Journal of Experimental Psychology:
Learning, Memory, and Cognition
1989, Vol. 15, No. 3, 443–455

0278-7393/89/$00.75

Mood Dependent Memory for Internal Versus External Events

Eric Eich
University of British Columbia
Vancouver, British Columbia

Janet Metcalfe
University of California, San Diego

Events that originate through internal mental operations such as reasoning, imagination, and thought may be more colored by or connected to one's current mood than are those that emanate from external sources. If so, then a shift in mood state, between the occasions of event encoding and event retrieval, should have a greater adverse impact on one's memory for internal than for external events. To investigate this inference, a series of studies was conducted that relied on a continuous music technique to modify mood, and on the generate/read procedures devised by Slamecka and Graf (1978) to distinguish internal from external events. Considered collectively, the results suggest that internal events are less likely than external events to be recalled after a shift in mood state. Discussion centers on both the empirical limitations and theoretical implications of the present results, as well as on prospects for future research.

This article addresses the state dependent effects of moods on memory for internal as opposed to external events. To paraphrase Johnson and Raye (1981), internal events are those that originate through mental operations such as reasoning, imagination, and thought, whereas external events refer to sensory stimuli that are apprehended, or brought into awareness, via the processes of perception. Though the distinction is neither rigid nor precise—thought tends to reflect perception, and perception, thought—differences between memories derived principally from internal versus external sources do exist. For instance, several studies have shown that the probability of recalling or recognizing a word is greater if the word had been generated rather than read by subjects (e.g., Slamecka & Fevreiski, 1983; Slamecka & Graf, 1978). Also, when asked to estimate the frequency with which they had either actively generated or been passively presented with a word, subjects appear more attuned to the relative number of generations than to the relative number of presentations (Johnson, Taylor, & Raye, 1977; Raye, Johnson, & Taylor, 1980). And although the prior presentation of a word has a weak effect on the later estimation of how often that word had been generated, prior generations strongly bias subsequent judgments about presentations (Johnson et al., 1977; Raye et al., 1980). These results suggest that, as Raye and her associates (1980) have remarked, there is something special about memory for internally generated events.

There may be still another respect in which memory for internal events is special. Recent research in the area of affect and cognition indicates that experimental manipulations of mood exert powerful effects on the performance of some, but not all, types of tasks. Among the tasks that seem most sensitive to mood manipulations are word association, narrative construction, and interpersonal assessment (see Blaney, 1986; Bower, 1981; Isen, 1984; Teasdale, 1983). Thus, for example, angry people produce angry associations, tell hostile stories, and tend to find fault in others. Among the least sensitive tasks are Stroop interference, speech shadowing, and perceptual identification (see Blaney, 1986; Bower, 1985). By way of illustration, Gerrig and Bower (1982) failed to find a reliable effect of hypnotically induced elation or anger on tachistoscopic thresholds for naming pleasant versus unpleasant words.

One plausible interpretation of this pattern of results is that tasks that are not particularly sensitive to experimental modifications of mood are those that place a premium on the automatic or data-driven perception of *external* events. In contrast, tasks that are especially sensitive are ones that involve the *internal* production of what Johnson and Raye (1981, p. 70) call "cotemporal thought ... the sort of elaborative and associative processes that augment, bridge, or embellish ongoing perceptual experience but that are not necessarily part of the veridical representation of perceptual experience." If this interpretation is correct, then there should be implications for memory. Specifically, a shift from one mood state to another—say, from happiness to sadness—should lead to a larger loss of memory for events that had been internally produced rather than externally presented.

What evidence is there to support the idea that internal events are particularly apt to be forgotten following a shift in mood state? At present, the answer is—not much. About all that one can point to are a few, scattered reports of mood dependent memory in situations involving the recall of subject-produced associations to words—a type of internally generated event (see Eich, 1980; Weingartner, 1978). Weingartner, Miller, and Murphy (1977), for example, found that affectively disturbed patients who had generated verbal associations during an episode of mania recalled 97% more of their associations when tested in a manic than in a normal mood state. Similar results were obtained by Goodwin, Powell, Bremer, Hoine, and Stern (1969) in a study involving alcohol—a drug that produces reliable, sometimes radical, alterations of affect (Persson, Sjoberg, & Svensson, 1980). In

Preparation of this article was aided by Grants U0298 and A0505 from the Natural Sciences and Engineering Research Council of Canada to Eric Eich and Janet Metcalfe, respectively, and by the advice and assistance provided by Jennifer Campbell, Jackie DiGeso, Heather McEachern, Jack Rachman, Jim Russell, Peter Suedfeld, and Ainslie Winter.

Correspondence concerning this article should be addressed to Eric Eich, Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y7.

their study, intoxicated medical students made 83% fewer errors than did sober subjects when tested for the recall of associations that had been generated under the influence of alcohol.

Interestingly, Goodwin and his colleagues detected a significant, but much less striking, state dependent effect of alcohol in the performance of both a visual-discrimination and a sentence-recall task, and they saw no sign of state dependence in recognition memory for pictures. The fact that all three of these tasks involved memory for experimenter-presented rather than subject-produced items prompted Goodwin et al. (1969, p. 1359) to suggest that "the word-association task, measuring single-trial, 'self-generated' learning, may be particularly useful in studying dissociation [a/k/a state dependence]." Though this suggestion makes sense, it is open to two serious objections.

First, the association task administered by Goodwin and his associates differed from their other tasks not only with respect to the source of the target items (i.e., internal vs. external), but also in terms of the nature of the items themselves (e.g., words vs. pictures) and the method by which memory was tested (e.g., recall vs. recognition). The presence of these confounds makes Goodwin et al.'s data moot on the issue of whether internal events are especially susceptible to state dependence.

Second, it is possible that what Goodwin and his collaborators demonstrated was not a state dependent effect of alcohol on word association recall, but rather a state-congruent effect. The distinction here parallels that between *mood dependence*, which implies that events encoded in a particular mood are most retrievable in that mood, irrespective of the events' affective valence or content, and *mood congruence*, which involves the enhanced encoding or retrieval of events whose affective content is congruent with one's current mood (see Blaney, 1986; Bower, 1981). Though mood dependence and mood congruence represent theoretically distinct phenomena, several studies have yielded results that can be viewed as support for either. One such study is that of Weingartner et al. (1977), and another may be that of Goodwin et al. (1969). Regarding the former study, Blaney (1986) has remarked that

> Weingartner et al.'s (1977) results—indicating that subjects experiencing strong mood shifts were better able to regenerate associations first generated in same as opposed to different mood—could be seen as reflecting either mood congruence or [mood] state dependence. That is, the enhanced ability of subjects to recall what they had generated when last in a given mood was (a) because what was congruent with the mood at first exposure was still congruent with it at subsequent exposure, or (b) because return to that mood helped remind subjects of the material they were thinking of when last in that mood, irrespective of content. (p. 237)

The same ambiguity arises in connection with Goodwin et al.'s (1969) results, inasmuch as the associations that people produce while inebriated may be qualitatively different from those elicited while sober.

To overcome both of the objections raised above, and thereby determine more directly whether mood dependent effects are indeed more pronounced for internal than for external events, the present research relies on procedures similar to those developed by Slamecka and his associates (e.g., Slamecka & Fevreiski, 1983; Slamecka & Graf, 1978). In our initial study, subjects either read a target item, such as GOLD, that is paired with a category name and a related exemplar (e.g., *precious metals*: *silver*–GOLD), or they generate (with a very high probability) the same item when primed with its initial letter, in combination with the category name and exemplar cues (e.g., *precious metals*: *silver–G*). In this manner, memory for one and the same target item can be assessed in relation to its source: either internal (the generate condition) or external (the read condition).

During the encoding session of the first study, subjects generate 16 target items and read 16 others while they are either happy or sad—moods engendered by means of a "continuous music technique," explained below. During the retrieval session, held 2 days later, subjects receive a surprise test of free recall of the 32 target items. The subjects' mood at retrieval either does or does not match the mood they experienced during encoding.

If memory is mood dependent, then a significant advantage of matched over mismatched mood conditions (e.g., encode happy/retrieve happy vs. encode happy/retrieve sad) should be evident in performance of the recall test. Whether this advantage is greater for items that had been generated, rather than read, is the first and foremost issue of interest in this research.

