to represent the fact that the subject's mood was reversed for learning of the second (B) list, and this second-list mood (B) recurred for the final recall and recognition of both lists. Half the subjects in each of the three conditions started with a happy mood, and half with a sad mood, making a total of six mood conditions. Whenever a mood was repeated, the subject was asked to create the mood by dwelling on a happy or sad memory which was different in content from the mood-memory used to induce that mood earlier.

Procedure

After prescreening for hypnotizability, subjects were tested individually. A hypnotic trance was induced by the eye-closure technique (see instructions in hypnotic test Form C, Weitzenhoffer & Hilgard, 1962). The appropriate mood was induced by asking subjects to recall a happy or sad emotional event from their lives. They were asked to experience the emotion of that recollection, and to adjust its intensity until it was intense but not overwhelming (described as a 6 or 7 on a ten-point scale). Subjects typically indicated such a mood level within about a minute. Subjects were then told they would be able to hold the emotion deeply for as long as needed while they learned the first list of words.

Subjects then heard the experimenter read one of the 16-word lists aloud twice, in a different random order each time, allotting 3 seconds per word. They were then asked to recall orally as many of the words as they could, in any order, for 2 minutes. At the end of the recall period, the experimenter read the list a third time, in yet a different random order, and the subject was again given 2 minutes to free recall it orally. After this 10-minute session subjects were then asked to rate their mood on a ten-point scale. The first mood was then removed, the second mood was induced, and the procedure was repeated as before for learning the second list. After this, a post-hypnotic suggestion for easy re-entry into hypnosis was given; subjects were then removed from trance.

Upon awakening, participants took a short break. For the remaining part of a 10-minute period (measured from the time subjects awoke from hypnosis), letter-cancellation tasks were administered. Subjects were then instructed to reenter trance, and to signal the experimenter when in trance. After the subject's signal, the experimenter administered trance-deepening instructions, and the final mood was induced by asking the participant to recall a happy or sad event from their life which was different from the memories they had used to induce the earlier moods.

Once the third mood induction was completed, the subject was allotted 3 minutes to recall orally the words from the first list, and immediately after, three more minutes to recall orally the words from the second list. These recalls were followed by a third mood rating. If mood intensity had dropped, subjects were instructed to reintensify their mood. The recognition test then ensued, with instructions given halfway and three-quarters through the test to deepen the mood. The recognition memory test consisted of 64 words written on a sheet: the 16 words from each learned list along with 32 distractors, selected from a similar word population. The subject's task was to identify each word by checking a blank, indicating its having been either on the first, second, or neither list that they had heard.

Subjects were then awakened, their mood was normalized, and they were debriefed. They were asked to telephone should they suffer any disturbing after-effects. (None ever did; for a discussion of the safety of hypnotic mood-inductions, see Mayer, Mauro, & Trebino, 1987).

Results

Efficacy of Mood Induction. Mood intensity averaged 5.99, 6.58, and 5.94 (out of 10) across the three lists. Moreover, subjects rated the similarity of their mood during testing to the "same mood" of training at 5 (on a 7-point scale) and to the "different mood" of training at 2. Thus, conditions existed for MDR.

Mood dependent retrieval. The two lists of words were learned equally well, so this list factor will be ignored. Table 1 reports three alternative dependent measures of retention. The first measure, "percent retention," is the final correct recall for a given list divided by the number of words correctly recalled at the end of original learning for that list. The second, "recall composite" measure, will be discussed below. The third measure, "recognition," was calculated for List 1 as the probability that when the subject said List 1 the test item was indeed from List 1, minus the probability that it came from the other list or the distractor pool. This measure was calculated similarly for recognition of List 2. This is roughly equivalent to "hits" minus "false alarms" for each list.

For MDR to occur, subjects in the AAA and ABA condition must retain more of the first list than subjects in the ABB condition. Table 1 shows the average retention percentages for List 1 in each of the three conditions. The three retention percentages are not ordered as expected, and in any event do not differ reliably.

Similar analyses were performed on retention percentages for List 2, where, according to MDR, subjects in the ABB and AAA conditions should recall more than those in the ABA condition. Table 1 shows that the three conditions for List-2 retention are ordered as expected but the differences are too small to be statistically significant. Also, difference scores for List 1 vs. 2 retention showed no reliable differences between the ABA and ABB groups (see Table 1).

We thought perhaps that MDR might be more clearly manifested if subjects experienced the moods strongly. Therefore, we carried out an analysis of mood-dependency in percentage retention scores for just those subjects who rated all three of their moods (during Lists 1, 2, and Retention) at an intensity of five or above. Difference scores for retention percentages for the 22 (of 48) subjects who passed this stringent mood criterion showed no reliable MDR ($t(20) = 0.49$, n.s.).

Other Memory Measures

Several similar analyses were conducted with different dependent measures. These measures included several alternative values in the numerator of the recall ratio. One value was the correct list recall plus extra-list intrusions. A second value consisted of the correct recall of words on a list, regardless of whether they were reported while attempting recall of List 1 or List 2 (but excluding any repetitions of a word). A third numerator value was the total number of words produced at recall, including both extra-list instructions and repetitions. None of these measures yielded a statistically significant effect in the predicted direction. A fourth measure which showed MDR ($p < .02$, one-tail in the Difference Scores of Table 1) was a recall composite which used the correct list recall during the final recall period, plus intrusions (counting repetitions) of those words while recalling the other list, plus extra-list intrusions, all divided by the recall score for that list at the end of original learning. This composite is roughly the total production of List 1 (or List 2) items anywhere plus distractors when recalling that list. This measure would be high, for example, if subjects repeated the same words when recalling either list.

Recognition Memory. Our recognition test comprised a mixture of item recognition and list discrimination. The corrected recognition scores in Table 1 show the likelihoods that the subject's judgments, saying "from List 1" or "from List 2," were correct minus the likelihood of a "false alarm" with that list response. Examining these scores, several effects are evident. First, subjects performed slightly better in remembering items from the second list than from the first one. Second, group differences appeared for each list. These were in the expected order (according to MDR) for List 2, but were the reverse of the expected order for List 1. Another summary of the data is to say that the ABB subjects outperformed the ABA subjects on both lists. Thus, the recognition and list discrimination results provide no evidence in favor of mood-dependent memory.

Influences of Happy and Sad Mood on Learning. To examine the influence of happy versus sad mood on learning, the free recall scores at the end of original learning for each list were entered into an ANOVA with list (first vs. second), and word content (pleasant vs. unpleasant). No main effects were significant. Thus, neither happy nor sad moods showed any advantage in overall level of learning, nor were pleasant words recalled overall more than unpleasant words. However, a statistically significant interaction occurred between mood and word content ($F(1,92) = 26.07$, $p < .0001$). Happy subjects learned 72% and 55% of pleasant and unpleasant words, respectively. Sad subjects learned 69%

TABLE 1 Group Means and SD's for Retention of Lists 1 and 2, and F-tests for Group Differences

	AAA	*ABA*	*ABB*	*F(2,45)*	*p*
List 1 Retention					
% Retention	.40 (.18)	.49 (.21)	.53 (.21)	1.71	.19
Recall comp.	.49 (.19)	.73 (.35)	.69 (.20)	4.68	.01
Recognition	.53 (.22)	.40 (.13)	.60 (.19)	5.15	.01
List 2 Retention					
% Retention	.71 (.18)	.58 (.16)	.67 (.21)	2.05	.14
Recall comp.	.87 (.25)	.68 (.26)	.90 (.28)	5.18	.01
Recognition	.55 (.23)	.47 (.17)	.61 (.18)	2.05	.14
Difference Scores (List 1 - List 2)					
% Retention		-.10 (.30)	-.14 (.27)	0.32	.57
Recall comp.		.05 (.41)	-.22 (.34)	6.03	.02
Recognition		-.07 (.12)	-.01 (.11)	3.12	.09

and 60% of unpleasant and pleasant words, respectively. This is a mood-congruent learning effect of the type seen before (Bower et al., 1981). In contrast to findings by Clark and Teasdale (1985), we found no sex-differences in mood-congruent learning. The percentage of mood-congruent vs. incongruent words learned by men (72% vs. 58%; N=21) was nearly equal to that learned by women (69% vs. 57%; N=27).

Influences of Happy and Sad Mood on Retention. To examine the influence of happy and sad mood on retention, a recall mood (happy vs. sad) x order (first learning or second learning) x learning mood (happy vs. sad) x word content (pleasant vs. unpleasant) ANOVA was conducted, for the 32 subjects who had two different moods in learning (ABA and ABB conditions). The dependent measure was percentage retention, which was final recall conditionalized on original learning. Such retention percentages showed no mood-congruity effect. In fact, none of the main effects were significant. Two higher-order interactions reached significance, one of which reflected better retention for the second list learned (a recency effect). The second significant interaction (of triple order) was not clearly interpretable.

Discussion

To summarize our results, the mood manipulations generated the intended moods. Depending on the recall measure used, there either was no evidence or only slight evidence for mood-dependent retrieval. Most measures showed no differences. Those statistical tests which had yielded significant results in the earlier study (Experiment 3 of Bower et al., 1978) indicated insignificant results here despite our having twice as many subjects per condition. A mood-congruity effect was found in the present study; this effect had not been found in the earlier experiment. A reanalysis of the earlier data, however, showed trends toward mood-congruity. About half the words used in the original experiment were replaced in our study to enhance the pleasant versus unpleasant contrast of the words in the lists, and this might have enhanced the mood-congruity effect.

Conflicting Evidence for Mood-Dependent Retrieval. The present experiment differs from the earlier one in several ways, any one of which could conceivably have caused the different outcomes. Perhaps some aspect of the word list changes caused MDR to disappear. If it can be so easily affected by subtle, unsuspected variations in materials, then MDR is surely an effect of little generality. Perhaps altering the task interpolated between learning and the retention test between the two studies created the difference. The original task was to read and study a book chapter. This was replaced here by a series of letter cancellation tasks, which might have resulted in less verbal interference in the present experiment, with a corresponding enhancement of a list-recency effect, as observed here. This stronger list-recency effect might have masked an MDR effect.

Our discomfort with the findings of the first experiment led to Experiment 2 which will be reported in only abbreviated form (for fuller details, contact the authors).

EXPERIMENT 2

Experiment 2 attempted to promote MDR by increasing the amount of interference surrounding the target memories. Subjects learned three lists: the target list (which subjects learned in a happy or sad mood) was the middle list, whose retention should suffer from both proactive and retroactive interference. Hopefully, mood-matching at retrieval should help overcome such interference.

Method

Twenty-four highly hyponotizable subjects were selected from the same source. They learned three lists of 16 abstract nouns while hypnotized. Neutral moods were induced for learning lists 1 and 3; happy or sad moods (M = 12 each) were induced during second list learning. Each list was presented on tape in three random orders at 3 sec/per word, followed by written recall. At the end of learning the three lists, half of each group of subjects was induced to feel happy or sad, and written recall of the second (critical mood) list was requested. After mood normalization, subjects were dismissed.

Results

The mood inductions were successful (6.3 average rating on 10-point scale); also, more interference in the form of extra list intrusion errors occurred here than in Experiment 1. Despite this, an MDR effect did not appear.[3] The level of final recall of List 2 was 22% for like-mood subjects versus 27% for opposing-mood subjects. Percentage retention conditionalized on original learning was 45% for like-mood subjects and 49% for opposing-mood subjects. These results were not changed when we examined only subjects scoring high on the mood inductions. Thus, MDR did not appear.

Degree of original learning of list 2 was identical for happy and sad learners. No test for mood-congruity was possible since the learning items here were all affectively neutral.

To summarize, increasing interference surrounding the target list did not produce an MDR effect. In another attempt, Experiment 3 reverted to the two-list design of Experiment 1, but was designed to learn more about the mood-congruity effect.

EXPERIMENT 3

Experiment 1 showed a strong mood-congruity effect, of subjects learning more items whose affective content agreed with their emotional state. The reliability of the mood-congruity effect persuaded us to design our next experiment to examine that effect, but with another test for mood-dependent retrieval also included.

Experiment 3 examined the extent to which mood-congruent learning was under voluntary control. We directly manipulated the attention and learning strategy of the participants. One group of subjects received no special instructions about which items to learn selectively. A second group was instructed to try to concentrate especially on learning the words that agreed in content with their mood. A third group was instructed to concentrate on learning the words that disagreed in content with their mood. Each subject learned two lists of mixed pleasant and unpleasant words, one list learned while happy, the other while sad. During the final retention test, a third happy or sad mood was admini-

[3] The design of Experiment 2 precludes calculation of the "Composite retention measure" used in experiment 1. For what it is worth, like-mood subjects produced about the same number of intrusions (1.67) as did opposing-mood subjects (1.75) in Experiment 2.

stered and was varied so as to permit another test for mood-dependent retrieval.

We assumed that the two groups instructed to concentrate on particular words during learning would do better in learning those words. We predicted further, however, that individuals in the "opposing-content" condition would learn fewer words overall, because they would be working against the natural tendency to concentrate and elaborate upon same-mood material. We also predicted that the "opposing-content" subjects would have inferior moods to the other groups because concentrating on opposing-mood words should partially offset their hypnotically-induced mood.

Method

Subjects

Subjects were 24 students prescreened for high hypnotizability from the General Psychology course. Eight were assigned to the three treatment conditions.

Stimuli

Two word sets composed of 16 high-concreteness, high-emotion content words were selected from the Paivio et al. (1968) word lists. The pleasant and unpleasant word lists were matched on imagery ($M = 5.6$ vs. 5.2, respectively), concreteness ($M = 4.8$ vs. 4.7), meaningfulness ($M = 6.6$ vs. 6.1), and frequency ($M = 3.7$ vs. 4.0).

Design

This experiment required three mood periods. Subjects learned a different list during each of the first two moods, and then recalled both lists during a third mood. Subjects received a happy and sad mood in one of two possible sequences over the two learning lists, followed by a mood at final recall. Thus, the moods were either ABA, or ABB, with initial mood (happy or sad) counterbalanced over conditions. One-third of the subjects in each mood-sequence were given no special instructions (Control Group), one-third were told to concentrate on learning the mood-congruent words (MC Group), and one third were told to concentrate on learning the mood-incongruent words (MI Group). Two subjects were assigned to each of the twelve conditions in the 2 x 2 x 3 factorial design.

Half the subjects in each condition began with a happy mood, half with a sad mood. Hypnosis, mood instruction, and word-list presentations were delivered by a tape recording. Subjects were told to learn the word lists for free recall. The MC and MI groups were told "try to lean as many words as you can, but you should make a special effort to learn the words that are [*same/or/opposite*] in emotional content as the mood you are in at the time..." Learning consisted of presenting each word list randomly three times (3sec/word) before a 2 minute recall was requested. Moods were shifted between learning of the first and second list, and between the second list and the retention test. Finally, subjects were asked to recall all words from both lists in any order; they were given 3 minutes. Thereafter, their moods were neutralized, they were debriefed, and dismissed.

Results

All three groups were equally hypnotized and entered the appropriate moods equally well. The two word sets were equally learned, so they will be ignored.

Mood-dependent retrieval. Each subject learned one list in a mood that matched his or her recall mood, and another list in a mood that mismatched. For same-mood retrieval, the percentage retention was 76%; for opposite-mood retrieval, it was 70% ($t = .69$, $df = 23$, n.s.). These results were unchanged when we examined only subjects who experienced strong mood inductions. Thur, the percent retention measures again showed no MDR. This absence occured here despite use of the MMFR procedure of Barnes and Underwood (1959) which minimizes list-discrimination problems at recall.

Replication of the word-congruity effect. The Control subjects showed superior learning for same-mood words during original learning (71% vs. 55%, $t = 2.54$, $df = 7$, $p < .02$, by a one-tailed test). Thus, the use of tape-recorded procedures excludes experimenter voice inflections, instructional emphases or varying word orders as possible artifacts in this effect.

Instruction/manipulation check. Performance of subjects in the MC and MI groups demonstrated that they had followed their instructed learning strategy. For original learning scores, the Group by Learning-Mood by Word-Content interaction was significant ($F(2,21) = 4.38$, $p < .03$). The mean difference in immediate recall for mood-congruent versus-incongruent words was 2.63 for the Control Group, 2.00 for the MC Group, and -1.63 for the MI Group, averaged over the two learning lists (one in each mood). The absolute values of these effects (ignoring sign), which reflects the degree of selectivity in learning, did not differ reliably ($F(2,21) = .21$, n.s.). Thus, our instructions produced a learning differential which mimicked the spontaneous mood-congruity effect.