Following the test of item recall, subjects are presented with a list of all 32 previously encoded targets, interspersed with 32 new items. The subjects are asked to respond "old" to any item they recognize as having generated *or* read during the encoding session and to respond "new" otherwise. Given that mood dependent effects rarely emerge when the retrieval of *external* events is prompted with explicit reminders—such as the "copy cues" that are available in a test of recognition memory (see Bower, 1981; Eich, 1980)—one would not expect to find a reliable advantage of matched over mismatched mood conditions in the probability with which read items are correctly classified as old. Whether such an advantage accrues to the prompted retrieval of *internal* events—or, more specifically, to the recognition of generated items—is the second issue at stake in this study.

Following the test of item recognition, subjects are again presented with all of the items they had identified as old and are asked to specify whether they had generated or read these items during the encoding session. The intent here is to investigate a third issue which, cast as a question, is: Even if a shift in mood state does not impair the subjects' ability to recognize the *occurrence* of a particular past event, does such a shift adversely affect their recognition of the source or *origin* of that event?

The fourth and final issue of current interest concerns the role of arousal in mood dependent memory. Drawing on their own and other's data, Clark, Milberg, and Ross (1983) have argued that (a) arousal can act as a cue for the retrieval of arousal-related material from memory (so that events experienced shortly after, say, a session of strenuous exercise are better remembered in a physically active than in a physically relaxed state), and (b) in the main, positive moods, such as happiness, are accompanied by higher levels of arousal than are negative moods, such as sadness. Thus, according to Clark and her colleagues, memory impairments incurred in the transition from happiness to sadness (or vice versa) may be the result of a change in arousal, rather than a change in mood. To check this possibility, we ask subjects to rate their current levels of both mood and arousal on several occasions over the course of the encoding and retrieval sessions. These ratings are then used to separate the subjects into groups; for instance, subjects who show a substantial shift between sessions in mood as well as arousal, as opposed to subjects who experience a large change in mood but only a small change in arousal. By determining the degree to which these groups differ with respect to item recall, source recognition, and other memory measures, we endeavor to evaluate Clark et al.'s (1983) claim that mood dependent effects may be mediated by alterations in level of arousal.

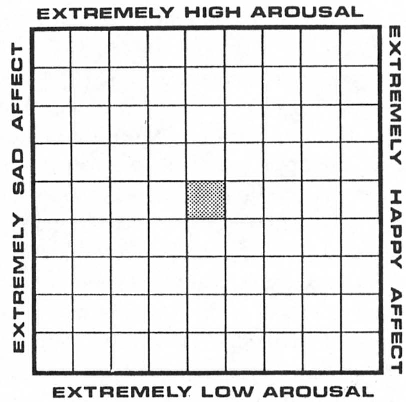

Figure 1. The matrix on which subjects placed a single mark to indicate their current levels of mood (horizontal axis) and arousal (vertical axis). (Adapted from "The Affect Grid: A New Single-Item Scale of Pleasure and Arousal" by J. A. Russell, W. Weiss, and G. A. Mendelsohn, in press, *Journal of Personality and Social Psychology*. Copyright by the American Psychological Association. Adapted by permission.)

Experiment 1

Method

Subjects and design. The experiment employed 48 University of British Columbia (UBC) undergraduates as subjects (each of whom received course credit in return for their participation) and entailed a 2 × 2 × 2 mixed design. One between-subjects factor was the mood (happy or sad) in which target items were encoded, and a second was the mood (again, happy or sad) in which retrieval of the items was tested. The crossing of these two factors defined four encoding/retrieval conditions, each of which was represented by 12 randomly assigned subjects. The third factor, item type, was varied within subjects.

Participants were tested individually throughout the course of the experiment, which was divided into two sessions. We will refer to the first of these as the *encoding session* and to the second as the *retrieval session*; the intersession interval was 2 days. Details on the materials and procedures used in each session are summarized in the following sections.

Encoding session. At the start of this session, subjects were told that the present experiment was part of a research program aimed at understanding how moods affect the performance of word association, speech production, and other cognitive tasks, and how the performance of such tasks in turn affects mood. It was explained that the experiment would be divided into two sessions, spaced 2 days apart, and that each session would entail a different set of tasks. The subjects were then provided with a copy of the matrix drawn in Figure 1—an adaptation of the "affect grid" designed by Russell, Weiss, and Mendelsohn (in press)—together with these instructions:

> We will use the matrix to measure your feelings at a particular moment. Two types of feelings are of interest: one is your level of affect or mood—that is, how happy or sad you feel—and the other is your level of arousal. The center of the matrix represents neutral feelings: you are neither happy nor sad, and you are neither aroused nor unaroused. As you move from the center column to the right, your mood changes for the better—from feeling slightly happy to moderately happy to very happy and finally to extremely happy—and as you move from the center column to the left, your mood changes for the worse—from feeling slightly sad to moderately sad to very sad to extremely sad. In a similar manner, as you move upwards from the center row, your level of arousal becomes progressively higher, and as you move downwards from the center row, your level of arousal becomes progressively lower. Thus, for example, if you now feel moderately sad and moderately aroused, you should mark the square located three rows down from the third column on the left. Alternatively, if you are experiencing an average level of arousal, and feel slightly happy, you should place a check mark just to the right of the center square. With these examples in mind, please mark the one square that best reflects your current feelings."

After making their mark, the subjects were informed of the following:

> In today's session you will listen to a selection of classical music that should help you develop a happy (sad) mood. However, music alone cannot create the desired affect, so you should try to think of something that makes you happy (sad). You may find it especially helpful to concentrate on happy (sad) events that you have personally experienced.
>
> When I think that you have developed an appropriate mood, we will begin today's main task, which will go like this. I will read to you the name of a category, such as *ice cream flavors*, and an item that fits that category, such as *chocolate*. I will then ask you to generate another item that belongs to the same category and that starts with a certain letter of the alphabet. For instance, if I said *ice cream flavors*: *chocolate–V*, you might think of VANILLA. Whatever it is that you do think of, you should say it out loud, so that I can keep track of the items you generate.
>
> While I will sometimes ask you to generate an item that begins with a particular letter, as in the *V–VANILLA* case just mentioned, I will occasionally read to you the complete second item, which you should then read back to me aloud. For example, if I said *natural earth formations*: *river–VALLEY*, you should respond by saying VALLEY out loud. These read-only items are included so that you will remain alert throughout the course of the task, and so that you will not know in advance whether you must generate or read the second item aloud."

On receiving these instructions, the subjects were seated in a specially designed lounge chair (a Sensory Environment Engineers Alpha Chamber) that contained recessed stereo speakers. Through these speakers was played, at a comfortable listening volume, one of two selections of "happy" music (a segment of *Eine Kleine Nachtmusik* [5 min: 10 s] or *Divertimento #136* [4:10], both by Mozart), or one of two selections of "sad" music (Albinoni's *Adagio in G Minor* [6:32] or Barber's *Adagio pour Cordes* [5:33]). The two selections representing each type of music were assigned at random, except that subjects who completed both the encoding and the retrieval session in the same mood (happy or sad) were played one selection of the appropriate music during encoding and the alternative selection during retrieval. Every selection had been repeatedly recorded on a 45-min cassette tape; this saved us the trouble of having to stop and rewind the tape, which might have impeded development of the mood we were seeking to instill.

Five minutes following music onset, and every 5 min thereafter, the subjects rated their current levels of mood and arousal on a new matrix. The music played on while the subjects made these ratings; indeed, the music, once started, did not stop until the subjects were dismissed at the end of the session (hence the term *continuous* music technique).

To begin the task of generating and reading items, the subjects were required, at a minimum, to rate themselves as being either very happy or very sad. This was a stringent criterion, and as will be shown shortly, the subjects took considerable time to attain the critical level of mood. Nevertheless, we thought that by imposing this criterion, the odds of demonstrating mood dependent memory would be improved.

Two additional points concerning the mood criterion merit comment. First, the subjects were not told that the start of the generate/read task was contingent on their achieving a certain level of happiness or sadness. This contingency was kept confidential because we did not want the subjects to try to rush matters by rating their current mood as being more extreme than it actually was. Second, because the specific aim of the continuous music technique was to alter the subjects' level of mood, rather than their level of arousal, ratings of the latter played no part in determining whether or when the subjects were ready to begin the generate/read task. Thus, for example, a subject who had been listening to happy music, and thinking happy thoughts, was considered "ready" once he or she placed a check mark in *any* of the squares included in either of the two right-most columns of the mood/arousal matrix (see Figure 1). Ratings of arousal, therefore, were free—not forced—to covary with those of mood.

On reaching the requisite level of mood, the subjects were auditorily presented with a list composed of 32 triads, arranged in a random order. Of these 32 triads, 16 were of the form category name: category exemplar–target item to be read (e.g., *precious metals: silver–*GOLD), and 16 were of the form category name: category exemplar–initial letter of target item to be generated (e.g., *musical instruments: drum–G*). As noted earlier, the subjects were asked to state aloud both the generated and the read items (to control for overt activity), and their statements were transcribed by the experimenter (to monitor response accuracy). The rate of item presentation was subject-paced, and the total time taken was recorded by the experimenter.

Four lists were prepared for purposes of the generate/read task. These lists were structured in such a way that (a) all of the materials (i.e., category names, category exemplars, and target items) included in one pair of lists were different from those included in the other pair, and (b) within either pair, any item that was to be generated in one list was to be read in the other. Every subject was presented with one of the four lists; to guard against item-selection artifacts, the assignment of lists to subjects was counterbalanced across encoding/retrieval conditions.

On completing the generate/read task, all subjects marked a new mood/arousal matrix. Those who had been listening to happy music were then discharged, with a reminder to return in 2 days to complete a number of new and different tasks; they were not informed that their memory for the items they had generated or read this day would be tested at that time. Subjects who had been listening to sad music stayed longer—snacking on cookies and chatting with the experimenter while happy music played in the background—to ensure that their mood was at least neutral before they left the laboratory.

Retrieval session. Procedures involved in manipulating and measuring mood during the retrieval session were identical to those employed during the encoding session.

When subjects attained the appropriate level of mood (either very happy or very sad, irrespective of current level of arousal), they were asked to recall the items that they had either generated or read 2 days earlier. The subjects recalled the items aloud, in any order, and without benefit of any specific reminders or cues within a 5-min period.

Following item recall, the subjects again rated their current levels of mood and arousal. Subjects were then presented with a list containing all 32 target items from the encoding session, randomly interspersed with 32 previously unencoded items. Presentation was auditory and subject-paced. The subjects were instructed to respond aloud by saying "old" to any item they remembered having generated *or* read 2 days before and by saying "new" otherwise.

The test of item recognition was followed by one of source recognition. For this purpose, the subjects were again presented auditorily, and at their own pace, with every item they had identified as "old" moments before, and they were asked to specify aloud whether they had either generated or read each of these items 2 days earlier.

Subsequent to source recognition, subjects marked another mood/arousal matrix and were then debriefed. As was the case at encoding, subjects who completed the retrieval session while sad were not dismissed until their level of mood was not less than neutral.

Results

Mood and arousal ratings. Mean ratings of mood recorded on seven selected occasions over the course of the experiment appear in Table 1. These ratings were derived by translating the marks made by subjects along the horizontal axis of the mood/arousal matrix into an integer score ranging from −4 (*extremely sad*) through 0 (*neutral*) to 4 (*extremely happy*).

Three aspects of the data deserve comment. First, as would be expected, ratings of mood registered at the beginning of either the encoding or the retrieval session (ratings ESB and RSB) were about the same, regardless of whether subjects

Table 1
Mood as a Function of Rating Occasion and Encoding/Retrieval Condition

Encoding/retrieval condition	n	Rating occasion						
		ESB	BET	AET	RSB	BRc	ARc	ARn
Experiment 1								
Happy/happy	12	0.7	3.1	2.4	0.9	3.0	2.3	1.8
Happy/sad	12	0.3	3.1	1.9	0.8	−3.0	−1.8	−1.5
Sad/happy	12	0.5	−3.0	−1.3	0.7	3.0	1.8	1.6
Sad/sad	12	1.1	−3.0	−1.2	0.9	−3.0	−1.3	−0.8
Experiment 2								
Happy/happy	24	0.8	3.0	2.4	1.3	3.1	1.9	1.8
Happy/sad	24	1.1	3.1	2.1	0.9	−3.1	−1.9	−1.3
Experiment 3								
Happy/happy	24	1.4	3.0	1.6	0.8	3.0	1.7	1.4
Happy/sad	24	1.0	3.0	1.5	0.8	−3.0	−1.5	−0.3
Experiment 4								
Happy/happy	48	1.0	3.1	2.0	1.1	3.0	2.1	1.7
Happy/sad	48	1.0	3.0	2.0	1.0	−3.0	−1.8	−1.2

Note. Ratings derive from a 9-point scale ranging from 4 (*extremely happy*) through 0 (*neutral*) to −4 (*extremely sad*). n = number of subjects per mean rating. ESB = encoding session baseline; BET = before encoding task; AET = after encoding task; RSB = retrieval session baseline; BRc = before recall task; ARc = after recall task and before recognition tasks; ARn = after recognition tasks.

were soon to start listening to happy or to sad music. Second, in keeping with the level-of-mood criterion stated earlier, subjects rated themselves as being, at a minimum, either very happy (3 or higher) or very sad (−3 or lower) just before they began the task of generating and reading words (rating BET), and again, just before they began the task of recalling these words (rating BRc). Third, the impact of the mood manipulation declined over time and across tasks. To clarify, inspection of Table 1 reveals that the mean difference in mood ratings made by encode-happy and encode-sad subjects (ns = 24) was 6.1 points at the beginning of the generate/read task (rating BET: happy = 3.1; sad = −3.0), but only 3.5 points at its finish (rating AET: happy = 2.2; sad = −1.3). Similarly, the mean difference in mood between retrieve-happy and retrieve-sad subjects (ns = 24) decreased from 6.0 points before the test of item recall (rating BRc: happy = 3.0, sad = −3.0) to 3.7 points after recall and before item recognition (rating ARc: happy = 2.1, sad = −1.6) to 2.9 points after the final test of source recognition (rating ARn: happy = 1.7, sad = −1.2). These diminishing differences may reflect a regression to the mean, or they may be due to the encoding and retrieval tasks themselves. Whatever the reason, it should be noted that although the mood manipulation lost some of its effectiveness, it did not lose it all: Differences between happy and sad subjects in their mean ratings of mood were significant at all posttask assessments (viz., ratings AET, ARc, and ARn), $ts(46) > 9.51$, $ps < .01$. Thus, it appears that once a state of happiness or sadness had been induced, it remained relatively, though not absolutely, stable over time and across tasks. This is an important point, and we will return to it in the final discussion.

Regarding the time required to reach the critical level of mood, a 2 × 2 (Happy/Sad Mood × Encoding/Retrieval Session) analysis showed that subjects took significantly longer to become very sad than to become very happy (18.7 vs. 15.1 min), $F(1, 92) = 5.18$, $MS_e = 58.13$, $p < .05$. This may reflect the fact that subjects generally felt good at the start of the encoding and retrieval sessions: Mean mood ratings recorded at assessments ESB (0.7) and RSB (0.8) were both reliably greater than zero, $ts(47) > 2.91$, $ps < .01$. Neither the main effect of experimental session nor the interaction of sessions with type of mood had a significant influence on the time taken to attain the requisite mood level ($Fs > 1$).

Ratings of arousal are summarized in Table 2. These ratings were assigned in accordance with marks made by subjects along the vertical axis of the mood/arousal matrix and varied from 4 (*extremely high arousal*) through 0 (*neutral*) to −4 (*extremely low arousal*).

Comparison of Tables 1 and 2 suggests that the continuous music technique caused substantial changes not only in the subjects' ratings of mood—which is what the technique was specifically designed to do—but in their ratings of arousal as well. More to the point, happy subjects rated themselves as being more aroused than did sad subjects, and this was true for all ratings taken after the music had started (viz., ratings BET, AET, BRc, ARC, and ARn in Table 2), $ts(46) > 4.24$, $ps < .01$.