Differences in Performance Among Groups. It had been predicted that because individuals in the MI Group would be concentrating on mood-incongruent material, they would learn less than the other two groups. The same Group by Learning-Mood by Content ANOVA reported above told us that exactly the reverse phenomenon had occurred. A Group by Learning-mood interaction did occur ($F(2,21) = 4.01$, $p < .03$), but it reflected the MI Group's learning more words overall than the other two groups (C: 10.06, MC: 10.38, MI: 11.19). It was also hypothesized that mood would be lower for the MI Group. While the three induced moods were lower for the Mood-Incongruent subjects after they had concentrated on mood-incongruent items ($M = 4.8$), their difference from the other two groups ($M = 5.5$ and 5.8), was too small to reach statistical significance ($F(2,21) = 1.60$, n.s.).

Clustering. Finally, we examined clustering in recall according to pleasant versus unpleasant words. Using a runs measure, we found that subjects were very definitely clustering their recalls by pleasant versus unpleasant words. However, the three groups did not differ in degree of clustering.

Questionnaire. We included a post-experimental check-off questionnaire to find out what people believed was the purpose of Experiment 3. Twelve of 16 subjects in the Control and MC conditions ranked the mood-congruity hypothesis as among their three highest choices. Exactly the same number chose the mood-dependent retrieval hypothesis.

As a third choice, 5 of 17 of the subjects picked, "to test if people learn more under happy or sad moods," and to check whether "going through moods increases learning." With somewhat less consensus, 4 of 8 students in the MI condition picked the mood-contrast alternative, which is the exact opposite of the mood-congruity hypothesis (which, incidentally, their recall confirmed). As their second choice, 4 of 8 again picked "to test if people learn more under happy or sad moods," and as their third choice, 3 of 8 chose the mood-dependent retrieval hypothesis among their top three hypotheses. Taking these responses at face value, subjects were quite accurate at guessing the hypotheses under test. Regrettably, once again, subjects' belief that MDR was being tested seemed not to bring about that phenomenon.

Discussion

The results of Experiment 3 confirm and extend the results of Experiment 1. First, the mood-dependent retrieval effect failed to appear, despite use of a modified free recall test ("MMFR") that did not require subjects to differentiate list-membership of candidate words for recall. None of the several recall or conditionalized retention measures showed any evidence of MDR. We have reluctantly concluded that the initial demonstration of MDR by Bower et al. (1978) was an unreliable, chance event, possibly due to subtle experimenter demand. Given similar procedures, the MDR effect simply does not replicate.

Second, a robust mood-congruity effect was found in immediate recall. The control group in this experiment essentially replicated our earlier results. The instructional manipulation had strong effects, specifically, instructions to concentrate on mood-incongruent words reversed the usual effect. Curiously, instructions to concentrate on mood-congruent words produced no more biased learning than did the standard instructions given to control subjects. This might be interpreted to mean that the strategy controlled by instructions engages essentially the same encoding mechanisms as occur "naturally"without specific instructions, due to the prevailing mood. These mechanisms may include: greater attention to mood-congruent items, greater or more persistent rehearsal of them, and greater associative elaboration connecting these items together and to the context for recall. The clustering results suggest that inter-relating items of the same affective quality may be a major contributor to the congruity effect.

The surprising outcome was the elevated recall of subjects instructed to concentrate on the mood-incongruent items. Their results were unexpected in two respects. First, their differential learning in favor of the instructed target items was just as great as for the other two groups. We had expected them to show a lesser differential, on the hypothesis that their natural mood would be promoting encoding processes in opposition to the instructions. This clash should have resulted in learning that reflected a compromise between "automatic versus controlled" directions of the encoding mechanisms. But no such cancellation of opposing learning processes occurred. Instead, the second unanticipated outcome occurred, namely, better learning overall by the MI subjects. While their learning of mood-congruent items nearly equaled that of the MC and Control Groups, it was greater learning of mood-incongruent items that led to their better overall scores. One interpretation is that the MI subjects' mood naturally ("automatically") engaged the mood-congruent items for more elaborated learning, whereas their instructions motivated their learning when incongruent items were presented. In contrast, MC and Control subjects were not motivated to learn the mood- incongruent items, so those suffered the poorest recall.

A second interpretation of the results is that the basic mood-congruity effect (in the control condition) is due to subjects following an implicit experimenter demand. This seems unlikely for several reasons. First, mood-congruous recall has been observed in circumstances which would seem to preclude implicit demand (e.g., Laird, Wagener, Halal, & Szegda, 1982). Second, on the questionnaire, our subjects indicated that they believed several hypotheses (MDR and a main effect of mood on learning) which were not evidenced in their performance. Thus, not even our subjects' beliefs in these effects were able to produce them, whereas mood-congruity occurred reliably.

Our results suggest that the "normal" mood-congruity effect results from the greater allocation of attention and learning processes (rehearsal, elaboration) to mood-congruent items to the detriment of incongruent items. This differential encoding hypothesis implies that a mood-congruity effect will appear only during learning of a mixed list of pleasant and unpleasant materials. Only in mixed lists can the subject allocate encoding resources differentially to mood congruent versus incongruent materials. No such learning bias should appear when different subjects learn unmixed lists of uniformly positive or negative materials. In line with this expectation, McDowell (1984) found that depressed subjects learned more depressing than pleasant words when these were presented in a mixed list but not when the pleasant or depressing words occurred in pure lists of consistent affective value. McDowell explained his results by the "encoding resources" idea noted above. Our results with instructional variations support the same hypothesis about the mood-congruity effect.

To summarize, the general conclusions from Experiment 3 are that MDR failed to appear again and that learning strategies under instructional control can facilitate, hinder or even reverse, the tendency of subjects to learn mood-congruent items.

EXPERIMENT 4

We became convinced that our two-list interference design for the mood-context experiment would not produce a reliable MDR by itself. We then asked whether there were some alterations in procedure that might increase the likelihood of obtaining MDR. Experiment 4 tests one promising hypothesis regarding a critical condition to produce MDR.

Examining closely the learning procedures of the earlier experiments, we may note that the mood was induced before the word list to be learned was presented. The subject was told to experience the mood as a continuing background while the items were being presented and learned. Importantly, the subjects were not to directly relate the items to their mood, but rather to view the items as unrelated events that occurred coincidentally (almost accidentally) while they were in a certain feeling state. Interestingly, in Brett Thompson's experiment that had produced evidence for MDR (cited in Bower, 1981), a key part of the learning instructions had asked subjects to explicitly relate each word to their mood, to think of each word as deepening or intensifying their mood.

Following up on this hint, we hypothesized that MDR might require the subjects to causally relate their aroused emotional response to the to-be-learned material. If subjects causally attribute their emotional reaction to the material—if they perceive the two as "causally belonging" together—then we believed that they should form a strong association in memory between the stimulus event and the emotion it evoked. According to this hypothesis, contiguity alone, without belongingness, produces either no connections or only weak ones. This hypothesis falls in the tradition of E.L. Thorndike's (1932) ideas on belongingness as a necessary condition for association formation. (For demonstrations of the insufficiency of contiguity alone for memory, see Glenberg and Bradley, 1979). The causal belonging hypothesis makes sense in terms of ecological validity. Organisms should associate their emotional reactions primarily to the specific objects, persons, and situations that evoked those emotions. In this way, they are enabled to react appropriately upon recurrence of those situations.

We wished to design a test for MDR in light of this causal belongingness hypothesis. To do so we had to help subjects create the belief that the intensity of their emotional state was increased as a result of their experiencing a specific learning item. We therefore changed the learning stimuli from word lists to one-to-three sentence descriptions of happy or sad episodes, asked the subjects to imagine themselves experiencing each episode and to imagine it causing an appropriate emotion in them. Each episode was presented for one minute in order to allow subjects to fully develop the episode imagery and experience the emotion it would cause. In order to make the causal attribution at all plausible, we had to have subjects pair happy or sad episodes only with the socially appropriate emotion. This design confounds learning mood with type of material (and precludes examination of a mood-congruity effect), but seems required by the hypothesis to create a plausible rationale for the experimental subjects.

Finally, to plausibly attribute one's emotional reaction to an event, one should be feeling nearly "neutral" before the event, and upon presentation of the event begin to experience the emotion. Ideally, then, we would like to have subjects in a neutral state, present a single episode and have them experience an appropriate emotion; and repeat this emotional cycle 20 or 30 times to build up a sample of memory items to estimate recall reliably. But using such a procedure for each item would create a major logistical problem. Emotional reactions due to imagined events are slow to develop and slow to dissipate. Multiple moods are draining for the subject, and rapid switches between moods create a "whipsaw" effect. We were therefore prevented from using this method.