Table 2
Arousal as a Function of Rating Occasion and Encoding/Retrieval Condition

Encoding/retrieval condition	n	Rating occasion						
		ESB	BET	AET	RSB	BRc	ARc	ARn
Experiment 1								
Happy/happy	12	0.8	0.6	0.9	0.7	1.0	1.2	1.6
Happy/sad	12	−0.3	0.8	1.6	0.8	−0.8	−0.5	−0.4
Sad/happy	12	−0.3	−1.8	−1.0	−0.3	1.6	1.6	1.8
Sad/sad	12	0.3	−1.8	−0.4	0.3	−1.9	−0.4	0.2
Experiment 2								
Happy/happy	24	0.2	0.3	1.0	0.6	0.7	1.1	1.3
Happy/sad	24	−0.2	0.4	1.3	0.3	−1.4	−1.0	0.0
Experiment 3								
Happy/happy	24	0.1	0.2	0.7	1.0	0.7	1.3	1.5
Happy/sad	24	0.0	0.3	1.2	0.5	−1.7	−0.2	0.4
Experiment 4								
Happy/happy	48	0.2	0.1	0.7	0.6	0.3	0.7	1.0
Happy/sad	48	0.3	0.6	1.2	0.5	−1.8	−0.9	−0.3

Note. Ratings derive from a 9-point scale ranging from 4 (*extremely high arousal*) through 0 (*neutral*) to −4 (*extremely low arousal*). n = number of subjects per mean rating. ESB = encoding session baseline; BET = before encoding task; AET = after encoding task; RSB = retrieval session baseline; BRc = before recall task; ARc = after recall task and before recognition tasks; ARn = after recognition tasks.

Additional evidence of covariation between ratings of mood and arousal was obtained by means of correlational analysis. For purposes of this analysis, we computed the product-moment correlation between the seven pairs of mood and arousal ratings (i.e., ratings ESB though ARn in Tables 1 and 2) provided by every subject. Of the 48 individual-subject correlations, 25 were significantly positive, $rs(5) > .75$, and only 1 was significantly negative. The mean correlation was .50, which exceeds zero by a reliable margin, $t(47) = 6.80$, $p < .01$.

Item-generation errors and encoding-task time. The mean number of times that subjects generated a target item other than the one we had intended was 1.00 under both encode-happy and encode-sad conditions. Thus, the overall error rate in the generation task was a modest 6.3%. Items for which such errors occurred were omitted from the scoring of the item recall, item recognition, and source recognition data.

Because psychomotor slowing is a salient symptom of depression, whether naturally occurring or experimentally induced (Clark, 1983), it seems reasonable to suppose that subjects would take longer to complete the generate/read task if they were sad than if they were happy. This was indeed the case: Whereas sad subjects required an average of 270 s to generate or read the 32 target items, happy subjects performed the same encoding task in an average of 243 s, $t(44) = 2.66$, $p < .05$; task-performance times were not available for 2 subjects.

Item recall. The mean probabilities with which generated and read items were recalled in each encoding/retrieval condition are presented in Table 3. A 2 × 2 × 2 mixed-design analysis of the recall data disclosed (a) a main effect of item type, so that generated items were more likely to be recalled than were read items, $F(1, 44) = 88.55$, $MS_e = 0.008$, $p < .01$), (b) an interaction between encoding mood and retrieval mood, so that the probability of recall, averaged across item types, was greater when these moods matched than when they mismatched, $F(1, 44) = 18.67$, $MS_e = 0.008$, $p < .01$, and most important, (c) an interaction among encoding mood, retrieval mood, and item type, $F(1, 44) = 6.37$, $MS_e = 0.008$, $p < .05$. Further analysis of this complex interaction indicated that (a) the simple interaction of encoding and retrieval moods had a relatively strong influence on the recall of generated items, $F(1, 44) = 14.84$, $MS_e = 0.013$, $p < .01$, with performance in either of the two matched mood conditions surpassing performance in either of the two mismatched mood conditions by a reliable margin ($Fs > 4.25$, $ps < .05$), and (b) the simple interaction of encoding and retrieval moods had a relatively weak influence on the recall of read items, $F(1, 44) = 4.07$, $MS_e = 0.003$, $p = .05$, with the happy/happy condition holding a significant, but less than striking, advantage over both the happy/sad and sad/happy conditions ($Fs > 3.90$, $ps < .05$). Taken together, these results square with the possibility, posed in the opening section, that mood dependent effects in memory may be more pronounced for internal than for external events.

Table 3
Probability of Recall as a Function of Item Type and Encoding/Retrieval Condition

Encoding/retrieval condition	n	Item type G/1	R/1	G/3	R/3
	Experiment 1				
Happy/happy	12	.32	.09		
Happy/sad	12	.17	.04		
Sad/happy	12	.17	.05		
Sad/sad	12	.27	.06		
	Experiment 2				
Happy/happy	24	.22			.25
Happy/sad	24	.14			.23
	Experiment 3				
Happy/happy	24			.35	.19
Happy/sad	24			.29	.10
	Experiment 4				
Happy/happy	24	.11		.36	
Happy/sad	24	.07		.29	
Happy/happy	24		.04		.25
Happy/sad	24		.03		.24

Note. n = number of subjects per mean probability. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

Earlier it was remarked that ratings of mood were correlated with those of arousal, so that happy subjects were more highly aroused than were sad subjects. Nevertheless, inspection of the ratings supplied by subjects assigned to the mismatched mood conditions (viz., encode happy/retrieve sad or encode sad/retrieve happy) suggested that although all of these subjects showed a substantial shift in mood between the encoding and retrieval sessions (as was mandated by our methods), only some manifested a marked change in arousal as well. This being so, it is meaningful to ask: Does a change in both mood and arousal impair item recall more than does a change in mood alone, and if so, is this impairment specific to the recall of generated as opposed to read items?

To address these questions, we derived two scores for each of the 48 participants in the experiment. One of these scores—*change in mood*—was defined as the absolute difference between (a) the mean of the mood ratings registered immediately before and immediately after the generate/read task (i.e., ratings BET and AET in Table 1) and (b) the mean of the mood ratings registered immediately before and immediately after the test of item recall (i.e., ratings BRc and ARc). The second score—*change in arousal*—was defined the same way, except that ratings of arousal substituted for those of mood.

Next, we divided the 24 representatives of the mismatched mood conditions into two groups: *large change in mood/large change in arousal* (the 12 subjects with the largest arousal change scores) and *large change in mood/small change in arousal* (the 12 subjects with the smallest arousal change scores). For purposes of comparison, a third group—*small change in mood/small change in arousal*—was formed, which consisted of the 24 representatives of the matched-mood conditions (viz., encode happy/retrieve happy and encode sad/retrieve sad), most of whom displayed little, if any, disparity between sessions in their ratings of either mood or arousal.

Table 4
Probability of Recall as a Function of Item Type and Mood/Arousal Change

Mood/arousal change	n	G/1	R/1	G/3	R/3
Experiment 1					
Small (0.3)/small (0.8)	24	.30	.08		
Large (4.6)/small (1.3)	12	.23	.03		
Large (4.8)/large (4.1)	12	.11	.06		
Experiment 2					
Small (0.6)/small (1.3)	24	.22			.25
Large (5.0)/small (1.1)	12	.15			.25
Large (5.1)/large (3.1)	12	.12			.21
Experiment 3					
Small (0.1)/small (1.1)	24			.35	.19
Large (4.0)/small (0.9)	12			.33	.09
Large (4.9)/large (3.0)	12			.26	.11
Experiment 4					
Small (0.3)/small (0.7)	24	.11		.36	
Large (4.9)/small (1.1)	12	.09		.33	
Large (4.8)/large (3.8)	12	.05		.25	
Small (0.5)/small (1.2)	24		.04		.25
Large (4.8)/small (1.0)	12		.03		.23
Large (5.2)/large (3.9)	12		.03		.25

Note. Mean absolute differences between encoding and recall ratings of mood and of arousal appear in parentheses. n = number of subjects per mean probability. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

Table 4 presents the mean measures of mood and arousal change, plus the mean probabilities of generated and read item recall, for each of the three groups. With respect to the recall of generated items, analysis by planned comparisons produced a clear pattern of results: Subjects who experienced a large change in both mood and arousal were significantly outscored by those who experienced an equally large change in mood but a much smaller change in arousal, $F(1, 45) = 7.73$, $MS_e = 0.011$, $p < .01$, and the latter subjects were, in turn, marginally outscored by those who experienced a small change in mood as well as arousal ($F = 3.20$, $p < .10$). With respect to the recall of read items, however, the pattern is perplexing: The large/large condition produced slightly better—not significantly worse—performance than did the large/small condition, whereas performance was reliably poorer in the large/small than in the small/small condition, $F(1, 45) = 6.69$, $MS_e = 0.003$, $p = .01$.