As a compromise procedure to meet these conflicting constraints, we used a "blocked staircase" method for presenting the episodes. This "blocked staircase" learning procedure was our attempt to help subjects believe that the imagined episodes were causing an increase in the subjects' emotional arousal. The method has the advantage of presenting many items while avoiding the problems of emotional whip-saw and inertia.

One other change was made in the procedure for this experiment. Our earlier demonstrations of MDR have been frequently criticized (in private as well as in print, e.g., Hasher, Rose, Zacks, Sanft, & Doren, 1985) for use of hypnosis as a mood-inducing procedure and for selection of hypnotically susceptible subjects to participate in our experiments. Without conceding the validity of those criticisms, we nonetheless felt that any demonstration of MDR involving hypnosis would not be generally accepted. Therefore, we used a different mood-induction procedure with unselected subjects. Specifically, the moods were induced by having subjects listen to a tape recording of happy or sad music, while imagining themselves in scenes designed to make them feel happy or sad. This induction has been tested successfully by Clark (1983), and it worked well with pilot subjects. We used the music tape to get subjects started on the appropriate mood at the beginning of the learning episodes. They then used the continuing music plus the imagined episodes to move up or down the emotional staircase.

As a control condition, some subjects learned all lists in a neutral mood, then were tested for recall while in a happy or sad mood. This condition should control for the possibility of experimenter-demand for selective recall conveyed by the induced recall mood. The mood-dependent retrieval hypothesis expects no selective recall in this condition since no mood-to-item associations would have been formed during the learning period.

Method

Subjects

Subjects were 32 students fulfilling experimental requirements for the undergraduate course in General Psychology. No prescreening for "mood susceptibility" was conducted. All passed a prescreening test for depression. Conditions 1 and 2 contained 12 subjects and Condition 3, the control, contained eight subjects.

Design

There were five mood periods during the present experiment. Subjects were mood-induced and exposed to the learning material in four blocks. The learning order was always ABAB, where the initial mood was happy in Condition 1, and sad in Condition 2. In Condition 3, the Control condition, subjects performed all learning in a neutral mood (N). Each condition was further divided in half according to whether the final mood during retention testing was happy or sad. In Conditions 1 and 2, happy and sad-mood induction/learning episodes were blocked in eights as described below; in Condition 3, they alternated. In sum, the three conditions run were ABABA, ABABB, and NNNNA, where N denotes a neutral mood and the A and B denote happy and sad moods in Condition 1 and denote the opposite moods in Condition 2. Final happy and sad memory moods alternated in the control condition.

Equipment

A Sony Walkman tape-cassette player with two sets of headphones was used to play music to one or two subjects being run at one time.

Music Selections

Music selections were recorded on tape and played on a Sony Walkman cassette player with headphones. The music consisted of two 10-minute happy selections, two 10-minute sad selections, and two 10-minute neutral selections. The happy selections were *Coppelia* by Delibes, and portions of the *Brandenburg Concerto* #2, by Bach. The sad selections were *Russia under the Mongolian Yoke*, from the opera *Alexander Nevsky* by Prokofiev, and *Nocturne #15 in F minor, Opus 55 No. 1*, by Chopin. The neutral music consisted of two additional selections from the Chopin nocturnes, #17, *Op.* 62, No. 1 and #18, *Op.* 62, No. 2. Any music selection less than 10 minutes was repeated in a tape loop so as to fill up the necessary time.

Mood Induction Episodes and Learning Stimuli

The mood induction and learning stimuli consisted of 16 happy and 16 sad episodes. Each episode was written in the second person and consisted of one to three sentences. These episodes had been rated on a 21-point scale (from -10, Very Sad, up to +10, Very Happy) by four judges. On the basis of its mean rating, each episode was then ordered on its respective happy or sad scale. Within each scale, episodes were divided into two equivalent sets of eight. Episodes were then ordered within each presentation block according to intensity such that the progression from episodes 1 through 8 within each block went from least to most intense. Each episode was printed in large type on 4 x 6 inch index cards. The four blocks of eight episodes, along with their happy or sad music, comprised the four mood induction/learning stimuli. For the control subjects, the same episodes were assigned to the four blocks, with the restriction that equal numbers of happy and sad episodes were interspersed throughout each block of eight episodes. Four happy or four sad photographs were used as the final mood-induction stimuli (these were not to be recalled). Photographs were used in this final induction in the hope that they would not interfere selectively with memory for the learning materials.

Procedure

Subjects in the experimental group were instructed to read each episode and "to try and imagine themselves actually involved in the particular scenes that were depicted." It was repeatedly emphasized that they were to try to feel the happy or sad emotion that would be evoked by each episode as if it were happening to them. They were told that, with each successive episode, they should move one step up the emotion scale, making their mood more pleasant (or unpleasant). The metaphor suggested was "as if going up (or down) a staircase," or "going up (or down) in an elevator." As in the earlier experiments, subjects were told not to experience the mood at a level they would find overwhelming. Between mood blocks, subjects were asked to enter into a neutral mood by remembering a neutral event from their life.

Instructions were identical for the Control subjects except that they were told to maintain a neutral mood throughout the learning portion of the experiment regardless of the mood evoked by the episodes they would be reading. To make their instructions comparable to the others, control subjects received the (admittedly puzzling) request to "take themselves one step more toward a truly neutral mood" with each episode they read. The neutral music tape was played throughout their learning period.

The first music selection appropriate to the experimental condition was then played for two minutes. Subjects were told that, during this time, they were to relax and try to feel the mood suggested by the music. After two minutes, subjects were handed the first block of eight mood induction/learning stimuli and were asked to read the first episode in the packet. They were signaled at one-minute intervals to advance to the next episode in the packet and imagine it; the music played throughout the block. At the end of the eight-episodes, participants rated their mood on a ten-point scale. A one-minute rest period followed, wherein subjects were urged to return their mood to a neutral state, and to rate their mood again. The above procedure was then repeated three more times, taking the subject through the three more staircase blocks of learning trials. During learning, the same happy music was used for both happy moods and the same sad music for both sad moods. After going through the four learning blocks and returning each time to a neutral mood, all subjects were placed into a final happy or sad mood for recall. For this final mood, subjects were asked to get happy or sad by listening to a new piece of happy or sad music for two minutes; they then looked at and thought about the events depicted in either four happy or sad photographs for a further three minutes while the music played.

At the end of this five-minute period, subjects rated their mood. They were then given 15 minutes to free recall as many of the 32 learning episodes as they could. The music of their final induction phase continued to play throughout this recall period, at the end of which subjects rated their mood one last time. Those participants receiving a final happy mood were then debriefed and dismissed; those who had received a final sad mood were put through a brief happy mood induction before being debriefed and dismissed. The entire procedure took slightly under two hours per subject.

Results

Mood manipulations. The mood manipulations involving music and imagined emotional episodes proved quite effective. For the experimental groups, mood ratings on a ten-point intensity scale obtained at the beginning of the session before the induction were at a neutral level (1.67). The four mood blocks produced average mood ratings of 5.56, 6.25, 5.77, and 6.37 in absolute values. Between learning periods, the instructions to return to a neutral mood produced mood intensity ratings of 0.63, 0.92, 0.71, and 0.84, where zero is defined as truly neutral.

The final mood induction for the retention test produced an average rating of 6.37. Thus, the moods induced by this music-plus-imagined-episodes induction were rated as strong as those induced by hypnosis in selected subjects (compare our 6.37 to 6.38 of Experiment 2 and 5.35 of Experiment 3). While the inductions created a moderately strong mood on average, some subjects were not greatly affected by one or more of the five inductions. Later, these subjects will be treated separately in our analyses.

Retention mood effect. The main data on final retention percentages are shown in Table 2 for the three groups. The effect of recall mood across the three groups was small, as indicated by an average retention percentage of 60% for happy recallers versus 57% for sad recallers. These percentages do not differ reliably. Similarly, pleasant episodes were recalled overall slightly but insignificantly more often than unpleasant episodes (60 vs. 57%).

Mood-dependent retrieval. Comparing the recall of items learned in a mood that matches versus a mood that mismatches the recall mood, Table 2 reveals a slight advantage for mood-matching recall in three of the four experimental groups (lines 2, 3, and 4). Individual t-tests on the recall levels for each condition (row of Table 2) finds significance in only one case, that for Line 3 ($t(5) = 2.01$, $p < .10$). If we combine these four groups, we obtain average recall of 64% for mood-matching items versus 56% for mismatching items. This 64 vs. 56% difference is marginally *significant* (*t*(23) = 1.85, *p* < 0.05 by a one-tailed test). However, examining individual recall scores for mood-match and mood-mismatch items, only 13 subjects showed more recall of match items, 9 showed less recall, and 2 showed equal recall. These frequencies do not differ from a .50:.50 split.

The marginal level of MDR in the experimental conditions may be contrasted to the complete absence of any recall effect of mood in the Control conditions (see Table 2). If subjects believed the recall-mood induction was an implicit demand to recall same-mood items, then better recall for mood congruent items should arise for these Control subjects. But in fact Control subjects had identical recall percentages (54%) for episodes whose affective quality matched or mismatched their recall mood. This absence of effect was expected in light of our earlier negative findings when mood was induced only at recall (Bower et al., 1981).

The results in Table 2 may be examined from other perspectives. One perspective notes that subjects who were happy during recall remembered more happy than sad episodes (68 vs. 59%), whereas subjects who were sad at recall remembered slightly more sad than happy episodes (59 vs. 54%). Another perspective notes that happy episodes were recalled better by happy subjects (68%) than by sad subjects (54%), whereas sad episodes were recalled equally well (59%) by both sad and happy subjects. The latter perspective is consistent with a proposal by Isen (1985) that positive emotions produce better recall of pleasant materials, whereas sad moods produce no systematic differences in recall of pleasant vs. unpleasant materials.

Next, we examined those subjects who were most affected by the mood inductions. One could argue that MDR should be expected only in those cases when the subject strongly feels happy and sad at the appropriate times. Failure to feel the instructed mood would attenuate the mood-to-item association during learning or the retrieval cueing power of the mood during recall. We therefore adopted a stringent criterion and selected only subjects who showed mood-intensity ratings of 5 or greater (out of 10) on all five learning and recall moods. The 13 selected subjects (of 24) provide a stronger test of the MDR prediction. Among these subjects, those who were happy at recall remembered 68% of the happy episodes and 46% of the sad episodes. Those who were sad at recall remembered 59% of the sad episodes and 52% of the happy episodes. Pooling these results, these subjects recalled 65% of the mood-matching episodes versus 50% of the mismatching episodes, a significant MDR effect ($t(12) = 2.41$, $p < 0.025$, one tailed). In contrast, subjects who failed to achieve the instructed moods at the requisite levels showed al-

TABLE 2 Average Recall-Score Percentages for the Three Groups in Experiment 4 Divided According to Recall Mood and Whether Items Matched or Mismatched the Recall Mood

Learning Moods	*Recall Mood*	*Matched Moods*	*Mismatch Moods*
HSHS	Happy	.71	.73
	Sad	.58	.51
SHSH	Happy	.65	.44
	Sad	.60	.57
Experimental Average		.64	.56
NNNN	Happy	.50	.52
	Sad	.58	.55
Control Average		.54	.54

most no MDR effect, at 62% versus 63% for the matching versus mismatching conditions, respectively.

Several subsidiary analyses were carried out. One analysis examined the influence on recall of the serial position of the items during study in Conditions 1 and 2. There were four successive blocks of eight episodes presented in the study list. Recall of these four successive blocks averaged 51, 59, 62, and 68%, respectively. These proportions differ significantly, revealing a simple recency effect in retention. The slope of this recency trend was larger for those episode blocks which mismatched the recaller's mood (percentages of 42, 60, 53, and 70%, respectively) than for those blocks which matched the recaller's mood (60, 59, 71, and 66%, respectively).

A second analysis examined recall of episodes depending on the level of emotional intensity that subjects were instructed to feel in association with the episodes. Recall that episodes within a block were presented in order of increasing emotional arousal, as judged by six norming raters. One might suppose that subjects would remember more of those episodes that were associated with more intense arousal. To assess this hypothesis, recall scores were tabulated separately for the eight episodes within each block which corresponded to increasing levels of instructed emotional intensity. To increase stability of estimates, successive episodes within each block were pooled (i.e., 1 & 2, 3 & 4, 5 & 6, 7 & 8). The recall percentages for the episodes experienced at those four levels of presumably increasing intensity were 70, 55, 54, and 61%, respectively. These percentages show a standard "serial position effect" within each block, with no apparent influence of instructed intensity. Moreover, the same serial position effect was observed for those episode blocks which matched versus those which mismatched the recaller's current mood. This outcome is understandable if instructed intensity per se has no influence on recall ability. Distinct breaks before and after each block of eight episodes occurred to permit the subject's mood to return to neutral. Consequently, the initial and final episodes within each block would be isolated and protected from crowded study of items on either side of them. In any event, we must conclude that the intensity manipulation via instructions and rated episodes produced no differences in memory of the episodes corresponding to their rated emotional arousal.

Discussion

Experiment 4 provided the only results in this series which revealed a significant MDR (discounting the one measure in Experiment 1). The new mood-induction procedure was reasonably successful, and the instruction to link the learning items to the mood by the blocked staircase method seemed to work reasonably well for our subjects. Moreover, Control subjects behaved as expected, showing no differential recall as a function of their mood during recall. These Control results make the point that MDR depends upon a match between mood during learning and during retention testing. Along with Experiments 1-3, they raise problems for a simple experimenter-demand explanation of these results. The demand explanation has to argue that subjects must be exposed specifically to the full MDR design in order to get the idea that they should produce MDR results.

The design of Experiment 4 precludes examination of the mood-congruity effect during learning of materials. In fact, to achieve face validity, the affective quality of the to-be-learned materials had to be perfectly confounded with the subject's mood while learning them. The MDR effect observed, however, cannot in any way be considered a "mood-congruity" effect, since in all cases the learning items agreed with the learner's mood at the time. The congruity effect requires comparing learning of materials which are congruent and incongruent with the learner's mood.

The most intriguing result here is that a small mood-dependent retrieval effect appeared to have returned under circumstances that are theoretically interesting and plausible, viz., when causal belongingness between the items and their emotion is promoted. That is, if subjects are induced to causally attribute their emotional reactions to the material to be learned, then a natural association may be forged between the two units in memory. This emotion-to-item association is then revealed when that emotion is reinstated during retention testing. The idea that associations require "belongingness" to be formed is hardly new. Thorndike's (1932) experiments repeatedly demonstrated such effects in a variety of memory tasks.

We note that the causal belongingness hypothesis would account for one general pattern of presence versus absence of MDR findings in different settings. Specifically, MDR fails or is unreliable when the mood is an incidental background feature during learning. But we've noted that MDR consistently occurs when people recall autobiographical events; they selectively retrieve more events that agree with their recall mood (e.g., see Bower, 1981; Snyder and White, 1982). This selectivity may not simply be an effect due to recall mood but rather may be an MDR effect. An experiment by Gilligan and Bower (reported in Bower, 1981), which had subjects keep a diary of daily emotional events, found that happy versus sad people later recalled a greater proportion of the matching-mood events recorded in their diary. Such emotional memories are characterized by a natural causal belongingness between the life-event and the emotional reaction, so an MDR would be expected.

However, before placing too much confidence in the results of Experiment 4, we should note several qualifications. First, in average percentages, MDR was small and barely reliable in a statistical sense. Second, the group-average MDR effect was not uniform across subjects; in fact, only 13 of 24 subjects recalled more mood-matching than mismatching episodes. The statistical significance of MDR at the group level was largely due to the extreme MDR shown by four subjects. Third, the episodes were not recalled at a level corresponding to the intensity of the emotion the subjects were instructed to feel in reaction to the episodes. In fact, a bow-shaped serial-position effect was observed for recall across the successive serial positions within each block. While not required by an MDR effect, a positive effect of intensity on recall would have increased confidence in the validity of our escalating mood manipulation.

EXPERIMENT 5

Because of the disquieting factors mentioned above, we decided to replicate Experiment 4 with a fresh pool of subjects and new experimenters.

Subjects

The subjects were 20 students fulfilling a service requirement for the undergraduate course in General Psychology. No prescreening for mood susceptibility was conducted.

Design

The design was identical to that of Experiment 4 except that the neutral mood (control) conditions were not run. Five subjects were assigned in random alternation to four groups. The four blocks of learning episodes were experienced in the order HSHS or SHSH; half the subjects who received each learning sequence were tested for retention in a happy mood (H), and half in a sad mood (S).