Let us return to the questions posed earlier: Does a change in both mood and arousal impair item recall more than does a change in mood alone, and if so, is this impairment specific to the recall of generated as opposed to read items? In view of the results summarized above, the answer to both questions seems to be yes. This answer must, however, be regarded with caution, for two reasons: First, as already noted, mood/arousal change did have a significant, albeit unsystematic, impact on the recall of read items; second, the answer reflects the outcome of only one study. Data bearing on both the generality and reliability of the present results are provided by the next three studies in this series.

Item and source recognition. Results of the test of item recognition, in the form of mean probabilities with which generated, read, and new items were classified as "old," are shown in Table 5 in relation to encoding and retrieval moods. Though generated items were more often recognized than were read items, $F(1, 44) = 60.06$, $MS_e = 0.020$, $p < .01$, this difference was unaffected by encoding mood, retrieval mood, or their interaction (ps $> .10$). The influence of these variables on the false recognition of new items was also negligible (ps $> .10$).

Table 6 shows the results of the item recognition test in relation to three conditions of mood/arousal change. These conditions were formed in the same fashion as described earlier in connection with the item recall test (see Table 4), except that in the current context, change in mood (or change in arousal) was defined as the absolute difference between (a) the mean of the mood (or arousal) ratings registered immediately before and immediately after the generate/read task (i.e., ratings BET and AET in Tables 1 and 2) and (b) the mean of the mood (or arousal) ratings registered immediately before and immediately after recognition testing (i.e., ratings ARc and ARn). Planned comparisons revealed no reliable differences among any of the three conditions of mood/arousal change with respect to the recognition of generated, read, or new items (ps $> .10$).

To assess performance in the source recognition test, we determined for every subject both (a) the proportion of generated items correctly called "generated" and (b) the propor-

Table 5
Probability of Recognition as a Function of Item Type and Encoding/Retrieval Condition

Encoding/retrieval condition	n	G/1	R/1	G/3	R/3	New
Experiment 1						
Happy/happy	12	.90	.64			.28
Happy/sad	12	.79	.63			.28
Sad/happy	12	.85	.62			.32
Sad/sad	12	.88	.64			.29
Experiment 2						
Happy/happy	24	.89			.84	.19
Happy/sad	24	.88			.89	.17
Experiment 3						
Happy/happy	24			.94	.76	.07
Happy/sad	24			.93	.77	.09
Experiment 4						
Happy/happy	24	.85		.94		.12
Happy/sad	24	.80		.94		.12
Happy/happy	24		.57		.91	.20
Happy/sad	24		.53		.82	.19

Note. n = number of subjects per mean probability. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

tion of new items incorrectly called "generated." These two proportions were then translated into a d' score, signifying the subject's ability to discriminate the source of target items that had been generated. An analogous d' score, indexing the source discriminability of read items, was also computed for each subject on the basis of the proportions of read and new items that were identified as having been read.

Mean d' scores representing the source discriminability of generated and read items are formatted as a function of encoding/retrieval condition in Table 7 and as a function of mood/arousal change in Table 8. In general, subjects were more accurate in discriminating the source of generated than of read items, $F(1, 44) = 7.14$, $MS_e = 1.01$, $p = .01$, and this was true irrespective of encoding mood, retrieval mood, or their interaction ($ps > .10$). Similarly, there were no significant differences among the three conditions of mood/arousal change in terms of the source discriminability of either generated or read items ($ps > .10$). Thus, the data displayed in Tables 7 and 8, together with those found in Tables 5 and 6, indicate that alterations of mood or arousal do not diminish recognition of either the occurrence or the origin of target events, whether internal or external.

Experiment 2

Aims and Method

The first study showed that a shift in mood state impaired the recall of generated items to a greater extent than that of read items. Is this because internally generated events are particularly prone to mood dependence, as was suggested at the outset, or because the level of read-item recall was too low (an average of about 6%, as indicated in Table 3) to allow as large a mood dependent effect to emerge in the recall of read items as was evident in the recall of generated items?

To find out, we conducted a second study that was identical to the first in all respects save three. First, in an effort to enhance the recall of read items relative to that of generated items, subjects in the second study were presented, at encoding, with 16 items to read three times each, and with 16 items to generate one time each. Repetitions of the same item to read were spaced at random throughout the presentation list, and any item that was thrice read by one subject was once generated by another. Both the manner in which read and generated items were presented at encoding, and the nature of the items themselves, were the same as previously described.

Second, the Alpha Chamber, in which subjects in the first study sat while they listened to music, was returned to the colleague from whom we had borrowed it, and was replaced with a more conventional lounge chair. Music was played to participants in the second study via two high-quality stereo speakers, which were situated on either side of this chair.

Third, because the mood dependent effect found in the first study was symmetric in form—that is, encode happy/retrieve sad subjects performed as poorly in recall as did their encode sad/retrieve happy counterparts (see Table 3 for the relevant data, and Eich, 1989, for a theoretical discussion of symmetric and asymmetric forms of mood dependence)—the design of the second study was simplified, such that some subjects completed both the encoding and retrieval sessions in a happy mood, whereas others were happy during encoding and sad during retrieval. Each of these two encoding/retrieval conditions (matched vs. mismatched mood) was represented by 24 randomly assigned subjects, all of whom were UBC undergraduates.

Table 6
Probability of Recognition as a Function of Item Type and Mood/Arousal Change

Mood/arousal change	n	G/1	R/1	G/3	R/3	New
		Experiment 1				
Small (0.9)/small (1.0)	24	.89	.64			.28
Large (3.7)/small (0.9)	12	.82	.60			.29
Large (4.3)/large (4.1)	12	.82	.65			.31
		Experiment 2				
Small (1.2)/small (1.5)	24	.89			.84	.19
Large (4.0)/small (0.6)	12	.91			.91	.17
Large (4.4)/large (3.0)	12	.84			.88	.16
		Experiment 3				
Small (0.8)/small (1.4)	24			.94	.76	.07
Large (2.6)/small (0.6)	12			.91	.73	.12
Large (3.6)/large (2.3)	12			.94	.80	.07
		Experiment 4				
Small (0.8)/small (0.7)	24	.85		.94		.12
Large (4.0)/small (0.8)	12	.81		.96		.10
Large (3.8)/large (3.1)	12	.79		.93		.14
Small (0.9)/small (1.5)	24		.57		.91	.20
Large (3.8)/small (0.6)	12		.56		.85	.18
Large (4.3)/large (3.2)	12		.50		.79	.21

Note. Mean absolute differences between encoding and recognition ratings of mood and of arousal appear in parentheses. n = number of subjects per mean probability. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

Results

Mood and arousal ratings. Results pertaining to ratings of mood are presented in Table 1. Inspection of these results indicates that the subjects were slightly happy at the start of the encoding session (mean ESB rating = 1.0, $n = 48$), very happy when they began the task of generating and (repeatedly) reading words (mean BET rating = 3.1), and moderately happy by the time they had finished (mean AET rating = 2.3). This pattern of encoding session ratings parallels the one found for happy subjects in the first study.

Though matched and mismatched mood subjects were equivalently happy at the beginning of the retrieval session (mean RSB rating = 1.1, $n = 48$), their ratings of mood differed markedly thereafter. On average, the former subjects outscored the latter by 6.2 points on the 9-point happiness/sadness scale prior to item recall (rating BRc), by 3.8 points prior to item recognition (rating ARc), and by 3.1 points at the conclusion of source recognition (rating ARn). All of these differences were significant, $ts(46) > 10.05$, $ps < .01$, which reaffirms the conclusion, reached in the first study, that the continuous music technique produced statistically reliable and relatively stable modifications of mood.

Data relating to ratings of arousal appear in Table 2. Comparison of these data with the mood ratings summarized

Table 7
Discriminability (d') of Source (Experiments 1–3) or Frequency (Experiment 4) as a Function of Item Type and Encoding/Retrieval Condition

Encoding/retrieval condition	n	Item type G/1	R/1	G/3	R/3
	Experiment 1				
Happy/happy	12	2.63	1.94		
Happy/sad	12	2.47	1.81		
Sad/happy	12	2.44	1.69		
Sad/sad	12	2.06	1.96		
	Experiment 2				
Happy/happy	24	3.04			2.77
Happy/sad	24	3.01			2.85
	Experiment 3				
Happy/happy	24			3.51	3.38
Happy/sad	24			3.31	3.30
	Experiment 4				
Happy/happy	24	2.84		2.99	
Happy/sad	24	3.08		3.18	
Happy/happy	24		2.17		2.58
Happy/sad	24		2.48		2.21

Note. n = number of subjects per mean d'. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

in Table 1 suggests that, as was the case in Experiment 1, ratings of arousal tended to covary with those of mood. The mean correlation between the seven principal pairs of ratings made by subjects in Experiment 2 was .34, which significantly exceeds zero, $t(47) = 4.77$, $p < .01$.

Item recall. The average error rate in the generation task was 5.6%; items contributing to this average were excluded from analyses of recall and recognition performance.

The mean probabilities with which once-generated and thrice-read items were recalled are presented in Table 3. The simple effects of encoding/retrieval condition and item type were both significant, indicating an advantage of matched over mismatched moods, $F(1, 46) = 5.68$, $MS_e = 0.010$, $p < .05$, as well as an advantage of thrice-read over once-generated items, $F(1, 46) = 9.11$, $MS_e = 0.010$, $p < .01$. The interaction between these effects was marginally reliable, $F(1, 46) = 2.75$, $MS_e = 0.010$, $p = .10$. Relative to subjects whose encoding and retrieval moods matched, those whose moods mismatched recalled fewer generated items, $F(1, 46) = 8.93$, $MS_e = 0.009$, $p < .01$, but the same number of read items ($F < 1$).

Table 4 recasts the recall data in relation to mood/arousal change. Though the three conditions comprising this variable yielded comparable levels of read item recall ($ps > .10$), they differed with respect to the recall of generated items. Specifically, subjects who experienced a small change between sessions in mood as well as arousal recalled more generated items than did subjects who experienced either a large change in both mood and arousal, $F(1, 45) = 8.76$, $MS_e = 0.009$, $p < .01$, or a large change in mood alone ($F = 3.65$, $p < .10$). The levels of generated item recall attained in the large/large and large/small conditions were statistically indistinguishable ($F < 1$).

Item and source recognition. Data derived from the tests of item and source recognition are summarized with respect to encoding/retrieval condition in Tables 5 and 7, and with respect to mood/arousal change in Tables 6 and 8. The methods by which these data were scored and analyzed were similar to those discussed earlier in reference to Experiment 1. No simple or interactive effects (involving encoding/retrieval condition or item type) and no planned comparisons (involving the three conditions of mood/arousal change) were reliable ($ps > .10$).

Experiment 3

Aims and Method

The results of the second study were similar to those of the first in several respects and different in others. As to their similarities, both studies showed that (a) the continuous music technique produced strong and stable modifications of mood, which were associated with alterations in arousal, (b) experimental manipulation of the match

Table 8
Discriminability (d') of Source (Experiments 1–3) or Frequency (Experiment 4) as a Function of Item Type and Mood/Arousal Change

Mood/arousal change	n	Item type G/1	R/1	G/3	R/3
	Experiment 1				
Small (0.9)/small (1.0)	24	2.34	1.95		
Large (3.7)/small (0.9)	12	2.44	1.83		
Large (4.3)/large (4.1)	12	2.47	1.67		
	Experiment 2				
Small (1.2)/small (1.5)	24	3.04			2.77
Large (4.0)/small (0.6)	12	3.11			2.73
Large (4.4)/large (3.0)	12	2.92			2.97
	Experiment 3				
Small (0.8)/small (1.4)	24			3.51	3.38
Large (2.6)/small (0.6)	12			3.13	3.17
Large (3.6)/large (2.3)	12			3.50	3.43
	Experiment 4				
Small (0.8)/small (0.7)	24	2.84		2.99	
Large (4.0)/small (0.8)	12	3.07		3.33	
Large (3.8)/large (3.1)	12	3.09		3.03	
Small (0.9)/small (1.5)	24		2.17		2.58
Large (3.8)/small (0.6)	12		2.67		2.14
Large (4.3)/large (3.2)	12		2.28		2.28

Note. Mean absolute differences between encoding and recognition ratings of mood and of arousal appear in parentheses. n = number of subjects per mean d'. G/1 = once generated; R/1 = once read; G/3 = thrice generated; R/3 = thrice read.

between encoding and retrieval moods had a more pronounced effect on the recall of generated in comparison with read items, (c) significantly fewer generated items were recalled after a large rather than a small shift, from encoding to retrieval, in arousal in addition to mood, and (d) a shift in mood, arousal, or both did not deter recognition of either the occurrence or the origin of a target item, regardless of whether the item had been generated or read.

As to differences between Experiments 1 and 2, subjects in Experiment 1 who experienced a sizable shift in mood, but only a slight shift in arousal, recalled significantly more generated items than did subjects who experienced a substantial shift in both measures. Although a trend in the same direction was seen in Experiment 2, it was not statistically reliable. A second, more striking difference is that whereas a small but significant mood dependent effect was evident in the recall of items read once in Experiment 1, no such effect was found in the recall of items read thrice in Experiment 2. If it is the case, as these and related results (see Eich & Birnbaum, 1982) seem to suggest, that repetition diminishes mood dependent effects in memory, then little or no evidence of mood dependence should emerge in the recall of target items that have been repeatedly read *or* repeatedly generated.

This inference provided the rationale for Experiment 3. During the encoding session of this experiment, every subject read 16 target items three times each and generated 16 others three times each as well. Repetitions of the same read or generated item were distributed at random throughout the presentation list. Forty-eight UBC undergraduates served as subjects in this experiment, of whom 24 were randomly assigned to the encode happy/retrieve happy condition and 24 to the encode happy/retrieve sad condition. Except for the fact that participants in Experiment 3 had three opportunities instead of one to generate certain target items, they were treated in exactly the same manner as subjects in Experiment 2.

Results

Mood and arousal ratings. Mean ratings of mood and arousal are shown in Tables 1 and 2, respectively. These averages accord well with those found in the first two studies, and so rather than discuss the new data in detail, we refer the reader to earlier accounts of mood/arousal ratings.

Item recall. Unless a given item was correctly generated three times by a given subject, it was omitted from all analyses of recall and recognition performance. On average, about 8.2% of the thrice-generated items were omitted on this basis.

Table 3 displays the recall data in relation to encoding/retrieval condition and item type. Both of these simple effects were significant: More items were recalled when encoding and retrieval moods matched than when they mismatched, $F(1, 46) = 11.64$, $MS_e = 0.011$, $p < .01$, and more thrice-generated items were recalled than thrice-read items, $F(1, 46) = 67.05$, $MS_e = 0.012$, $p < .01$. In clear contrast to the results of the first two studies, the interaction between encoding/retrieval condition and item type did not approach statistical significance ($F < 1$).

In Table 4, the recall data are reformatted as a function of mood/arousal change. Relative to subjects who experienced a small shift in mood as well as arousal, those who experienced a large shift in both mood and arousal recalled fewer thrice-generated items, $F(1, 45) = 4.65$, $MS_e = 0.015$, $p < .05$, and fewer thrice-read items, $F(1, 45) = 6.25$, $MS_e = 0.007$, $p < .05$. Recall of repeatedly read items was also poorer in the large/small than in the small/small condition ($F = 10.17$, $p < .01$), but there was no difference between these conditions in the recall of repeatedly generated items.

Item and source recognition. Although recognition of the occurrence of thrice-generated items exceeded that of thrice-read items (see Table 5), $F(1, 46) = 44.98$, $MS_e = 0.016$, $p < .01$, recognition of the origin of the former type of item was neither better nor worse than that of the latter type (see Table 7; $F < 1$). Both the probability with which generated, read, and new items were identified as "old" and the discriminability of the source of the target items remained constant regardless of encoding/retrieval condition (Tables 5 and 7; $ps > .10$) or mood/arousal change (Tables 6 and 8; $ps > .10$).