Procedure

The materials and procedures were either identical or closely similar to those of Experiment 4. A few changes were: (1) we re-ordered many of the 32 episodes because a new group of 10 norming subjects ranked these in intensity slightly differently than did raters in the earlier experiment; (2) the instructions were revised to make clearer the rather complex imaginings we asked of subjects. The experiment was described as a study of the effectiveness of imaginative exercises and music in inducing increasingly happy or sad moods. The intensity of the subject's external reaction to the imagined episode was supposed to increase one step as they went through the eight successive episodes within a block. No mention was ever made of learning or memory testing until the surprise recall instructions given after the final mood induction.

The recall protocols were scored with each episode receiving a score of 0, .5, or 1 depending on whether its gist was not recalled, partly recalled, or fully recalled. The scores of the primary judge were compared across the 32 episodes with those of a second judge, and achieved a reliability of .94. Consequently, the scores assigned by the primary coder were used for all analyses.

Results

The data are presented in Table 3 in terms of two scores for each group. One score gives partial credit (of .5) for partly recalled episodes, the other score ("Number of Items") gives full credit for any part of an episode recalled (so it is a higher score). Regardless of which measure is examined, the same dreary conclusion is inescapable: there was no evidence for mood-dependent retrieval. None of the four subgroups of Table 3 show a significant MDR, and the overall average recall percentages for mood-match and mismatch items are virtually identical. Comparing the match and mismatch recall scores for individuals, ten recalled more match items, six more mismatch items, and four recalled equal numbers. The ten-of-sixteen preference for match recall is not a significant discrepancy from chance of 50%, $X^2(1) = 1, p>0.20$).

TABLE 3 Average Recall-Score Percentages and Percent Items Recalled Wholly or Partly for the Four Groups Divided According to Recall Mood and Matching Versus Mismatching Recall Mood

Learning Moods	*Recall Mood*		*Match Items(%)*	*Mismatch Items(%)*
HSHS	Happy	Score	66	63
		Items	75	78
SHSH	Happy	Score	73	73
		Items	78	81
HSHS	Sad	Score	55	69
		Items	68	74
SHSH	Sad	Score	63	55
		Items	70	59
Overall Average		Score	64	65
		Items	73	73

Next, we examined recall for just those subjects whose moods were appreciably shifted by each induction. An induction was judged effective if the subject's mood rating immediately after the induction was 5 or greater on the10-point scale. Moreover, to be selected for this analysis the subject had to have had effective moods during each of the four learning blocks and at recall. Ten subjects passed these stringent criteria, four who recalled happy and six who recalled sad. Examining their recall scores, we found no evidence whatsoever for mood-dependent retrieval. The happy recallers had percentage recall scores of 78% for happy episodes and 80% for sad episodes; the sad recallers scored 64% in recalling sad episodes and 61% for happy episodes. Overall, the recall scores are 69.7% for matching episodes and 69.1% for mismatching episodes.

Since there was so little evidence of MDR, we did not pursue more complex analyses of the data, e.g., recall as related to serial position and/ or instructed intensity levels.

Discussion of Experiment 5

We have a failure to replicate the marginally significant MDR-effect found using the "staircase" procedure of Experiment 4. The replication failure provokes a search for explanations, perhaps in terms of differences in procedures. The differences are: (1) the presentation order of many of the episodes was altered somewhat; (2) at reinduction of the mood for recall, a briefer time (1 vs. 2 minutes) of music listening was given before the first mood-inducing picture was presented to the subjects; (3) new experimenters ran a new set of subjects; and (4) the lengthy instructions were revised to clarify the subjects' imagination tasks. None of these procedural differences appears consequential enough in theory to create major differences in results. Therefore, we are left to conclude that the significant MDR result obtained in Experiment 4 was spurious or at least statistically non-replicable. We would point, again, to the fact that the MDR effect found in Experiment 4 was only marginally significant and indeed depended on the extreme MDR-like data of four subjects (out of 24).

With the failure to replicate MDR with the staircase procedure, the causal belongingness hypothesis can now be viewed as only an interesting conjecture which was not confirmed in this arrangement of imagined causal connections. Perhaps the experiment was not a proper test of the hypothesis since the "imaginative pretense" aspect of the events and emotions may not have been an effective manipulation of the relevant variables. Conceivably, the belongingness hypothesis could be confirmed in experiments which tested memory for actual events that caused actual emotional reactions. That remains a possibility; but to arrange an appropriate experimental test poses logistical difficulties which we have not overcome. We suggest that memory of sports fans for emotional events in sporting events might be a useful naturalistic arena for testing such ideas.

EXPERIMENT 6

In our search for conditions that might produce reliable MDR, a recent procedure of Eich (in press) came to our attention. Eich obtained MDR with a music mood induction but only when memory was measured for verbal items which subjects had generated to cues during learning. The generation technique followed that used by Slamecka and Graf (1978) wherein the to-be-remembered word was elicited by presenting the subject with a category cue and the unique first letter. An example is "Flavor of ice cream: V———" as a cue for the subject to generate VANILLA. While in a happy or sad mood, Eich's subjects generated some items mixed in with others which they merely read (with the same category cues). Their recall 24 hours later in the same or opposite mood showed a strong MDR for the Generated words but none for the Read words. Eich replicated this interaction in a second experiment designed to increase recall levels for the words in the Read condition. He speculated that Generated items might show more MDR because changing the subject's mood during retention testing may alter the ability to reconstruct the internal processes that caused a specific item to be generated during the initial exposure to the items.

Hoping to follow up on this promising lead, we conducted a "systematic replication" of Eich's experiment. That is, we changed several procedural details which a priori and in theory should not be critically important to the MDR result. One change was that our subjects only received Generate items rather than a mixture of Read and Generate items. This seemed appropriate in our search for MDR since Eich had observed MDR only in the Generate condition. A second difference was our use of a within-subjects learning design wherein each subject generated one word-list while happy and one list while sad, whereas Eich had used a between-subjects design. Half his subjects had generated their learning list when happy and half when sad, with half of each of those groups recalling the next day after either a happy or sad mood induction (a 2 x 2 design). We expected that the addition of the two-list interference design could only enhance the effect. A third difference is that we used the mood music induction as did Eich but added the Velten mood-induction cards. Eich had continued the mood induction until a subject attained a criterion mood level before he or she was advanced to the Generation task or the recall task. In contrast, we advanced subjects to the Generation (learning) task regardless of the impact of our fixed mood induction procedure.

The question of interest was whether the MDR result observed by Eich would be replicated with these variations in procedures.

Method

Subjects

The subjects were 20 summer students at Stanford paid $8.00 for their participation. They were solicited by ads posted around campus.

Materials

The learning materials were patterned after those of Eich (in press) and Slamecka and Graf (1978). Sixty instances of the taxonomic categories of Battig and Montague (1969) were selected such that the category plus the first letter designated a unique high-frequency instance. Examples were "Metal-I(ron)", "A relative-U(ncle)", and "A Fruit-A(pple)." These 60 category-plus-first letter cues were arranged into two learning lists of 30, which were seen equally often by happy or sad subjects.

Recognition-memory test sheets were composed from these 60 learning items together with 60 lures chosen as high-frequency instances of the same taxonomic categories. The 120 words were scrambled and printed on two sheets of paper. Beside each word was a scale representing "List 1, List 2, or Neither". The subject was to circle one choice indicating his judgment for each word on the test list.

The mood induction used the music selections of Experiments 4 and 5 (plus some new mood music) together with the Velten self-referent mood statements. The mood-induction statements of the Velten refer to feelings of elation or depression which the subject is asked to experience, e.g., "I'm feeling very tired and listless today," "I'm very sad and depressed today." We chose two sets of 20 sad statements and two sets of 20 happy statements from Velten's master set of 60 sad and 60 happy statements. These sets of 20 statements were presented in order of increasing intensity of the suggested mood. One set of happy statements and one set of sad statements (identical for each subject) were used during the initial generation session. The appropriate alternate set of Velten statements were then used (along with different music) to induce that subject's assigned mood during retention testing.

Design and Procedure

Each subject generated 30 words while happy, 30 while sad, and returned 24 hours later for free recall and recognition memory tests on both lists while feeling happy or sad. Ten subjects recalled when happy and ten when sad. Half of each set of ten subjects had generated words first while sad and then while happy, whereas the other subjects experienced the moods in the reverse order.