Experiment 4

Aims and Method

The results of the third study threw us a curve. Not only was a robust mood dependent effect evident in the recall of thrice-read items, but the effect was, if anything, stronger than that seen in the recall of thrice-generated items. This pattern represents a radical departure from the one observed in the first two studies, which is strange considering the many methodological similarities that all three studies shared.

Short on insight into why the results of Experiment 3 turned out as they did, we undertook Experiment 4. Subjects in this study were 96 UBC undergraduates, of whom 24 were randomly assigned to each of four groups defined by the crossing of (a) encoding/retrieval condition (happy/happy vs. happy/sad) and (b) type of encoding task (item generation vs. item reading). A third variable, item presentation frequency (once vs. thrice), was within subjects. Within each encoding/retrieval condition, then, half of the subjects generated 16 target items one-time each and generated 16 additional items three-times each, whereas the other half read 16 targets once and 16 others thrice.

Because subjects in Experiment 4 either generated or read targets items during the encoding session (rather than doing some of both, as was the case in the earlier experiments), the test of source recognition was replaced by one of frequency recognition. Thus, during the retrieval session of the fourth study, subjects initially completed a test of item recall and then a test of old/new recognition (identical to the one administered in the prior studies). Afterwards, the subjects were reminded of every item they had identified as "old" and were asked to indicate whether that item had been presented (for purposes of generation or reading) once or thrice during the encoding session.

In all other important respects, the fourth study matched the methods of the first three. In particular, Experiment 4 employed the same lists of target items (with their corresponding category name and category exemplar cues), the same retention interval (2 days separating encoding and retrieval), and the same techniques for mood induction and measurement as those involved in the three previous studies.

Results

Mood and arousal ratings. As is apparent in Tables 1 and 2, ratings of mood and arousal registered in the fourth study corresponded closely to those recorded in the first three.

Item recall. Owing to errors of item generation, 7.9% of the once-generated and 11.9% of the thrice-generated items were excluded from all analyses of memory performance.

Tables 3 and 4 feature the mean probabilities of item recall as functions of encoding/retrieval condition and mood/

arousal change, respectively. In reference to Table 3, an advantage accrued to the recall of (a) generated over read items, $F(1, 92) = 20.90$, $MS_e = 0.010$, $p < .01$, (b) thrice-presented over once-presented targets, $F(1, 92) = 275.87$, $MS_e = 0.009$, $p < .01$, and (c) matched over mismatched mood conditions, $F(1, 92) = 5.52$, $MS_e = 0.010$, $p < .05$. None of the double interactions was reliable, and neither was the triple ($Fs < 2.13$, $ps > .10$).

In reference to Table 4, planned comparisons showed that subjects who experienced a large as opposed to a small shift in both mood and arousal recalled fewer once-generated items, $F(1, 90) = 5.74$, $MS_e = 0.005$, $p < .05$, and fewer thrice-generated items, $F(1, 90) = 7.19$, $MS_e = 0.013$, $p < .01$. All other comparisons were nonsignificant ($Fs < 2.48$, $ps > .10$).

Item and frequency recognition. With respect to Table 5, a 2 × 2 × 2 (Encoding/Retrieval Mood × Item Type × Item Frequency) analysis revealed that the probability of a recognition hit was greater for generated than for read items, $F(1, 92) = 52.85$, $MS_e = 0.029$, $p < .01$, and greater for thrice-presented than for once-presented targets, $F(1, 92) = 132.72$, $MS_e = 0.017$, $p < .01$. The interaction between these variables was significant, $F(1, 92) = 27.54$, $MS_e = 0.017$, $p < .01$, indicating that the advantage in recognition of generated over read items was magnified if both types of items had been presented once rather than thrice. Oddly, the overall recognition-hit rate attained by subjects whose encoding and retrieval moods matched was marginally above that achieved by subjects who shifted from happiness to sadness, $F(1, 92) = 3.63$, $MS_e = 0.029$, $p < .10$. It is unclear why this study should show some semblance of mood dependent recognition when none of the others did. Reassuringly, the overall false-positive rate was unaffected by encoding/retrieval condition ($F < 1$), though it was higher if subjects read rather than generated target items during the encoding session, $F(1, 92) = 7.51$, $MS_e = 0.020$, $p < .01$).

With respect to Table 6, analysis by planned comparisons disclosed only one significant difference: Subjects who experienced a large change in mood as well as arousal recognized fewer thrice-read items than did subjects who experienced a small change in both mood and arousal, $F(1, 90) = 7.44$, $MS_e = 0.016$, $p < .01$. Inasmuch as no comparable difference was detected in either Experiment 2 or Experiment 3, the significance of the present result is probably illusory.

Mean d' scores derived from the test of frequency recognition are arranged in relation to encoding/retrieval condition and mood/arousal change in Tables 7 and 8, respectively. Two d' scores were computed for every subject, one of which was based on (a) the proportion of once-generated or once-read items that the subject correctly classified as having been presented one time and (b) the proportion of new items that the subject incorrectly classified as having been once presented. The second d' score reflected the frequency discrimination of thrice-generated or thrice-read items, and was based on the proportions of thrice-presented and new items that were identified as having been presented three times.

Relative to subjects who read target items during the encoding session, those who generated them were more accurate in discriminating the frequency with which the targets had been presented, $F(1, 92) = 15.80$, $MS_e = 1.35$, $p < .01$. Accuracy of discrimination was not affected by item frequency, encoding/retrieval condition, or their interaction ($Fs < 1$), and there were no differences among the three conditions of mood/arousal change ($ps > .10$).

General Discussion

We will frame the discussion of this series of studies around four principal points. First, the present results provide good—not great—support for the proposition that events that are generated through internal mental processes such as reasoning, imagination, and thought may be more colored by or connected to one's current mood than are those that emanate from external sources, thereby leading to a larger loss of memory for internal than for external events after a shift in mood state. We say the support is "not great" for the simple reason that, from a statistical standpoint, only the first two of the four studies showed that the recall of generated items was significantly more impaired by a shift in mood than was the recall of read items. What is more, in one of the studies (Experiment 3), the advantage in recall of matched over mismatched mood conditions was slightly greater for read than for generated items—the antithesis of the anticipated pattern of results.

Still, a case for "good" support can be made, especially if the four studies are considered collectively rather than separately. To this end, we reapportioned the recall data in such a way that each of the 144 subjects who served in either Experiment 1, 2, or 3 contributed two scores (one reflecting the recall of once- or thrice-generated items, and the other, the recall of once- or thrice-read items), whereas each of the 96 participants in Experiment 4 supplied a single score (the mean probability of recall of either once- and thrice-generated or once- and thrice-read items). The total number of observations, then, was 384, of which half pertained to the recall of generated items and half to the recall of read items. Within each type of item, half of the observations corresponded to matched mood conditions (either encode happy/retrieve happy or encode sad/retrieve sad), and half to mismatched mood conditions (encode happy/retrieve sad or the reverse).

The aggregate data appear in Table 9. Analysis of these data showed that the simple effects of encoding/retrieval condition and item type were significant, $Fs(1, 380) > 27.57$, $MS_e = 0.012$, $ps < .01$, as was their interaction, $F(1, 380) = 3.76$, p

Table 9
Probability of Recall as a Function of Item Type and Encoding/Retrieval Condition (Experiments 1–4 Combined)

Encoding/retrieval condition	n	Item type: Generated	Read
Matched moods	96	.28	.16
Mismatched moods	96	.19	.13

Note. n = number of subjects per mean probability. Matched moods correspond to encode happy/retrieve happy and encode sad/retrieve sad conditions; mismatched moods correspond to encode happy/retrieve sad and encode sad/retrieve happy conditions.

= .05. Whereas matched and mismatched mood conditions differed by 3% in the recall of read items, $F(1, 380) = 5.49$, $p < .05$, they differed by 9% in the recall of generated items, $F(1, 380) = 25.85$, $p < .01$. Neither of these differences is particularly impressive, but they do imply that internal events are more apt to be rendered unrecallable in the transition from one mood state to another.

Although the source of an event seems to play a part in the occurrence of mood dependent memory, it is not the only factor that matters, or even the one that matters most. This, the second main point, is plainly illustrated by the fact that of the several tests of retention—free recall, item recognition, and source or frequency discrimination—that were taken by subjects in this research, only free recall afforded consistent evidence of mood dependence. Perhaps the more one must rely on internal resources, rather than external aids, to generate both the target events at encoding and the cues required for their retrieval, the more likely is one's memory for these events to be mood dependent. In any case, the finding that deserves emphasis is that mood dependent effects are determined not only by the source of the target events (internal vs. external), but also by the manner in which retention is measured.

The third issue of interest has to do with the strength and stability of the moods in which encoding and retrieval take place. Intuition suggests that in order to show that a shift in mood state significantly impairs memory, two conditions must be satisfied. First, the shift must be substantial. Just *how* substantial is an open empirical question, but based on the present results, it seems that the shift from a *very happy* to a *very sad* state (or vice versa) is sufficient to reduce the recall of internally generated events. It is possible, indeed probable, that a less severe shift—say from feeling *slightly happy* to feeling *slightly sad*—would have no significant impact on the recall of events, even those that had been produced internally. Second, the mood that exists at the start of the encoding or retrieval task must still be present at its end. The importance of this proviso is implied by the results of a small pilot study, in which we sought to manipulate mood using the Velten (1968) technique, whereby subjects are asked to internalize the affect suggested by a series of elation or depression statements. Subjects in this study (16 UBC undergraduates) undertook a generate/read task, similar to the one used in Experiment 1, after the induction of either an elated or a depressed mood, and were tested for free recall 30 min later in either the original or the opposite state.

Relative to subjects who received the depression induction, those exposed to statements suggesting elation rated themselves as being significantly happier immediately *before* they began both the generate/read and the item recall tasks (mean ratings of 5.2 vs. 3.0 on a 7-point scale), $t(14) = 5.08$, $p < .01$. However, there was no reliable difference between elation and depression inductions in happiness ratings registered immediately *after* the performance of either task (means of 4.8 vs. 4.6; $t < 1$). Thus, the effects of the Velten technique on self-reported happiness, though strong initially, proved to be short lived (for similar results, see Isen & Gorgoglione, 1983). Perhaps for this reason, our pilot study provided no evidence of mood dependence in the recall of either generated or read items (mean probabilities of .33 and .10 under matched mood conditions; .39 and .09 under mismatched mood conditions). Because the mood effects engendered by the continuous music technique are reasonably stable over time and across tasks (see Table 1), and because far more people are responsive to the musical than to the Velten induction (see Clark, 1983), the former appears to be the preferred method for manipulating mood in future studies of mood dependent memory.

Up to now, the focus of discussion has been on the state dependent effects of moods on memory for internal as opposed to external events. The last point we wish to pursue concerns the memorial consequences of changes in arousal. As a prelude to further remarks, consider Table 10, in which the recall results of all four studies have been combined (following the same strategy outlined in connection with Table 9) and classified according to three levels of mood/arousal change. Analysis of the combined results revealed that in comparison with subjects whose encoding and retrieval states were well matched with respect to both mood and arousal, those who experienced a large change in mood but only a small change in arousal recalled fewer generated items, $F(1, 189) = 5.07$, $MS_e = 0.013$, $p < .05$, and fewer read items, $F(1, 189) = 5.07$, $MS_e = 0.011$, $p < .05$. These differences denote "pure" mood dependence and argue against the possibility, suggested by Clark and her associates (1983), that mood dependent effects are mediated by alterations in arousal. However, in partial support of Clark et al.'s (1983) claim that arousal cues arousal-related material in memory, a large change in both mood and arousal resulted in poorer recall of generated items than did a large change in mood alone, $F(1, 189) = 9.19$, $MS_e = 0.013$, $p < .01$, though such was not the case for read items ($F < 1$). Thus, the most striking difference detected in the present research was in the levels of generated-item recall achieved by subjects who experienced a small rather than a large alteration in arousal as well as mood, $F(1, 189) = 33.09$, $MS_e = 0.013$, $p < .01$. It is proper to describe this difference as demonstrating *mood* dependent memory for internal events, or is *mood-and-arousal* dependence a more appropriate account?

The answer depends on how one views the relation between mood and arousal. If mood is considered to be exclusive of arousal—so that the former corresponds to a certain degree of happiness or sadness, and the latter, to a certain degree of liveliness or languor—then mood-and-arousal dependence is clearly the characterization of choice. If, however, mood is

Table 10
Probability of Recall as a Function of Item Type and Mood/Arousal Change (Experiments 1–4 Combined)

Mood/arousal change	n	Item type: Generated	Read
Small (0.4)/small (1.0)	96	.28	.16
Large (4.6)/small (1.1)	48	.23	.12
Large (5.0)/large (3.5)	48	.16	.13

Note. Mean absolute differences between encoding and recall ratings of mood and of arousal appear in parentheses. n = number of subjects per mean probability.

considered to be inclusive of arousal—so that mood corresponds to a subjective state specified by a particular level of both happiness/sadness and activity/inactivity—then the simpler description, mood dependent memory, is sufficient.

Although on first impression the difference between the "exclusive" and "inclusive" views of mood and arousal seems to be simply a matter of semantics, there is more to it than that. In particular, Russell (1980) has assembled strong empirical support for a circumplex model of affect, in which different moods, such as delight or distress, are represented as points on a circle in two-dimensional bipolar space, the axes of which are pleasure/displeasure (which parallels the present distinction between happiness/sadness) and high/low arousal (or activity/inactivity). Stated in terms of Russell's model, the present research suggests that whereas a shift along only the pleasure dimension of mood results in a modest reduction in the recall of internally generated events, a shift along both the pleasure and arousal dimensions produces a much more marked impairment. Future research may therefore profit from the development of new experimental methods for manipulating both pleasure and arousal, for it is possible that through the use of such methods, the prospects of demonstrating mood dependent memory would be improved.

References

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99*, 229–246.

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

Bower, G. H. (1985, September). *Review of research on mood and memory.* Paper presented at the Symposium on Affect and Cognition, British Psychological Society, Oxford, England.

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

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

Eich, E. (1989). Theoretical issues in state dependent memory. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 331–354). Hillsdale, NJ: Erlbaum.

Eich, E., & Birnbaum, I. M. (1982). Repetition, cuing, and state dependent memory. *Memory & Cognition, 10*, 103–114.

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

Gerrig, R. J., & Bower, G. H. (1982). Emotional influences on word recognition. *Bulletin of the Psychonomic Society, 19*, 197–200.

Goodwin, D. W., Powell, B., Bremer, D., Hoine, H., & Stern, J. (1969). Alcohol and recall: State dependent effects in man. *Science, 163*, 1358–1360.

Isen, A. M. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–230). Hillsdale, NJ: Erlbaum.

Isen, A. M., & Gorgoglione, J. M. (1983). Some effects of four affect-induction procedures. *Personality and Social Psychology Bulletin, 9*, 136–143.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*, 67–85.

Johnson, M. K., Taylor, T. H., & Raye, C. L. (1977). Fact and fantasy: The effects of internally generated events on the apparent frequency of externally generated events. *Memory & Cognition, 5*, 116–122.

Persson, L. O., Sjoberg, L., & Svensson, E. (1980). Mood effects of alcohol. *Psychopharmacology, 68*, 295–299.

Raye, C. L., Johnson, M. K., & Taylor, T. H. (1980). Is there something special about memory for internally generated information? *Memory & Cognition, 8*, 141–148.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (in press). The affect grid: A new single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology.*

Slamecka, N. J., & Fevreiski, J. (1983). The generation effect when generation fails. *Journal of Verbal Learning and Verbal Behavior, 22*, 153–163.

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

Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behaviour Research and Therapy, 5*, 3–25.

Velten, E. (1968). A laboratory task for induction of mood states. *Behavior Research and Therapy, 6*, 473–482.

Weingartner, H. (1978). Human state dependent learning. In B. T. Ho, D. W. Richards, & D. L. Chute (Eds.), *Drug discrimination and state dependent learning* (pp. 361–382). New York: Academic Press.

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

Received June 30, 1986
Revision received August 15, 1988
Accepted August 26, 1988 ■