Procedure

Subjects were run individually. They were told the experiment concerned the influence of mood upon the ability to generate answers to questions. The nature of the mood manipulations and the generation tasks were then described. Specifically, the subject was asked to get into a happy or sad mood by listening to music and by concentrating on feeling the emotions suggested by the 20 Velten statements typed on index cards. The subject read the statement and imagined experiencing the feelings suggested in the statement, proceeding to the next card every 45 seconds (signaled by a click). At the end of the 20 Velten statements, the music continued while the subject rated his or her mood on a -10 (very sad) to +10 (very happy) scale, and proceeded to complete the first list of 30 Generate questions. These were presented three to a page, and subjects wrote in their answers. A second mood rating was taken. Then the music was stopped. Subjects were asked to get themselves into a neutral mood for one minute and rate their mood again. Then they were asked to get themselves into the opposite mood, by listening to the

TABLE 4 Recall Percentages of Each Learning List According to the Mood of the Recall Subject

	Recall Mood	
	Happy	*Sad*
List learned while		
Happy	27.4	24.6
Sad	26.7	21.2

opposite-mood music and by reading the Velten statements for the opposite mood. After this second mood induction, they rated their mood again, then proceeded to generate answers to the second list of 30 category-plus-first letter cues. After this, another mood rating was taken. Then happy subjects were dismissed. Sad subjects were instructed to neutralize their mood for three minutes to remove their sadness before they were dismissed. All were scheduled to return the next day for "further tests of our mood-induction procedure."

Upon return 24 hours later, subjects were assigned to receive a happy or sad mood induction. A new music selection was played and the other, unused set of Velten statements (appropriate to the emotion) were given to induce the mood. Subjects rated their mood, then received instructions for free recall of both lists of target words in any order that they came to mind. They had as long as they wanted to write their recall, but were forced to try to recall for at least ten minutes. The mood music played throughout the entire test period. A new mood rating was taken, and subjects proceeded to the test for recognition memory. This test consisted of 60 old target words and 60 new words in scrambled order. For each word, subjects circled their choice as to whether it had appeared the day before in List 1, in List 2, or in neither. They performed this test at their own pace. After completing this test, a final mood rating was taken. The sad subjects were given some pleasant fantasies for five minutes to elevate their mood; then all subjects were debriefed and dismissed.

Results

The primary results are the recall percentages of the two groups of subjects; these are shown in Table 4.

With only one incidental trial and a 24-hour delay, subjects still were able to recall 25% of the target words, which level is sufficient to reveal any MDR should one be present. Regrettably, the recall percentages in Table 4 provide no evidence of mood-dependent retrieval. None of the four percentages differ significantly from one another. Moreover, the poorest recall is in the sad-sad condition, which should have produced the better recall. Pooling across conditions, recall of items generated while in a matching mood was 24.3%, whereas recall of items generated in a mismatching mood was 25.6%. These obviously do not differ reliably. Nine of the subjects recalled more of the mood-matching targets, ten subjects recalled more of the mood-mismatch targets, and one recalled the two lists equally well. The results on MDR were no better considering only subjects who experienced intense moods; similarly, none of the measures of recognition memory revealed any MDR or group differences.

Discussion of Experiment 6

We have failed to replicate the promising results of Eich. As noted earlier, there are several differences between the procedures in the two experiments. The most conspicuous differences are that (1) our subjects learned two lists, one in each mood, whereas Eich's subjects learned but a single list in one mood, and (2) our mood induction procedure was different and less extensive than that used by Eich. Future research will have to track down the reason for the different outcomes in these two experiments. For now, we will merely note that the MDR effect produced with self-generated material was not robust enough to survive a systematic replication which altered slight details of the procedure.

GENERAL DISCUSSION

In six very disappointing studies we were unable to find stable evidence for the MDR effect. We tried to give MDR every fair chance to appear. Each of several experimental enhancements tested some suggestion about how to strengthen the cue or retrieval power of moods. For example, in Experiments 1, 2, and 3, we increased interlist interference in order to decrease the discriminative power of extraneous contexts such as the experimental setting, thus to make mood a more salient retrieval cue. In Experiments 4 and 5 we introduced attributions of causal belongingness in order to forge stronger bonds between the to-be-remembered, imagined episodes and the emotion they were causing. In Experiment 6, the items to be remembered were subject generated in the hope that memory for self-generated material would be more closely associated to subjects' moods. With the exception of the flicker of hope over the results of Experiment 4, all our manipulations were to no avail in producing an MDR effect.

In contrast to the unreliable MDR effect, a robust mood-congruity effect was obtained in those experiments (1 and 3) where it could be observed. Experiment 3 found that explicit instructions to learn selected materials could oppose the "natural" tendency to learn more about mood-congruent material. Interestingly, subjects instructed to concentrate on learning incongruent items learned more items overall, as though their mood facilitated learning of mood-congruent items while their instructions facilitated learning of mood-incongruent items. The robust mood-congruity effect found here with laboratory-induced moods has not been found by Hasher, Rose, Zacks, Sanft, and Doren (1985) comparing college students who scored moderately high or low on the Beck Depression Inventory (BDI). The lack of mood-congruent learning in the Hasher et al. experiments may result from an insufficient difference in the active emotional state of depression for their high versus low scorers on the BDI. Alternative interpretations of that result in relation to successful demonstrations of mood-congruent learning were stated in papers by Hasher et al. (1985), Hasher, Zacks, Rose, Sanft, and Doren (1985), and by Mayer and Bower (1985).

Returning to our search for mood-dependent retrieval, over the course of our experiments we have undergone several transitions in our thinking about MDR and context effects in general. Long ago, we started with the common assumption that context effects would be strong, omnipresent, and ubiquitous, and that mood could serve as such a memory aid. In fact, the first author was occasionally chided by research friends for even bothering to demonstrate such an "obvious" triviality as that one's emotional state could serve as such a context for learning. All the major learning theories (e.g., the venerable drive-stimulus or emotion-stimulus theories of Estes, 1958, Hull, 1943, and Miller, 1950) expected the result. Within that setting, our original failure to find MDR in a one-list design (see Experiments 1 and 2 in Bower et al. 1978) caused some consternation. But that was soon quelled when we observed MDR with a two-list interference design wherein mood became a discriminative stimulus for the target items to be recalled. And within that experimental design, which after all simulates everyday experience of multiple learning episodes in multiple moods, we believed that the mood-to-item associations arose automatically by contiguity alone and occurred largely without voluntary intention or control by the subject.

Our present findings do not support this interpretation. Four experiments (1, 2, 3, and 6) failed to find any MDR when mood was a prevailing background context that was unrelated to the items being learned. We are convinced that this condition produces either no MDR or, at best, weak unreliable evidence for MDR. In particular, we now believe that the positive results in the earlier report of MDR (Bower et al., 1978, Study 3) were either an unreplicable, chance outcome or possibly due to the experimenter demand produced by an expert hypnotist who had close rapport with his subjects. Thus, we are no longer surprised by other failures to find MDR (e.g., Wetzler, 1985).

As noted, the failure to find a mood-context effect in these "standard context" experiments impacts negatively upon many theories which expect it. The failure contradicts not only the first author's earlier theory of mood as a retrieval cue (Bower, 1981). The failure impacts more generally upon any learning theory which supposes that internal states act as contexts which by their presence can become associated automatically by contiguity to memories of coincident events, thus to later cue their retrieval. The disconfirmed theories include not only the drive-stimulus theories noted above but also the arousal-as-cue theory of Clark, Milberg, and Ross (1983). Moreover, to the extent that mood influences the encoding of verbal material, the failure of MDR on measures of recall and recognition impacts negatively upon the principle of encoding specificity (Tulving & Thomson, 1973). Clearly many theorists have been wrong in expecting or explaining MDR. Some consolation may be found in the fact that the unreliability of MDR may only be an illustration of the unreliability of the "experimental-room" context effect in free recall (Fernandez & Glenberg, 1985).

A consequence of our negative conclusion is that we now have no way to interpret reports of positive MDR effects. What is needed now is some insightful analysis of the reasons, if any, for the positive and negative results on MDR. The field clearly needs to move beyond the current reporting of simple demonstrations of positive or negative findings. We confess an inability to make sense out of the crazy patchquilt of findings on MDR. We are frankly puzzled by the inconsistencies in outcomes.

Experimental psychology rests upon a foundational belief in causal regularities, the idea that by repeating the "essentials" of an experiment, the same outcomes will result. But a psychological experiment, particularly one on induced moods, is a complex, elaborate social drama or "playlet" that the experimenter arranges to act out for the benefit of the subject and with his collaboration (Rosenthal, 1966; Goffman, 1959). In terms of speech acts, unconscious intentions, vocal inflections, subtle messages, bodily communications, framing of the task, and staging, the playlet has probably hundreds of variables or choice points at which something different could have been done or said to the subject.

Most research fields develop a certain mode of describing and analyzing their experimental paradigms, and so it is too in psychology. At the customary standard level of psychological description, we have hit an impasse in our experiments on MDR. We have searched in vain for some variable that we consider "significant" which would distinguish among experimental arrangements that produced positive versus negative results on MDR. One begins with the hope that the outcomes do not depend on "insignificant" irrelevancies such as the experimenter's appearance, tone of voice, exact room temperature, and so on. However, the inconsistent results in this area have caused us considerable despair. Since we cannot even identify the critical variables, it seems that we are far from understanding theoretically what is going on in such experiments. We can only hope that future research will clarify what now appears to us as a very muddled scene.

MUSICAL REFERENCES

Delibes, L. (1870). *"Coppelia" Ballet Suite.* von Karajan, H. (Conductor). Berliner Philharmoniker Deutsche Grammophon Records.

Bach, J.S. (1718-1721). Brandenburg Concerto No. 1 in F. *James Levine Plays and Conducts Bach.* Produced by Jay David Saks. RCA Records: New York, NY.

Prokofiev, S. (1939). *Alexander Nevsky, Op. 78,* Schippers, T. (Conductor). New York Philharmonic Y31014 Columbia Records.

Chopin, (1843). *Opus 55.* Rubenstein, A. (Pianist). The Young Rubenstein, Album 3, recorded 1936-37, IB-6133, Capitol Records, Inc.

Chopin, (1846). *Opus 62.* Rubenstein, A. (Pianist). The Young Rubenstein, Album 3, recorded 1936-37, IB-6133, Capitol Records, Inc.

REFERENCES

Barnes, J.M., & Underwood, B.J. (1959). "Fate" of first-list association in transfer theory. *Journal of Experimental Psychology, 58,* 97-105.

Bartlett, J.C., Burleson, G., & Santrock, J.W. (1982). Emotional mood and memory in children. *Child Development, 50,* 513-518.

Bartlett, J.C., & Santrock, J.W. (1979). Affect-dependent episodic memory in young children. *Child Developement, 50,* 513-518.

Battig, W.F., & Montague, W.E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph, 803,* 1-46.

Bower, G.H. (1981). Mood and memory. *The American Psychologist, 36,* 129-148.

Bower, G.H., & Cohen, P.R. (1982). Emotional influences in memory and thinking: Data and theory. In S. Fiske, & M. Clark (Eds.), *Affect and social cognition.* Hillsdale, NJ: Erlbaum Associates.

Bower, G.H., Gilligan, S.G., & Monteiro, K.P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General, 110,*451-473.

Bower, G.H., & Mayer, J.D. (1985). Failure to replicate mood dependent retrieval. *Bulletin of the Psychonomic Society, 18,* 39-42.

Bower, G.H., Monteiro, K.P., & Gilligan, S.G. (1978). Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior, 17,* 573-578.

Clark, D.M. (1983). On the induction of depressed mood in the laboratory: Evaluation and comparison of the Velten and musical procedures. *Advances in Behavior Research and Therapy, 5,* 27-49.

Clark, M.S., Milberg, S., & Ross, J. (1983). Arousal cues material stored in memory with a similar level of arousal: Implications for understanding the effects of mood on memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 633-649.

Clark, D.M., & Teasdale, J.D. (1985). Constraints on the effect of mood on memory. *Journal of Personality and Social Psychology, 48,* 1595-1608.

Eich, E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory and Cognition, 8*, 157-173.

Eich, E. (in press). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Ellis, E.A. (1983). *Emotional mood as a context for state-dependent retention: Some limitations of the phenomenon.* Unpublished Senior Honors Thesis, University of Toronto, Canada.

Ellis, H.C., Thomas, L., & Rodriguez, I.A. (1984). Emotional mood-states and memory: elaborative coding, semantic processing, and cognitive effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 470-482.

Estes, W.K. (1958). Stimulus-response theory of drive. In M.R. Jones (Ed.), *Nebraska Symposium on Motivation.* Lincoln, NE: University of Nebraska Press.

Fernandez, A., & Glenberg, A.M. (1985). Changing environmental context does not reliably affect memory. *Memory and Cognition, 13*, 333-345.

Glenberg, A.M., & Bradley, M.M. (1979). Mental contiguity. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 88-97.

Godden, D.R., & Baddeley, A.D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology, 66*, 325-331.

Goerss, J.C., & Miller, M.E. (1982). *Memory and mood: State-dependent retention and induced affect.* Meetings of the Midwestern Psychological Association, Chicago.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, NY: Doubleday.

Hasher, L., Rose, K.C., Zacks, R.T., Sanft, S., & Doren, B. (1985). Mood, recall, & selectivity effects in normal college students. *Journal of Experimental Psychology: General, 114*, 106-120.

Holoway, F.A. (1978). State-dependent retrieval based on time of day. In B.T. Ho, D.W. Richards, III, & D.L. Chute (Eds.), *Drug discrimination and state-dependent learning.* New York: Academic Press.

Hull, C.L. (1943). *Principles of behavior.* New York: Appleton-Century-Crofts.

Isen, A.M. (1985). Asymmetry of happiness and sadness in effects on memory in normal college students. *Journal of Experimental Psychology: General, 114*, 388-391.

Isen, A.M., Shalker, T.E., Clark, M., & Karp, L. (1978). Positive affect, accessibility of material in memory and behavior: A cognitive loop? *Journal of Personality and Social Psychology, 36*, 1-12.

Laird, J.D., Wagener, J.J., Halal, M., & Szegda, M. (1982). Remembering what you feel: The effects of emotion on memory. *Journal of Personality and Social Psychology, 42,* 646-657.

Leight, K.A., & Ellis, H.C. (1981). Emotional mood states, strategies, and state-dependency in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 251-266.

Macht, M.L., Spear, N.E., & Levis, D.J. (1977). State-dependent retention in humans induced by alterations in affective state. *Bulletin of the Psychonomic Society, 10,* 415-418.

Mayer, J.D., & Bower, G.H. (1985). Naturally occurring mood and learning: Comment on Hasher, Rose, Zacks, Sanft and Doren. *Journal of Experimental Psychology: General, 14,* 396-403

Mayer, J., Mauro, R., & Trebino, L. (1987). A preliminary report on the safty of hypnotic mood inductions. Unpublidhed manuscript.

McDowell, J. (1984). Recall of pleasant and unpleasant words in depreesed subjects. *Journal of Abnormal Psychology, 93,* 401-407.

Mecklenbrauker, S., & Hager, W. (1984). The effects of mood on memory: experimental tests of a mood state dependent retrieval hypothesis and of a mood congruity hypotheses. *Psychological Research, 46,* 355-376.

Metzger, R.L., Boschee, P.F., Haugen, T., & Schnobrich, B.L. (1979). The classroom as learning context: Changing rooms affects performance. *Journal of Educational Psychology, 71,* 440-442.

Miller, N.E. (1950). Learnable drives and rewards. In S.S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley.

Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values of 925 nouns. *Journal of Experimental Psychology, Monograph Supplement, 761,* Part 2.

Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.

Schare, M.L., Lisman, S.A., & Spear, N.E. (1984). The effects of mood variation on state-dependent retention. *Cognitive therapy and research, 8,* 387-408.

Schor, R.E., & Orne, E.C. (1962). *Harvard Group Scale of Hypnotic Susceptibility, Form A,* Palo Alto, CA: Consulting Psychlogists Press, Inc.

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

Smith, S.M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 460-471.

Smith, S.M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory and Cognition, 10,* 405-412.

Smith, S.M., Glenberg, A., & Bjork, R.A. (1978). Environmental context and human memory. *Memory and Cognition, 6,* 342-353.

Snyder, M., & White, P. (1982). Moods and memories: Elation, depression and the remembering of the events of one's life. *Journal of Personality, 50,* 149-167.

Thorndike, E.L. (1932). *The fundamentals of learning.* New York: Teacher's College.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352-373.

Weingartner, H., Miller, H., & Murphy, D.L. (1977). Mood-dependent retrieval of verbal associations, *Journal of Abnormal Psychology, 86,* 276-284.

Weingartner, H., & Silberman, E. (1982). Models of cognitive impairment: Cognitive changes in depression. *Psychopharmacology Bulletin, 86,* 276-284.

Weitzenhoffer, A.M., & Hilgard, E.R. (1962). *Stanford Hypnotic Susceptibility Scale, Form C,* Stanford, CA: Stanford University Press.

Wetzler, S. (1985). Mood-state-dependent retrieval: a failure to replicate. *Psychological Reports, 56,* 759-765.